SOCIAL PSYCHOLOGY

SOCIAL PSYCHOLOGY

HANDBOOK OF BASIC PRINCIPLES

SECOND EDITION

edited by
ARIE W. KRUGLANSKI
E. TORY HIGGINS

Printed in the United States of America

This book is printed on acid-free paper.

Last digit is print number: 9 8 7 6 5 4 3 2 1

About the Editors

Arie W. Kruglanski, PhD, is Distinguished University Professor of Psychology at the University of Maryland–College Park. He is a recipient of the Donald T. Campbell Award for Outstanding Contributions to Social Psychology from the Society for Personality and Social Psychology, the Distinguished Scientific Contribution Award from the Society of Experimental Social Psychology, and the Senior Career Award from the National Institute of Mental Health. Dr. Kruglanski is widely recognized for his distinguished research contributions, disseminated in over 200 articles and chapters, which focus on how people form judgments, beliefs, impressions, and attitudes, and the consequences for their interpersonal relations, group interactions, and intergroup relations. He has served as editor of the *Journal of Social and Personality Psychology: Attitudes and Social Cognition* and of the *Personality and Social Psychology Bulletin*. He is now a codirector of the START National Center for the Study of Terrorism and the Response to Terrorism at the University of Maryland.

E. Tory Higgins, PhD, is the Stanley Schachter Professor of Psychology, Professor of Business, and Director of the Motivation Science Center at Columbia University. He has received a MERIT Award from the National Institute of Mental Health, the Thomas M. Ostrom Award in Social Cognition, the Donald T. Campbell Award for Outstanding Contributions to Social Psychology from the Society for Personality and Social Psychology, and the Lifetime Contribution Award from the International Society for Self and Identity. Dr. Higgins has also received the Distinguished Scientist Award from the Society of Experimental Social Psychology, the William James Fellow Award for Distinguished Achievements in Psychological Science from the American Psychological Society, and the American Psychological Association Award for Distinguished Scientific Contributions. He is a Fellow of the American Academy of Arts and Sciences and a recipient of Columbia University's Presidential Award for Outstanding Teaching.

Contributors

Susan M. Andersen, PhD, Department of Psychology, New York University, New York, New York

Roy F. Baumeister, PhD, Department of Psychology, Florida State University, Tallahassee, Florida

Bianca Beersma, PhD, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Jennifer Berdahl, PhD, Rotman School of Management, University of Toronto, Toronto, Ontario, Canada

Monica Biernat, PhD, Department of Psychology, University of Kansas, Lawrence, Kansas

Irene V. Blair, PhD, Department of Psychology, University of Colorado, Boulder, Colorado

Jim Blascovich, PhD, Department of Psychology, University of California, Santa Barbara, California

Marilynn B. Brewer, PhD, Department of Psychology, Ohio State University, Columbus, Ohio

Pablo Briñol, PhD, Department of Social Psychology, Universidad Autónoma de Madrid, Madrid, Spain

Linnda R. Caporael, PhD, Department of Science and Technology Studies, Rensselaer Polytechnic Institute, Troy, New York

Chi-yue Chiu, PhD, Department of Psychology, University of Illinois at Urbana–Champaign, Champaign, Illinois

Gerald L. Clore, PhD, Department of Psychology, University of Virginia, Charlottesville, Virginia

Philip R. Costanzo, PhD, Department of Psychology, Duke University, Durham, North Carolina

David De Cremer, PhD, Department of Economic and Social Psychology, Tilburg University, Tilburg, The Netherlands

Carsten K. W. De Dreu, PhD, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Roland Deutsch, PhD, Department of Psychology, University of Würzburg, Würzburg, Germany

David Dunning, PhD, Department of Psychology, Cornell University, Ithaca, New York

Scott Eidelman, PhD, Department of Psychology, University of Maine, Orono, Maine

Melissa J. Ferguson, PhD, Department of Psychology, Cornell University, Ithaca, New York

Klaus Fiedler, PhD, Department of Psychology, University of Heidelberg, Heidelberg, Germany

Ayelet Fishbach, PhD, Graduate School of Business, University of Chicago, Chicago, Illinois

Susan T. Fiske, PhD, Department of Psychology, Princeton University, Princeton, New Jersey

Jens Förster, PhD, Department of Psychology, International University Bremen, Bremen, Germany

Geoffrey Haddock, PhD, School of Psychology, Cardiff University, Cardiff, Wales

E. Tory Higgins, PhD, Department of Psychology, Columbia University, New York, New York

Denis Hilton, PhD, Department of Social Psychology, University of Toulouse, Toulouse, France

Michael A. Hogg, PhD, School of Behavioral and Organizational Sciences, Claremont Graduate University, Claremont, California

Ying-yi Hong, PhD, Department of Psychology, University of Illinois at Urbana–Champaign, Champaign, Illinois

Gita Venkataramani Johar, PhD, Department of Marketing, Columbia University Business School, Columbia University, New York, New York

Eric J. Johnson, PhD, Department of Marketing, Columbia University Business School, Columbia University, New York, New York

Neil P. Jones, MA, Department of Psychology, Duke University, Durham, North Carolina

John T. Jost, PhD, Department of Psychology, New York University, New York, New York

Norbert L. Kerr, PhD, Department of Psychology, Michigan State University, East Lansing, Michigan

Arie W. Kruglanski, PhD, Department of Psychology, University of Maryland, College Park, Maryland

John M. Levine, PhD, Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania

Nira Liberman, PhD, Department of Psychology, Tel Aviv University, Tel Aviv, Israel

Gregory R. Maio, PhD, School of Psychology, Cardiff University, Cardiff, Wales

Amy Noll McLean, MA, Department of Psychology, Duke University, Durham, North Carolina

Kari A. Merrill, PhD, Department of Psychology, Duke University, Durham, North Carolina

Mario Mikulincer, PhD, Department of Psychology, Bar-Ilan University, Ramat Gan, Israel

Gordon B. Moskowitz, PhD, Department of Psychology, Lehigh University, Bethlehem, Pennsylvania

Brian A. Nosek, PhD, Department of Psychology, University of Virginia, Charlottesville, Virginia

Andrzej Nowak, PhD, Department of Psychology, Florida Atlantic University, Boca Raton, Florida, and Department of Psychology, Warsaw University, Warsaw, Poland

Kevin N. Ochsner, PhD, Department of Psychology, Columbia University, New York, New York

Allen M. Omoto, PhD, School of Behavioral and Organizational Sciences, Claremont Graduate University, Claremont, California

Daphna Oyserman, PhD, Institute for Social Research, University of Michigan, Ann Arbor, Michigan

Richard E. Petty, PhD, Department of Psychology, Ohio State University, Columbus, Ohio

Michel Tuan Pham, PhD, Department of Marketing, Columbia University Business School, Columbia University, New York, New York

Thane S. Pittman, PhD, Department of Psychology, Colby College, Waterville, Maine

Jo-Ellen Pozner, PhD, Kellogg School of Management, Northwestern University, Evanston, Illinois

Neal J. Roese, PhD, Department of Psychology, University of Illinois at Urbana–Champaign, Champaign, Illinois

Alexander J. Rothman, PhD, Department of Psychology, University of Minnesota, Minneapolis, Minnesota

Peter Salovey, PhD, Department of Psychology, Yale University, New Haven, Connecticut

Brandon J. Schmeichel, PhD, Department of Psychology, Texas A&M University, College Station, Texas

Norbert Schwarz, DrPhil, Institute for Social Research, University of Michigan, Ann Arbor, Michigan

Mark D. Seery, PhD, Department of Psychology, University at Buffalo, Buffalo, New York

Gün R. Semin, PhD, Royal Netherlands Academy for Arts and Sciences, Free University, Amsterdam, Netherlands

Eldar Shafir, PhD, Department of Psychology, Princeton University, Princeton, New Jersey

Phillip R. Shaver, PhD, Department of Psychology, University of California, Davis, California

Jeffrey W. Sherman, PhD, Department of Psychology, University of California, Davis, California

Jeffry A. Simpson, PhD, Department of Psychology, University of Minnesota, Minneapolis, Minnesota

David Sleeth-Keppler, PhD, Department of Psychology, Winthrop University, Rock Hill, South Carolina

Mark Snyder, PhD, Department of Psychology, University of Minnesota, Minneapolis, Minnesota

Wolfgang Steinel, PhD, Department of Psychology, Leiden University, Leiden, The Netherlands

Elena Stephan, PhD, School of Psychology, University of Southampton, Southampton, United Kingdom

Fritz Strack, PhD, Department of Psychology, University of Würzburg, Würzburg, Germany

Timothy J. Strauman, PhD, Department of Psychology, Duke University, Durham, North Carolina

Philip E. Tetlock, PhD, Haas School of Business, University of California, Berkeley, California

Leigh Thompson, PhD, Kellogg School of Management, Northwestern University, Evanston, Illinois

Zakary L. Tormala, PhD, Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana

Yaacov Trope, PhD, Department of Psychology, New York University, New York, New York

Tom R. Tyler, PhD, Department of Psychology, New York University, New York, New York

Robin R. Vallacher, PhD, Department of Psychology, Florida Atlantic University, Boca Raton, Florida

Eric Van Dijk, PhD, Department of Social and Organizational Psychology, Leiden University, Leiden, The Netherlands

Gerben A. Van Kleef, PhD, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Paul A. M. Van Lange, PhD, Department of Social Psychology, Free University, Amsterdam, The Netherlands

Mark Van Vugt, PhD, Department of Psychology, Keynes College, University of Kent at Canterbury, Kent, United Kingdom

Kathleen D. Vohs, PhD, Department of Marketing, Carlson School of Management, University of Minnesota, Minneapolis, Minnesota

Duane T. Wegener, PhD, Department of Psychological Sciences, Purdue University, West Lafayette, Indiana

Robert S. Wyer, Jr., PhD, Department of Psychology, University of Illinois at Urbana–Champaign, Champaign, Illinois

Kate R. Zeigler, BA, Department of Psychology, Colby College, Waterville, Maine

Preface

More than 10 years have gone by since the first edition of *Social Psychology: Handbook of Basic Principles* made its appearance. It has been an exciting decade for social psychology, with dynamic developments transforming the conceptual and empirical landscape of our field. New topics and principles have been discovered and vigorously explored. Classic issues have been approached from novel vantage points and via new methodologies.Young generations of social psychologists have emerged whose work has ushered in fresh, often boldly irreverent, perspectives on traditional issues. And, amidst all the excitement of progress, along with the strong winds of scientific change that have been blowing, some fundamental constancies have not only remained but have become increasingly rich in their implications. The rigor of social psychological research and the self-critical attitude that has become our stock in trade have continued to characterize our work. Another constant presence that we editors have certainly felt has been the need to take stock of the teeming, positively accelerated research activity in our field, and to clarify to ourselves what principles of social behavior seem to emerge out of the fast-flowing stream of effects, phenomena, and results that our research endeavors bring forth. Accordingly, the aim of the second edition of our handbook has been to capture the unfolding shifts in ways and means of social psychological research and to distill to the extent possible the principles of human social behavior that lurk beneath the surface of our manifold empirical observations.

Whereas the emphasis on principles is what ties the present edition of the handbook to its predecessor, the present contents, as well as the contributors, were unconstrained by their earlier counterparts. Of the 41 topics in the present handbook only 14 were represented in the previous volume. And of the 81 contributors to the present handbook only 10 contributed to the former project. In part, that is because the 1996 volume left out some research endeavors that were only beginning at that moment and that now have actualized their potential, having matured to full-fledged areas of active inquiry. In such a category belong the topics of "goals," "social metacognition," "the role of impulse," "information ecology," and "dynamic systems approach" or "social cognitive neuroscience." Other topics, though by no means novel historically, had been neglected by social psychologists for a period but now appear with a fresh impetus and with the benefits of theoretical and methodological developments that have visited our field in the meantime. In the latter category belong such topics as "value," "social power," "basic human needs," and "psychological distance."

To the extent that our choice of topics reflects the ongoing research activities in the field, it is apparent that the cognitive approach continues to grow as a significant presence in social psychology. Twelve contributions reflect this perspective—more than any other category. This strong showing is indicative of the richness and diversity of social psychological topics approached these days from the cognitive perspective. Still, the personal motivational emphasis continues to grow as well, suggesting its fruitfulness and utility for understanding social psychological phenomena.

A significant innovation in the present edition is the inclusion of chapters dealing with the applications of social psychology to a variety of applied domains such as "health behavior," "clinical psychology," "organizational behavior," "politics," "marketing," "law and the justice system," and "social action." Indeed, it is encouraging to witness how useful and influential social psychological principles have been across wide swaths of practical domains, attesting once again to the wisdom of Kurt Lewin's famous dictum that "there is nothing as useful [to applied concerns] as a good theory"—good theoretical social psychological principles, that is.

As in the prior edition of our handbook, all our contributors were asked to orient their discussions toward principles and issues rather than toward a detailed examination of specific studies. Indeed, it is our impression that they have done so quite admirably, tempered by the degree to which clear principles or issues were naturally emerging in the specific domains being reviewed. At least in spirit, then, the quest for social psychological principles was apparent in all our varied contributions. We feel and hope that such a quest will offer valuable blueprints to interested students and researchers, indicating where the field is at the moment, and where it needs to go to gain the kinds of insight that would *make a difference* to science and society. In this sense, then, we are pleased to offer our readers a handbook that holds a mirror to social psychology as we have begun our voyage into the 21st century and that reflects the state of our discipline and its exciting research directions.

ARIE W. KRUGLANSKI
E. TORY HIGGINS

Contents

III. PERSONAL MOTIVATIONAL SYSTEM

IV. INTERPERSONAL SYSTEM

PART I
BIOLOGICAL SYSTEM

Evolutionary Theory for Social and Cultural Psychology

Linnda R. Caporael

Humans are a fundamentally social species. They are unable to reproduce and survive to reproductive age without a group. They lack natural defenses such as slashing canines, bodily armor, or fleetness of foot; they have an extended infancy followed by years of further development before they reach reproductive maturity. Many of the evolved characteristics that have permitted humans to adapt to a wide range of physical environments, such as omnivory and tool making, create dependence on collective knowledge and cooperative information sharing among individuals and between generations. Humans are an *obligately interdependent* species (Brewer & Caporael, 2006a; Caporael & Brewer, 1995). Our evolution is a product of the coevolution of genetic endowment, social structure, and culture (Boyd & Richerson, 1985; Caporael, 2001a; Fiske, 2000; Janicki, 1998; Li, 2003; Mesoudi, Whiten, & Laland, 2006).

The dominant metaphor in the human sciences, "selfish gene theory" (Dawkins, 1976) or the "gene's-eye view" of evolution, has limited resources to theorize groups and the consequences of group living for the human mind. Human sociality is largely conceived in terms of Machiavellian intelligence, deception, exploitation, and coalition formation for intergroup conflict (Alexander, 1989; Byrne & Whiten, 1988; Dawkins, 1976; Flinn, Geary, & Ward, 2005), primarily because the gene's-eye view uses familiar economic explanations tied to a genetic currency that cannot be cashed out. This approach can still broaden our understanding of human cooperation, and a movement to do so is under way in a number of related disciplines such as economics (Gintis, Bowles, Boyd, & Fehr, 2003) and anthropology (Henrich et al., 2005), as well as in social psychology (Baumeister & Leary, 1995; Sedikides & Skowronski, 1997; Simpson, Schaller, & Kenrick, 2006).

This chapter aims to expand such approaches by drawing on multilevel selection theory, which offers a framework more appropriate than genes for psychological levels of analysis. *Coordination*, rather than the evolution of biological altruism (genetic self-sacrifice), is the central problem of a human psychology informed by biology and culture (Brewer & Caporael, 2006a; Caporael & Baron, 1997; Kameda & Tindale, 2006). Coordination refers to the skillful integration of diverse elements into a harmonious operation. This could mean, literally, a surgical operation, with its diverse artifacts, human skills, body movements, and the plan for the surgery. However, a broad range of human activities involves coordination from knowledge creation to walking down an urban street, attending a dinner party, or training a dog. Similarly, exploitation, intergroup conflict, and coalition formation presuppose highly developed coordination skills. While there is a sense in which all coordination benefits "selfish genes," a genetic level of explanation skips over the psychological and social processes that compose human cognition and behavior.

The following sections offer an evolutionary framework drawing on the implications of multilevel evolutionary theory. An expanded evolutionary theory has a crucial advantage in the human sciences: It enables us to point at recurrence without getting embroiled in nature–nurture controversies. A vocabulary for the repeated as-

sembly of diverse elements does not require making unsupportable claims about specific genetically based adaptations. Following the evolutionary framework is a cursory review of current knowledge about the "ancestral environment." In view of that research, drawn primarily from paleoanthropology, I propose a model of core configurations in face-to-face groups as the selective environment of uniquely human mental systems and the reorganization of earlier evolved "building blocks." If the model is valid, we should expect human mental systems to correlate with core configurations. We should also expect that capabilities specialized for specific configurations can be extended and recombined for novel behavior and situations, including life in the modern world.

Many psychologists will find a fit between their research and the core configurations model (Caporael, 1997). An important advantage of the model over other adaptationist approaches (Andrews, Gangestad, & Matthews, 2002; Schmitt & Pilcher, 2004) is that it can enable greater precision in psychological theory. One could argue not only that some particular feature is functional but also that it is functional with respect to particular configurations of structure and activity. The model also has the connective tissue to extend the scope of psychological theory and research. Brewer and Caporael (Brewer, 1997; Brewer & Caporael, 2006b; Caporael, 2001b) have illustrated such integration and extension for research and theory on social identity. Finally, the model rejects the traditional reductionist view of biology as the basis for psychology, which in turn is the basis for culture. These three arenas are more reasonably viewed as coconstructive in human evolution. Human evolution—and human psychology—is mediated by artifacts and situated in sociohistorical contexts (Cole, 1996, 1997). An interesting result of such coconstruction is that *it is human nature to invent human nature*, most visibly in terms of locally distinctive folk psychologies (Lillard, 1998). The researcher's job is to understand how humans invent and coordinate mind and activity. Analytically, our three arenas–biology, psychology, and culture—must lean on each other for validation, correction, and inspiration.

In short, this chapter is broadly synthetic. It draws on diverse disciplines far from psychology for a generalized framework constructed with the conviction that social and cultural psychology have a great deal to give back to the evolutionary efforts to understand humankind in other disciplines (Brewer, 2004, 2005).

MULTILEVEL EVOLUTIONARY THEORY

The gene's-eye view of evolution is based on a technical definition drawn from population genetics—changes in the frequency of alleles (genes) in a population. The major goal of the gene's-eye view of human evolution is to identify adaptations—gene-based traits that evolved through a process of natural selection in the evolutionary past. To use the gene's-eye view of evolution in any scientifically respectable way, psychologists would have to shoulder certain "burdens of proof" from biology

(Lloyd, 1999). These include specifying a well-defined trait, eliminating other explanations for the trait such as genetic drift, developmental constraints, correlated traits, and, in the case of humans, cultural explanations. The most critical evidence, however, is the demonstration that variations in the phenotypical trait are correlated with variations in reproductive or genetic fitness. Although biologists have a variety of methods for determining such covariation, few are practically or ethically possible with humans, and the most successful cases have been made for physiological characteristics such as sickle cell anemia or lactose intolerance (Durham, 1976). Demonstrating covariation between a human psychological trait and fitness would also be a logical challenge.

Certainly one major difficulty in the population geneticists' view of Darwinism is that it has little to say about development, environment, or culture. It lacks the theoretical resources to engage psychology because the gene's-eye view works *below* the level of the organism. Its generality is such that it works equally well for oysters, ants, birds, and humans because it is about changes in gene frequencies. Crucially, it tells us little about human (or oyster, ant, and bird) phenotypes in the absence of what we already know or think we know about them. Evolution then becomes an explanatory project (Smith, 2006) that might add another layer of understanding (DeKay & Buss, 1992), but could fulfill little in the way of Darwin's revolutionary promise of a new foundation for psychology or light on human origins. Moreover, the gene's-eye view is blind to the distinction between social and asocial species, a critical distinction for hypothesizing sociality at the phenotypical level (Brewer, 2004; Caporael, Dawes, Orbell, & van de Kragt, 1989). This is not to say that the gene's-eye view of evolution is false; merely that for Darwinism to be productive in psychology, we need an expanded view of evolution appropriate to our current levels of analysis.

Multilevel Selection Theory

In 1987, evolutionary biologist Leo Buss published a groundbreaking book, *The Evolution of Individuality.* Evolutionists going back to Darwin had recognized that selection could occur at different levels of the biological hierarchy wherever there was heritable variation and differential replication (Brandon, 1990; Buss, 1987; Campbell, 1974; Gould, 1980; Wilson & Sober, 1994). What the gene's-eye view overlooked (or rejected [Dawkins, 1982]) was that levels of biological hierarchy—chromosomes, cells, and multicellular individuality—also had to evolve. Buss's (1987) contribution was to detail the evolution of multicellularity, that is, how individuals composed of reproductive gametes and nonreproductive somatic cells that "gave up" reproduction could evolve.

This work ushered in a new era in biological theory. Levels of biological hierarchy roughly represent major transitions in the evolution of life on earth, especially in the way that information is stored and transmitted (Maynard Smith & Szathmáry, 1995; Szathmáry & Maynard Smith, 1995). They are also represented in the

bodies of multicellular organisms (Buss, 1987) and serve as rough guides to levels of analysis. The major characteristics of such transitions are that entities that were capable of independent reproduction prior to the transition are able to reproduce afterward only as part of a larger unit. Cellular mitochondria, for example, are believed to have been independent cellular organisms that continue to have their own DNA lineage. Selection on multiple levels is associated with increased task specialization (energy production by mitochondria in the cell) and new forms of information relevant to evolution and development (e.g., the genetic code, sexual reproduction, epigenetic inheritance, and human language) (Jablonka & Lamb, 2005; Szathmáry & Maynard Smith, 1995). At each level, the environment for an entity is the next level, which can also have downward causal effects (Campbell, 1974). For example, the environment for genes is the cell, and cellular products influence which part of the gene will be active. Humans, as individuals and as groups, constitute two levels of biological organization of major interest to social and cultural psychologists.

All such transitions pose the same "free-rider" question. What prevents lower-level entities from "defecting" and disrupting higher-level organization? The giving up of reproduction *can* be explained in terms of genetic self-interest. But far more predictive power is gained from considering conflicts and synchronies over different levels of organization, where there are different types of payoffs, obstacles, path dependencies, and locked-in systems (Caporael, 2003). Sex, for example, is ubiquitous, not for the single reason of benefiting the genes but because it is multiply synchronous, selected at the genic, cellular, individual, and group levels (Buss, 1987, p. 181, fn). What once seemed to be competing alternative explanations for the evolution of sex, taken together, are the discovery of synchrony over different levels of organization. In addition to synchrony, we should expect conflict among levels as well. Cancer is one example: A cell lineage reproduces itself at a cost to the organism. In effect, it has defected, although at a cost to itself in the end.

Multilevel evolutionary theory offered insights for an evolutionary theory at a psychological level of analysis. First, it provided a view of evolutionary processes as hierarchically structured; the process of selection is *situated*, *relational*, and *embodied* in specific contexts (Eldredge & Grene, 1992; Endler, 1986). The situated environment for genes is the cellular machinery with which DNA and its products interact. The situated environment for humans is the face-to-face group, through which they are able to interact with the habitat. Stressing "situatedness" in a context focused on selection as the statistical result of *relations* between entity and ecology (Caporael & Baron, 1997; Endler, 1986) in place of the anthropomorphic metaphor of selection favoring a trait, a locution that obscures the conditions from which selection may result. To take a psychological example, there is no basis for assuming, as the trait view implies, that a dominant individual will be so across interactions. An individual may act socially dominant in a dyad but not in a group of five people working on a shared problem. The situated context of behavior is reliably different, and the func-

tional relation between dominant behavior and the context may well differ, too (Mischel & Shoda, 1995). Functional relations refer loosely to the fit or match between entity and context in achieving certain ends (e.g., movement and nutrition). The focus on relations also opened the door to artifacts as part of the relation. In humans (and some other animals; Dawkins, 1982) functional relations can be achieved artificially. Humans can have disabilities that in other species would be terminal, but they can also use "artificial traits" (e.g., wheelchairs, canes, and eyeglasses) to achieve basic functional relationships such as locomotion or vision. The functional relations of an organism and environment are also *embodied*, that is, available to empirical investigation (Andrews et al., 2002; Caporael & Baron, 1997; Levins & Lewontin, 1985).

Multilevel evolutionary theory also accommodates groups as a level of organization. Darwin (1871/1981) and later Williams (1966; whose book is an extended argument against group selection and for the logic of costs and benefits accruing to the gene) believed that humans might have traits resulting from group selection and informally worked out the argument. However, Williams's (1966) book might have been too successful. By the early 1970s, group selection was not merely dropped; it was passionately damned (Pollock, 1989). Unfortunately, psychologists have followed suit, although Brewer and Caporael (2006a) have urged reconsideration. Wilson and Sober (Sober & Wilson, 1998; Wilson & Sober, 1994) formally worked out the argument for the evolution of "altruistic genes" from the perspective of multilevel selection.

Briefly, subgroups ("trait groups") of a larger group vary in the proportion of altruists and nonaltruists. Some groups will have low ratios of altruists to nonaltruists; other groups will have high ratios. Within all the groups, the nonaltruists increase relative to the altruists. However, if groups with higher proportions of altruists also produce more offspring, then the total number of altruists in the global population will increase when groups disperse and form new groups. In other words, within-group selection favors the evolution of self-interest whereas between-group selection favors the evolution of cooperation (because altruists benefit themselves as well as others). The important difference between a gene's-eye view of altruism and the multilevel evolutionary view is really in the next step. The selfish gene theorist seeks an explanation in terms of genetic self-interest. The multilevel theorist looks further for the evolution of individual mechanisms for developing and maintaining group memberships and group-level mechanisms for enforcing groupish behavior in the face of individualism. (Some readers may want to interpret the distinction as a parallel to the distinction between ultimate evolutionary causes and proximate causes. However, this dichotomy is insufficient. The ultimate cause will typically be an explanation about how some particular piece of behavior benefits the genes of an individual. The proximate explanation is typically about the stimulus that activates an adaptation [e.g., dust or some other stimulus causes the sneeze] rather than about the mechanism that responds

to the stimulus, which presumably has an evolutionary history [Kitcher, 1985].)

Repeated Assembly

One of the difficulties of evolutionary language is that it lacks a good vocabulary for waffling. An adaptationist vocabulary makes it impossible to avoid sounding like a genetic determinist. By definition, calling a trait an adaptation is a commitment to genes as a primary explanation of its recurrence and operation. Even for most descriptions of gene–environment interactionism, the possible alternatives for the trait are in the gene, and the environment merely fills gaps in "open programs," "shapes" innate proclivities, or "shunts" behavior from one option to another (Buss, 1999; Crawford & Anderson, 1989; Tooby & Cosmides, 1992). One reason social and cultural psychologists have been reluctant to turn to evolutionary theory is an uneasiness about making genetic claims when they are not biologists. At the same time, the phenomenon of recurrence, between generations and across cultures, is one of the appeals of an evolutionary approach.

Multilevel selection helps to solve this problem. Expanding neo-Darwinism calls attention to organisms' interaction with the environment and to phenotypical development. Phenotypes repeatedly assemble generation after generation from constructive interactions among multiple recurrent resources, including genes, cellular machinery, social resources, and the reliable presence of critical features of the habitat, from caregivers to oxygen in the air (Gottlieb, 2002; Lickliter & Honeycutt, 2003; Oyama, 1985). As a result, the concept of heritability is expanded to include a range of resources that mediate phenotypical outcomes (Cole, 1996; Griffiths & Gray, 1994). As humans, we inherit genes, but also attitudes, practices, place, nationality, expectations for behavior, and so forth. There may be change (e.g., in traditions) within lifetimes, but this simply indicates low reliability of recurrence. Heritability itself is thus no indicator of the universality, fixity, or difficulty of changing organism–environment relations.

A vocabulary of *repeated assembly* (Caporael, 2003) points to recurrences that we can observe—recurrences from generation to generation as well as recurrence across cultures. Not only do organisms repeatedly assemble, but so also do the products of organisms, including human ideas, artifacts, and cultural practices (Dawkins, 1982). Repeated assembly enables us to escape the dichotomy of biology and culture, nature and nurture, genes and environment; it does not require us to make commitments about genes in the absence of hard evidence; and it is consistent with the Darwinian model of selection, variation, and retention. More specifically, repeated assemblies are recurrent entity–environment relations composed of hierarchically organized, heterogeneous components having differing temporal scales and cycles of replication (Caporael, 2003). That description is quite a mouthful: The following discussion is meant to provide a brief sense of the nature of repeated assemblies.

Recurrence is fairly obvious, although we are most interested in a particular kind of recurrence, that which is recursive, where the seeds of the next cycle are part of a current cycle. Entity–environment relations are, at the psychological level, organism–environment relations, where the environment may be features of the physical habitat or configurations of conspecific groupings. Hierarchical organization (actually, heterarchical because causal effects may be downward and upward) is built into the evolving, jury-rigged structure of living organisms. Heterogeneous components refer to the multiple resources that compose a repeated assembly, although which resources are the objects of research depends on the focus and perspective of the research question. Components can be heterogeneous partly because they recur over different time scales. DNA recurs over macro-evolutionary time; at the other extreme, social rituals can rise and fall within a lifetime (affecting phenotype, but having no systematic selective effects).

As an example, consider the fertilized egg or zygote. A zygote is composed of multiple resources. There are two sources of DNA from the gametes, centrosomes from the sperm, maternal information in the egg cell, and an appropriate setting in the body. If all these resources are in the right place at the right time, a zygote automatically results. There is no genetic program; in fact, the genes are inactive during the first few cell divisions of the fertilized egg. The difference that matters for the differentiation of structure and specialization of tissue in the first cell divisions is time and space—recall, all DNA in cells are the same. Cells in the zygote adhere to each other, the lose adhesion, and migrate to different regions. The cellular activity influences genetic activity, which in turn affects cellular activity.

The zygote is a repeated assembly, and like other repeated assemblies, it is inherently interactionist. It emerges from the constructive interaction of genes, temporal and spatial relations, and the changing cellular machinery. Its continued development and existence depends on the reliable recurrence of appropriate contexts—from cellular machinery to social events to regularities in the habitat to persistent resources such as gravity and atmosphere (Griffiths & Gray, 1994). Genes are always part of the repeated assembly of an organism, but so also (in the human case) are group memberships, customs, artifacts, learned skills, languages, and worldviews (see Figure 1.1). These resources may vary wildly in their specifics but are nevertheless critical to human survival and reproduction.

Within historical, ecological, and developmental constraints, repeated assemblies are almost infinitely variable. Change in any repeating component of an assembly, whether genetic or epigenetic, can increase or decrease the probability of future assembly. This is because repeated assemblies are concretely situated in specific flows of interaction. The situatedness of activity makes the unit of analysis the organism-in-setting, where the focus is on the relation between entity and context, or person and setting. To illustrate the variability of concretely situated activity, consider any everyday routinized activity—say going to a supermarket (Lave, 1988). The su-
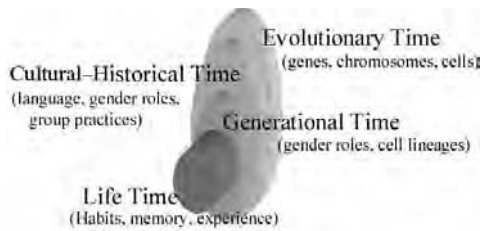
FIGURE 1.1. Recurring multiple resources having different scales of time contribute to the repeated assembly of unique individuals. Resources (which are usually repeated assemblies themselves) differ in their scale, meaning how long they have been repeated (e.g., the DNA code has been repeated far longer than any particular cultural resource), and their cycles of generation.

permarket has a more or less invariant set of practices and structure of shelves, aisles, and grouping of products. Yet, despite the routine, each trip is unique. The unit of analysis is the pattern of the assembly (shopping), until some feature that makes the pattern unique is picked out. Repeated assemblies are like Eleanor Rosch's (1977) prototypical categories. They have components with varying degrees of membership. Some components, however, will be more prototypical or central than others.

Although some traits of organisms do inspire marvel for their complexity of design and are designated as adaptations, evolutionary theory is ultimately a theory about the history of use and not design (Millikan, 1984). The same is true for repeated assemblies. Some recur because in the past, some components, in some specific relation to their contexts, contributed to the replication of the assembly. Specialized functions describe what an assembly is "supposed to do" under historically ideal conditions. (This is not to say that they cannot serve other functions as well.) For example, a function of sperm is to fertilize an egg cell, even though ova are rare in the world of sperm and most fail to fulfill their function. Nevertheless, because some ancestor sperm did find and fertilize ova, sperm are repeatedly assembled. Historically ideal conditions are not necessarily statistically average or modal conditions.

Specialized functions may be quite awkward, inefficient, and clumsy, far from optimal design, and even constrain the spread of better designs should they appear. They reflect the meandering path dependencies that occur in history. A cultural example of the result of such path dependencies is the QWERTY keyboard. Its design originated as a way of fitting the striking mechanisms for the keys into the machine to reduce jamming by slowing down fast typists (e.g., by separating TH). As the machine constraints were altered, other keyboard designs with better ergonomic features for users were developed. However, these better designs were unable to invade the keyboard niche. We know (or are able to know) much more about the recent past than the ancestral past, which serves as a caution. Claiming that a particular psychological mechanism or process has a specialized function depends on our best guesses about history. All things considered, we may be more like the QWERTY keyboard than the more optimally designed Dvorak layout.

Some repeated assemblies have no function; they are like the superstitious behavior of Skinner's circling pigeons. Some can result from illusory contingencies or once useful behavior. Other repeated assemblies, including biological adaptations, may have had a function at one time but become dissociated from it. Langer (1989) tells the story of a woman who would always cut a slice off the end of a roast before placing it in the oven. To make a long story short, the woman said she did this because her mother did it; mother said she did it because her mother did it; and grandmother did it because the pan she used for roasting was too small to accommodate the whole piece of meat. Of course, such dissociations can occur in evolutionary time (then they are vestiges) or, as Boyd and Richerson (1985) argue, in cultural–historical time, resulting in practices that have no function or are even costly.

Repeated assemblies can be characterized in terms of scales of time (i.e., persistence, cycles, and the frequency of repetition). Some components, such as the DNA code, have been parts of repeated assemblies since the beginning of life. Other components, such as a cultural environment, have a temporal scale measured in historical terms. Still others recur on the scale of a lifetime. "In the past" does not necessarily refer to a dichotomy between the Pleistocene and the urban present. It may refer to time measured in terms of ontogeny (in "habits of a lifetime"), to cultural–historical time (as in "2,000 years of Western civilization"), or to geological–evolutionary time (the temporal scale of DNA). Paradoxically, time expands the notion of heritability; we inherit genes but also attitudes, practices, place, nationality, expectations for behavior, and so forth. Assemblies may persist in evolutionary time, among Americans, in the Smith family, or in J. Smith. A more extensive discussion of reproduction, development, and culture on different scales is in Wimsatt and Griesemer (in press).

What is *not* a repeated assembly? The answer often depends on level of analysis and perspective. Every individual—even one of a monozygotic twin set—is unique. No twin has precisely the same experiences as the other. However, organisms in general are repeatedly assembled. Often the product of a repeated assembly with specialized functions is unique. For example, the processes involved in the healing of wounds are repeated assemblies in organisms. However, the specific scars that result are the unique results of an individual's history.

By definition, errors would not be repeated assemblies. Despite considerable redundancy in biological systems, errors do occur in development and produce organisms unable to survive or reproduce. Hence, their lineages cease to repeatedly assemble. At first glance, it seems that there should be few errors that repeatedly assemble in artifactual systems because these can be corrected when discovered. However, the poor design of artifacts can cause undesirable outcomes to repeatedly assemble because of ignorance or even outright deceit, as in the cases of the Ford Pinto's faulty gas tank and the

Dalkon shield. As in biology, complex, poor design can persist, although the mechanisms will be clearly and identifiably different from those sustaining artifacts.

An adaptive response to a novel event may be unique; but even though it is useful, it may fail to repeatedly assemble. In his examination of "cognition in the wild," Hutchins (1996) observed the failure of a ship's navigational systems. Through a process of trial and error, two men eventually solved the problem of plotting location; however, the innovation was not saved in the cultural system. The situation for using the solution again is rare because the failure of redundant navigational systems is rare.

Rethinking Evolution, Sociality, and Psychology

An expanded evolutionary theory immediately generates a set of interesting predictions and guidelines. Because of group selection and downward causation, we should expect the evolution of a variety of specialized processes for coordination among group members. (Unlike cooperation, coordination does not necessarily imply prosocial behavior—nor does it exclude it.) Some coordinating processes will have to do with motor coordination (Boinski & Garber, 2000), interactional synchrony (Bernieri, Davis, Robert, & Knee, 1994), and other phenomena related to rhythmicity, coordination, and body movement. Research in automaticity may also point to another system or set of related systems of general social coordination. For example, in a well-known study by Bargh, Chen, and Burrows (1996), priming students with words such as *Florida* evoked a stereotype of the elderly that reduced the speed of students walking down a hallway compared to students who did not receive the prime. The phenomenon of *stereotype threat*, when high performance is lowered when a group stereotype is made salient, illustrates tacit control of behavior (Fiske, 1993; Steele & Aronson, 2003). Such research suggests the types of mechanisms that function as a coordinated/coordinating stream of behavior.

Multilevel selection also suggests that there should be conflicting individual and social motives that result from selection on those levels. Humans are ambivalently social because both individual and group levels are under selection. The psychology of optimal distinctiveness and group-oriented motivations (Brewer, 1991; Brewer & Roccas, 2001) makes a great deal of sense as a regulator of individual and groupish behavior.

The emphasis on selection arising from relations between organism and environment maps to a growing emphasis in cognitive science on embodiment and adaptive responding to the environment (Agre, 1997; Clark, 1997; Hendriks-Jansen, 1996), the evolution and operation of socially situated cognition (Caporael, 1995; Smith & Semin, 2004), and a revival and reassessment of Gibson's (1979) concept of affordances (Norman, 1990; Zebrowitz & Montepare, 2006; Zhang & Patel, 2006). Dijksterhuis and Bargh (2001) posit a "perception–behavior expressway" that might also be part of this revival (given further consideration about the role of representation in cognition and action). Affordances originated in evolutionary considerations (unlike the computational metaphor) and bring with them concepts such as attunement and action boundaries that could deal with humans and their interactions with objects (Fajen & Devaney, 2006; Fajen & Turvey, 2003; Núñez & Freeman, 1999).

THE MIND'S NATURAL ENVIRONMENT

The previous sections established a general warrant for sociality and sketched an evolutionary approach that is suitable for psychological levels of analysis. In evolutionary biology, that might be enough. The details about genotypical variation and selection history of a single species tend to be sparse. The usual assumption is that the current traits of an organism evolved in the past and are maintained in the present by stabilizing selection, which shifts the problem from history to ecology. The field biologist can then demonstrate natural selection through a variety of methods, including covariation between trait and reproductive fitness (Rose & Lauder, 1996).

Things are tougher in the human case. We are well aware from sciences such as geology, climatology, paleoanthropology, and archaeology that the lives of modern humans are different from what they were before the rise of urban ultrasociality (Campbell, 1983). One clear consensus is that humans lived as foragers, and another is that the physical constraints of the body's evolved interaction with the habitat (e.g., the mechanics of vision or caloric requirements) would remain the same between the ineffable past and the present. Not all such characteristics necessarily would be *human* adaptations. Some would be adaptations of an ancestral organism (e.g., color vision); others would be the result of developmental constraints that arise from highly conserved regulatory genes. Other characteristics of humans that are adaptive and do show measurable variation, such as reading, are recognized as having appeared so recently as to not enter into an evolutionary explanation (although some of the components of reading probably would). These issues aside (and they are nontrivial), the big question is, what was life like in the evolutionary past, or more to the point, what was psychology like?

Life in the Pleistocene

Donald Symons (1992) wrote that life in modern industrial societies looks like unplanned evolutionary experiments. In fact, the whole of human evolution looks like unplanned experiments (Caporael, 2004; Potts, 1998; Tattersall, 1995). Ancestral populations contracted, expanded, appeared, dispersed and became extinct against the turbulent background of global climate change. Today some paleoanthropologists recognize as many as 23 species.[1] Deep-sea drilling for sediment cores shows dramatic climate changes beginning about 50 million years ago, when a major cooling phase began leaving the temperatures of today about 12–15° C cooler. This was not a smooth transition. There were major climate reversals, the polar ice caps formed, and the Mediterranean Basin dried up and refilled. Regional climates were affected by

the Earth's orbital peculiarities, the advance and retreat of glacial ice, and geological activity such as volcanoes, faults, uplifts, and rain shadow. During this period an enormous variety of Miocene apes became extinct. One surviving lineage included the common ancestor of chimpanzees and bipedal apes, including the australopithecines. These were small-brained animals whose locomotor specializations (splayed toes, curved fingers) indicated adaptation to both walking and climbing. Turbulent climatic change continued through the period of human evolution, which occurred primarily in glacial conditions, interspersed with shorter, warmer interglacials (the most recent beginning about 10,000 years ago).

The phylogenetic pattern (a phylogeny is a diagram depicting ancestor–descendant relationships in geological time) for human evolution is quite unlike the widely depicted image of uphill evolutionary progress that begins with a knuckle-dragging ancestor. Instead, the pattern shows branching and overlaps. About 2 million years ago, as many as five bipedal species may have coexisted, one (perhaps even one yet to be found) is ancestral to humans. One nonancestral species had an amazingly rapid 30% expansion of the brain in less than 700,000 years, an event that would occur again in the *Homo* lineage, between 600,000 and 200,000 years ago, when brain size approximates the modern size, but with considerable variation.

A later rash of lineages, sometimes collectively called archaic *Homo sapiens*, ebbed and flowed through Africa and Eurasia, and eventually Europe. Arguably, all modern humans are descended from a limited population, perhaps as small as 10,000 breeding individuals of *Homo sapiens* that survived a major population crash about 130,000 years ago (Lahr & Foley, 1998). Although there are important exceptions in Africa, *Homo sapiens* were behaviorally (based on artifactual data) indistinguishable from Neanderthal populations until about 50,000 years ago (Klein, 1999). Then there is an inexplicable outburst of modern behavior in art, personal ornamentation, cultural variation in tool use, and the working of materials, such as shells, previously unseen in the archaeological record. Modern humans eventually replaced the Neanderthals in Europe about 30,000 years ago, spread through Asia, probably constructed watercraft for several migrations to Australia and crossed the Bering Strait. Inconceivably, we modern, industrialized humans overlapped for 75% of our evolutionary history with another intelligent species.

In the view of evolution that sees it as an uphill march of progress, new stages in human evolution were expected to have greater intelligence and, of course, would have the latest new tools. That view dropped away when new techniques for dating materials became available. Humans were bipedal for approximately 2 million years before the earliest known stone tools. New tools do not appear at the same time as new hominins, which makes sense—tools have to appear within species that already exist (Tattersall, 2000). One way of characterizing ancient tools is by characteristic modes of production (Foley, 1987). Early artifacts are remarkably stable though time.

The same mode of tool production is used by different species with little variation and over long distances. The first tool user, *Homo habilis*, was essentially an australopithecine with a slightly larger brain, used Mode 1 (Oldowan) choppers and handaxes. Although the form is refined through time, Mode 1 tools were produced for over 1 million years and are found from the southern tip of Africa to the eastern reaches of China to the western edges of the British Isles.

In Africa, these are eventually replaced with bifaces (flakes removed from both sides of a stone core). Their use persisted, again with some refinement, for another million years. In the Far East, however, Mode 1 tools are replaced directly by Mode 5 tools, which may be evidence that *Homo sapiens* replaced existing *Homo erectus* populations. In Africa and Europe, the changeover in modes of production begins to speed up with Mode 3 (Mousterian). These tools are usually associated with large-brained (1,200 cc) hominins,[2] especially the Neanderthals. There was a shorter period of production, about 100,000 years, and a greater variety of form. Modes 4 and 5 are associated with modern humans, have a high turnover of different types of tools, none lasting more than 5,000–10,000 years. Some of this later variability may be attributed to stylistic differences between groups. Stone transport, diet, hearths, structures and eventually pigments, ornaments and composite tools (tools made with two materials, for example, obsidian flakes set into a wooden haft to make a knife) also appear in the fossil record. How to relate artifacts in the archeological record to psychological function is unclear, largely because, with few exceptions (Cole, 1997; Hutchins, 1996), artifacts have not played a major role in psychological theories until recently and represent a vast area for future research.

To sum up, research on the last 3 million years of human evolution indicates shifting habitats, with different regions responding differently to large-scale climatic changes. Fragments of populations would become isolated. They might reunite after long periods of time (if they had not become reproductively isolated); speciate from the parent population through gradual selection to a new or changing environment; stay unchanged by following their preferred habitats; or simply become extinct (Lahr & Foley, 1998). Globally, humans lacked the stable habitats supporting ever more complex cognitive adaptations. For psychology, the fossil evidence suggests that we should expect the evolution of small-grained adaptations, more like Lego blocks (Bechtel, 2003; Scher, 2004) than large-grain, informationally encapsulated adaptations that would require exceedingly long periods of environmental stability for their genetic evolution.

The conventional wisdom since Darwin was that the lack of natural defenses such as claws and canines meant that bipedalism implied tools used for defense, and tools implied intelligence. (It was partly this deeply held assumption that made the Piltdown Hoax so successful for so long.) Fossil finds of the 1970s overturned the conventional wisdom. "Lucy" (*Australopithecus afarensis*) lacked natural and artificial defenses and was bipedal. Whatever the details, australopithecines must have lived in groups;

in a remarkable find, at least nine adults and four juveniles who may have died at the same time, perhaps in a flash flood, were unearthed. Group living also formed the context for the changes between *Homo habilis*, and *Homo erectus*. There was an important shift to hunting, and higher-quality nutrition may have allowed the redeployment of energy from a large gut to a growing brain (Aiello & Wells, 2002). There is some evidence of group buffering (individuals with injuries that would have required support to survive). A reduction in sexual dimorphism appears to be the result of females growing bigger (Klein, 1999); no data seem to exist as to whether the change was accompanied by a reduction of male conflicts. Considerable evidence exists for the manufacture, transport, and use of stone tools and even raw materials. Presumably stone tool technology was repeatedly assembled and refined in group contexts; however, the sheer stability, especially of mode of manufacture in earlier hominin species, is difficult to explain or grasp with our current psychological understanding. There are also few fossil data that tell us about social organization, parenting, sexual preferences, mating behavior, division of labor, and so on until quite recently in the archaeological record.

Some of these questions might be resolved by turning to the ethnographic literature on hunters and gatherers. There are certainly some universals, for example, food sharing (Gurven, 2004) and other behaviors closely tied to ecology. When these can be connected to material remains, such as a pattern of waste materials from chipping stone tools (Binford, 1972), ethnoarchaeology offers additional insight about practices in the past. However, even this approach is not as fruitful as hoped because the archeological record of hunter–gatherers is itself so variable (Kusimba, 2005). Moreover, the closer foragers come to issues of keen interest (e.g., sex and gender) to the early missionaries and anthropologists who first studied their lives, the more bias can be expected from the data (Leacock, 1981). Later studies of traditional hunter–gatherer groups are frequently of peoples "tethered" to water, missions, welfare rations, government posts, agricultural plots of their own or their neighbors, and sources of cigarettes and alcohol (Kelley, 1995). Finally, recent research (Henrich et al., 2005) shows that indigenous peoples are surprisingly variable in their economic choices (measured in structured games) and frequently fail to choose strategies and outcomes in the self-interested ways predicted by traditional game theory.

A Minimalist Scenario

The "bones and stones" of the human past provide important negative evidence, in that paleoanthropological research fails to support the assumptions that most people make about the evolutionary past and, if anything, should make us wonder where wonderfully detailed claims about the past do come from. Research by Landau (1984) and Moser (1998), both anthropologists, suggests that narrative and imagery have played a long and unrecognized role in scientific and popular conceptions of the past, and the work by Green and her colleagues on transportation in narrative and imagery may provide the relevant psychological processes (Green & Brock, 2000; Green, Strange, & Brock, 2002). For the present, the best scientific strategy appears to lie in constructing minimalist scenarios that are transparent with respect to their idealized and fictional origins and in apposition to popular narratives and images. (This strategy is similar to one common in social psychology, where scientific research is tacitly in apposition to folk psychology, and may often raise our "discomfort index" [Fiske, 2003].) Given the considerable individualism that infuses selfish gene theory, and the undervalued significance of group living as a selective factor in coordination, focusing on face-to-face groups in human evolution is a good start (Byrne & Whiten, 1988; Geary, 2005).

Evolutionists often write as if organisms calculated genetic costs and benefits and made individual choices about whether or not to group. Nonetheless, most animals are group living (Pulliam & Caraco, 1984), even if for no other reason than the failure of kin groups to disburse for some part of their life history. However, there is considerable variability across species in the intensity of interdependence of group living (Avital & Jablonka, 2000; Boinski & Garber, 2000). Quite often, sociality may consist of little more than a group of opportunistic and individualistic cooperators (Norris & Schilt, 1988; Williams, 1966). Schools of fish or herds of fleet deer are *aggregate* groups. There may be safety in numbers, and the risks of predation might be spread out among a group of animals. In some species, individuals that form aggregates are adaptively specialized for knowing when to group and when to be solitary. Other kinds of groups cannot break up. The individuals that form such groups *must* be part of a group in order to reproduce and survive to reproductive age. These intensely *social* groups organize individual efforts, communicate within the group, have tasks and roles for group members, and definable boundaries (Brewer, 1997). In extreme cases, such as wolf packs, naked mole rats, and meerkats (a social mongoose), only a single pair of individuals in the group reproduces, somewhat like the "germ line" of the body. Other group members (usually relatives) may help care for the offspring by feeding, guarding, and tutoring them in hunting skills. Human sociality is nowhere so extreme as to have a single breeding pair within a group. Nevertheless, human sociality is at the extreme end in that humans are unable to reproduce and survive to reproductive age without a group context. Grouping is an appropriate starting point for constructing a minimalist scenario.

Given their morphology and ecology, the ancestors of evolving hominins must have survived as groups rather than as individuals. Finding food, defense from predation, moving across a landscape—these matters of coping with the physical habitat—would have been largely group processes that began among the earliest arboreal primates, possibly as a result of changing patterns of dispersal in response to climatic and ecological changes. Over time, if exploiting a habitat is more successful as a collective group process than as an aggregate group or

individual process, then not only would more coordinated groups persist but so also would individuals better adapted to group living. Because local ecologies limit group size, the number of "niches" in a group are also limited, resulting in a feedforward ratchet and group fission when groups become too large to exist on the local ecology.

If individual humans cannot survive outside of groups, then the structural requirements for sustaining groups create systematic constraints on individual biological and psychological adaptations. Such downward causation (Campbell, 1974, 1990b) across system levels operates whenever structural requirements at higher levels of organization determine or shape some aspects of structure and function at lower levels. Thus, individual human cognitive systems do not have to be—nor can they be— fundamentally self-interested under all conditions. Humans are frequently motivated by outcomes of purely personal value; however, a variety of preconscious, conscious, group-level, and institutional mechanisms curb self-interest, albeit not always successfully. In the latter instance, when self-interests cannot be curbed, a group itself may disintegrate and be replaced by better coordinated groups.

The result of downward causal processes would be a shift to face-to-face groups as the selective context for uniquely human mental systems. Selection within groups would lead to the evolution of perceptual, affective, and cognitive processes that support the development and maintenance of membership in groups (Caporael et al., 1989). Without a group, the probability of reproduction and survival to reproductive age is lowered for humans. To say that humans are a social species is much more than saying that they aggregate or form alliances for the mutual exchange of benefits; they are a group-living, obligately interdependent species.

The psychological implications of just this minimalist scenario are profound. Rather than individuals simply perceiving "reality" and then compromising or conforming their different versions, processes of social verification make reality valid, reliable, and objective (Hardin & Higgins, 1996). However, group-level and individual processes might still conflict. Campbell's (1990a) reinterpretation of Asch's conformity studies illustrates how they might. He argued that these studies should be viewed in terms of the dependence we have on others for knowledge. Accordingly, the Asch studies should be seen as a structural conflict between the respect we have for each other's reports (group-level trust) and our duty to report our observations honestly (individual-level honesty).

So far I have sketched an oversimplified and general model of downward causation and its consequences for grouping. There are various dimensions of groups that can serve as theoretical foci to take a minimalist scenario further and in different directions. For example, Rodseth, Wrangham, Harrigan, and Smuts (1991) developed a theory of social structure based on phylogenetic data about mating patterns. Bugental (2000) and Fiske (1991) theorized domains of relationships, and Flinn and colleagues (2005) and Kurzban and Leary (2001), coali-

tion formation. The following scenario differs from these in that it draws on considerations of body, group size, and task.

The Evolution of Human Coordination: Core Group Configurations

The idea of so-called magic numbers for functional subdivisions of groups appears mostly in anthropology (Lee & DeVore, 1968) but is also mentioned independently in other literatures (Sale, 1980). Dunbar (1993) identifies three groupings drawing partly on ethnographic records of contemporary hunter–gatherers. Conversation groups of about four individuals evolved as a substitute for grooming as a means to service social relationships; "overnight camp groups" or bands with a mean size of 38; and "tribes" with a mean size of 1,150. An intermediate grouping of 150, the predicted group size based on the covariation between primate group size and neocortical size, is proposed as the cognitive limitation on the number of relationships an individual can personally maintain. With the exception of tribes (which are not face-to-face groups), Dunbar's groupings approximate the usual divisions in anthropology.

Two recent reviews of foraging peoples concluded that there was only support for one consistently recurrent group size, but these differed, with one approximating a small work group (Binford, 2001) and the other a small band (Kelley, 1995). Although there is some variation in subgroup size and in what subgroupings are called (e.g., a four- to six-person group may be a hunting group or a gossip group), various models and ethnoarchaeological data (Hassan, 1981) roughly converge on face-to-face groupings shown in Table 1.1, which adds the dyad as a subgrouping. What is significant about anthropology's magic numbers is that they can map on to considerations of body and task. Only one person at a time can hold an infant, and feeding one ensures a specific and constant view of the human face. About four or five people are the most that can closely view a small object in the palm of the hand or examine an animal track on the ground.

Table 1.1, it should be remembered, is an idealized model, a fiction, doing theoretical service. However, it is not a complete departure from human possibility. In their ethnoarchaeological research on the sociospatial organization and decision-making processes of subarctic Chipewyan Indians, Jarvenpa and Brumbach (1988) identified three recurrent spatiotemporal phases of organization: a concentrated summer band of about 300 individuals, winter staging and domestic settlements of 4–10 related families, and, from the latter groups, dispersed bands of hunters, usually two to three men, but also nuclear families or older husband–wife pairs. In their view, the decision-making dynamics of subgroupings is implicit structural organization. One implicit "rule" is a sine qua non for a conjugal pair or nuclear family to ally itself with any cluster of 20–50 people where they had kin in order to survive the long frozen winter.

The congruence between subgroup size, body, and task suggests the hypothesis of four core group configurations as the topography of the selective environment

TABLE 1.1. Core Configurations

Core configuration[a]	Group size	Modal tasks	Proper function
Dyad	2	Sex, infant interaction with older children and adults	Microcoordination
Work/family group	5	Foraging, hunting, gathering, direct interface with habitat	Distributed cognition
Deme (band)	30	Movement from place to place, general processing and maintenance, work group coordination	Shared construction of reality (includes indigenous psychologies), social identity
Macrodeme (macroband)	300	Seasonal gathering, exchange of individuals, resources, and information	Stabilizing and standardizing language

Note. From Caporael (1995). Copyright 1995 by Linnda R. Caporael. Reprinted by permission.
[a]Core configurations are a function of *both* size and task. Except for dyads, these numbers should be considered as modal estimates.

for humans. These are the dyad, task group, deme (or band), and macrodeme (or macroband), organized as a nested hierarchy, or *demic* structure (Hull, 1988). A core configuration is the *joint* function of group size and activity. Configurations provide a context for the evolution of adaptive social cognitive mechanisms. Each group configuration affords functional possibilities and coordination problems that do not exist at any other level. Table 1.1 lists the configurations, along with an approximate group size, examples of modal tasks for the configuration, and an example of a proper function that could have evolved given the configuration.

If uniquely human sociocognitive processes evolved (or earlier building blocks were reorganized) for dynamic coordination in core configurations, then traces of the mind's natural environment and the processes that evolved in that context should persist despite substantial differences in large-scale social organization (Caporael & Baron, 1997). Core configurations and associated processes are hypothesized to recur ontogenetically as a developmental system, in day-to-day life, and presumably in human evolutionary history. There are, of course, other types of human groupings, but they are not necessarily "core." Also, we would expect that a variety of different mechanisms—not just single mechanisms—would evolve because any configuration is suitable for a variety of recurrent tasks.

The smallest configuration, the dyad, is the most ancient of configurations, minimally necessary for all forms of internal fertilization (although mechanisms across species will vary). Among humans (and perhaps primates more generally), dyads are probably evolutionarily significant not so much because new capacities appear but because this configuration functions in (and is influenced by) the initial social organization and entrainment of biological clocks, rhythmicity, and temporal patterning (Jones, 1976; Jones & Boltz, 1989; McGrath & Kelly, 1986). Sometimes this synchrony of movement—or microcoordination—maintains "twoness," even when it is a problem. Imagine two strangers walking toward each other in a direct line. They dance and jerk trying to avoid bumping into each other until one or both manages to

break the coordination. Learning motor tasks from games such as golf or tennis to scientific tasks such as blowing a pipette or using a complex piece of equipment are frequently hands-on jobs.

The earliest cultural overlap between generations occurs in early development with joint attention, shared reality, and storytelling (Higgins, 2005; Nelson, 1993; Tomasello, 1999). Dyadic interaction is generatively entrenched in development, meaning that the dyad's proper functions are necessary conditions for processes that develop later (Wimsatt, 1999). Higgins (2005) argues that self-consciousness begins in dyadic interaction, and Nelson's (2003) work on children's talk to themselves before going to sleep shows them weaving emerging self-concepts against a background of cultural experiences. Some research suggests that adult "dyadic selves" exist as cognitive representations in which concepts of the self and other overlap (Aron, Aron, Tudor, & Nelson, 1991), as transactive memory systems (Wegner, 1986; Wegner, Raymond, & Erber, 1991), and in the coordination of bodily motion or interactional synchrony (Newbern, Dansereau, & Pitre, 1994). Interpersonal networks, such as friendship groups, also engage relational identification processes, as extensions of the basic dyadic unit (Brewer & Caporael, 2006b).

More than other configurations, the work/family or task group appears to be a major point of sense making of the material environment. It affords possibilities for distributed cognition; this means that cognitive tasks such as perception, classification, inference, and contextually cued responses are distributed over group members, particularly when the group is confronted with ambiguous or anomalous environmental information (Cole, Hood, & McDermott, 1980). In the idealized hunter–gatherer group, a task group might be a foraging party interpreting signs of animal movements over a landscape. A modern example would be control tower personnel at airports interpreting signs of possible danger from ambiguous signs on a radar screen. Group size alone does not determine a configuration. Consider a collection of five strangers in an elevator. Under typical conditions, they are an aggregation; each person is an individual ab-

sorbed in his or her own thoughts. If the elevator gets stuck between floors, the same five people form a task group configuration. They jointly explore the opportunities for putting the elevator in motion, recall previous incidences of and solutions for coping with stuck elevators, and point to buttons (shared attention) that might suggest different possibilities of danger and safety. For cognition to be functionally distributed in small groups requires not only a (somewhat) shared focus of attention and representation or schema but also a division of cognitive labor. Wegner (1995) points out that there can be progressive differentiation of memory "specialties" in a group over time. Liang, Moreland, and Argote (1995) found groups trained together to assemble radios outperformed individually trained subjects tested in newly formed groups. Group training enhanced not only recall about assembly procedures but also specialization for remembering distinct aspects of the procedure and trust in one another's knowledge about the task.

The task group is also a primary site for the repeated assembly of culture between generations, such as learning subsistence modes in hunter–gatherer cultures or polite dinner manners in modern cultures. Vygotsky (1978) coined the phrase "zone of proximal development" to describe how children participate with adults in activities slightly beyond the young learner's competence. This zone is a dynamic cognitive region of heightened responsiveness to the tools, skills, and practices in a culture, which children must learn to participate fully as adults (Rogoff, 1990, 2003).

The terms "deme" and "macrodeme" in Table 1.1 are used to indicate a greater generality than the anthropological terms "band" and "macroband" (although these are used here when the anthropological context is intended). Biologists use *deme* to refer to a breeding population; however, it derives from the Greek term, *demos*, which refers to a political unit defined by geography rather than kin ties. Hull (1988) used the term "demic structure" to refer to the organization of scientific communities, which is similar to the structure in Table 1.1. Demes are the locus of practical skills and common knowledge, some of which may be mythical, some of which may be acutely attuned to local conditions, from detailed knowledge of other people to the local ecology. It is also a locus for articulated social identity, "we-groupness" communicated in terms of stories and songs and the basic economic unit, the first configuration that can be self-sustaining for survival and childrearing (but not sexual reproduction). It is also the staging ground for domestic life, including task group coordination, local norm enforcement, and cooperative alliances, which are the basis for fissioning when the community exceeds resources or is fractured by conflict (Olsen, 1987).

In our hunter–gatherer fiction (and sometimes in real-life) related bands meet seasonally, forming macrobands that exchange marriage partners and disgruntled deme members, gifts, and information and perform rituals, competitive games, and stories. Macrobands are generally (and often loosely) related by common history, origin accounts, customs, and, most enduringly, language. They also complete the cycle of biological and social reproduction. A Monte Carlo simulation of paleodemographics by Wobst (1974, cited in Hassan, 1981) indicates that about 175–475 people, or seven to nineteen 25-person bands, are needed to maintain genetic viability by providing mates for members reaching sexual maturity in a population. Macrobands are also historically transitional. They were probably seasonal in the prehistoric past because of limitations of resources, but as agriculture took root, settled macrodemes simply became settlements. Macrobands are rare in the modern world, but macrodemes with analogous group structures are discernible. For example, scientific conferences are often seasonal meetings where information and young people are exchanged, and where the standardization and stabilization of distinctive terminology and the reaffirmation of group identity occurs.

The most important modern function of macrodemes is probably the construction of symbolic group identity (Brewer & Gardner, 1996), shared reality (Hardin & Higgins, 1996), social representations (Deaux & Philogène, 2001), stereotypes (Fiske, 1993; Steele & Aronson, 2003), artifacts (Norman, 1990), and the gathering of them all into worldviews, paradigms, or *social imaginaries*. Taylor (2002) writes of the social imaginary as the way that ordinary people imagine their social surroundings, a collectively shared landscape that makes common practices normative and possible and provides a widely shared sense of legitimacy. Imaginaries are the epistemic products of groups (Kruglanski, Pierro, Mannetti, & De Grada, 2006). For some peoples, the stars in the sky are holes in the floor of heaven; for others, stars are gaseous balls of flame. In both cases, the sensory experience of stars is presumably the same, but the explanatory paradigm differs.

There are a few general points to be made about the core configurations model. First, core configurations repeatedly assemble, in evolutionary time, in ontogeny, and in daily life. As infants develop, their widening scope of interaction increases demands for reciprocity, skills, memory, social judgment, and so on. Second, core configurations differ in how deeply entrenched they are (Wimsatt, 2001). Dyads are deeply entrenched; a change in their evolved functions should predict poor developmental outcomes. In contrast, macrodemes are shallowly entrenched and are relatively easy to modify. Second, humans have made dramatic changes in their lifestyles over the past 10 millennia, and especially in the last 300 years. Clearly the functions that evolve and develop in core configurations are capable of being extended, combined, and used in new domains. For example, a heart surgery team combines microcoordination and distributed cognition. Third, technology can also provide bridges between the functions of configurations. A group of 500 people given an order to march on a football field are likely to clump and straggle, but if a rousing marching song is broadcast, they can hardly avoid keeping time. In fact, artifacts are extremely important, not only for coordinating cognition and behavior (Hutchins, 1996) but also for their role in creating value-saturated environments (Berkowitz, 1997; Hodges & Baron, 1992).

Fourth, core configurations and their psychological correlates can operate independently of the demic structure in which they are hypothesized to evolve. This independence is especially characteristic of urban life where there may be multiple, cross-cutting groups, many assuming the functions of different demic structures. Demic structure is characteristic of certain bureaucratic institutions (e.g., military and clergy) and systems with self-organizing tendencies, like science and the firm, although more often corporate and bureaucratic institutions are likely to suppress self-organizing tendencies of lower-level groups (Campbell, 1982).

In summary, the model of core configurations is built on a set of constraints. It includes an evolutionary framework suitable to psychological levels of analysis; a vocabulary, repeated assembly, for grasping recurrence without making genetic commitments or creating the illusion that a phenomenon of interest has been explained, and an awareness of the limitations of its fictions. Although behavioral data about the past are limited, they do warrant an assumption that humans evolved to be obligately interdependent. Within that framework, there are a variety of ways in which individualism and interdependence may be played out, and the core configurations model is one of these. It organizes a range of experimental findings and lends itself to deriving additional hypotheses in diverse areas of psychology, including development, social cognition, and cultural psychology. It also organizes existing research in social psychology and generates hypotheses that can be tested, including tests of its assumptions through carefully designed experimental and naturalistic fieldwork.

CONCLUSION:
INVENTING HUMAN NATURE

In the 1970s, sociobiology burst onto the scene and into the living room with great fanfare, threatening to "cannibalize" the social sciences, psychology, and ethics (Wilson, 1975). Its proponents claimed that neo-Darwinism would explain self-interest, altruism, aggression, cooperation, knowledge, conflict, reciprocity, and sympathy among other human attributes (Hamilton, 1975; Trivers, 1971; Wilson, 1978). A rarely mentioned feature of this research is that most theorists failed to cite the empirical literature in psychology. What then, was being explained? *Not* the illuminating, albeit counterintuitive, data respected by psychologists (Ross & Nisbett, 1993). Sociobiologists offered the "gene's-eye view" of a widely shared and understood folk psychology. Everyday human experience served as circle of inspiration and evidence for easily decoded scientific categories, beginning with money (resource acquisition), sex (mating), and status (dominance hierarchy) and moving onto the "terrible twos" (weaning conflict), divorce (mate desertion), rape (forced copulation), adolescent rebellion (parent–offspring conflict), and declining reproductive value (menopause). Darwinism itself was tamed (Greenwood, 1984) by a biological rendition draped on a story skeleton (Schank & Abelson, 1995) from Christian theology:

Like the soul, the immortal genes are the essence of the individual, whereas the body, vessel of the soul or vehicle of the genes, is transitory and ephemeral (Caporael, 1994; Dawkins, 1976; Oyama, 1985). God "favors" and "acts" on preferred individuals and natural selection does the same; man was cast out of the Garden of Eden, a state of primitive innocence, and has "lost his natural environment." The meaning implications of this view of Darwinism are hardly lost on the brightest of students (Brem, Ranney, & Schindel, 2003).

In this chapter, I have adopted a view of evolution, psychology, and culture as leaning on each other for inspiration, validation, correction, and interpretation. None of these areas can be reduced to the other because human biology, psychology, and culture coevolved. Each research domain is essential for its contributions to understanding the evolution of human mental systems. Darwinism, conceived at an appropriate level of analysis, provides a theoretical framework that constrains and guides hypothesis development. Social psychology has an experimental methodology that is particularly effective for identifying deviations from conventional wisdom and raising our "discomfort index" (Fiske, 2003) with counterintuitive findings. Cultural psychology, in addition to its range of observational and interpretive methods, trains us to focus on and understand the construction of meaning (Bruner, 1986, 1990), which helps to distinguish evolution as it is used in scientific and social domains. Understanding both domains is both a goal and a significant contribution of social psychology for understanding humanity past, present, and future.

ACKNOWLEDGMENTS

NOTES

1. Caporael (2004) has a more detailed review based largely on Klein (1999). Other resources are Foley (1987, 1995), Tattersall (1995, 1999), and Potts (1996). Potts (1998) offers a brief primer particularly detailed with respect to climatic change. Other primers are Tattersall (2000) and Lahr and Foley (1998).
2. Most readers know *hominins* by the earlier nomenclature, *hominids*. The new taxonomic status of humans includes the australopithecines and *Homo*. Hominids now includes the hominins, plus the great apes and extinct members of their lineages.

REFERENCES

Agre, P. (1997). *Computation and human experience*. Cambridge, UK: Cambridge University Press.
Aiello, L. C., & Wells, J. C. K. (2002). Energetics and the evolution of the genus *Homo*. *Annual Review of Anthropology*, *31*, 323–338.
Alexander, R. D. (1989). Evolution of the human psyche. In P.

Mellars & C. Stringer (Eds.), *The human revolution* (pp. 455–513). Princeton, NJ: Princeton University Press.

Andrews, P. W., Gangestad, S. W., & Matthews, D. (2002). Adaptationism—How to carry out an exaptationist program. *Behavioral and Brain Science*, *25*, 489–553.

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including other in the self. *Journal of Personality and Social Psychology*, *60*, 241–253.

Avital, E., & Jablonka, E. (2000). *Animal traditions: Behavioural inheritance in evolution*. New York: Cambridge University Press.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Bechtel, W. (2003). Modules, brain parts, and evolutionary psychology. In S. J. Scher & F. Rauscher (Eds.), *Evolutionary psychology: Alternative approaches* (pp. 211–227). Dordrecht, The Netherlands: Kluwer.

Berkowitz, L. (1997). Some thoughts extending Bargh's argument. In R. S. Wyer, Jr. (Ed.), *The automaticity of everyday life* (pp. 83–94). Mahwah, NJ: Erlbaum.

Bernieri, F. J., Davis, J. M., Robert, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin*, *20*, 303–311.

Binford, L. (1972). Archeological systematics and the study of culture process. In L. Binford (Ed.), *An archeological perspective* (pp. 195–207). New York: Seminar.

Binford, L. (2001). *Constructing frames of reference: An analytical method for archaeological theory building using hunter–gatherer and environmental data sets*. Berkeley: University of California Press.

Boinski, S., & Garber, P. A. (Eds.). (2000). *On the move: How and why animals travel in groups*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.

Brandon, R. N. (1990). *Adaptation and environment*. Princeton, NJ: Princeton University Press.

Brem, S. K., Ranney, M., & Schindel, J. (2003). Perceived consequences of evolution: College students perceived negative personal and social impact in evolutionary theory. *Science Education*, *87*, 181–206.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Brewer, M. B. (1997). On the social origins of human nature. In C. McGarty & A. Haslam (Eds.), *The message of social psychology* (pp. 54–62). Oxford, UK: Blackwell.

Brewer, M. B. (2004). Taking the social origins of human nature seriously: Toward a more imperialist social psychology. *Personality and Social Psychology Review*, *8*, 107–113.

Brewer, M. B. (2005). Reaching out: Across disciplines, across cultures. *Journal of Experimental Social Psychology*, *41*, 217–219.

Brewer, M. B., & Caporael, L. R. (2006a). An evolutionary perspective on social identity: Revisiting groups. In J. A. Simpson, M. Schaller, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 143–161). Philadelphia: Psychology Press.

Brewer, M. B., & Caporael, L. R. (2006b). Social identity motives in evolutionary perspective. In R. Brown & D. Capozza (Eds.), *Social identities: Motivational, emotional, cultural influences* (pp. 135–152). Philadelphia: Psychology Press.

Brewer, M. B., & Gardner, W. (1996). Who is this "We"? Levels of collective identity and self representation. *Journal of Personality and Social Psychology*, *71*, 83–93.

Brewer, M. B., & Roccas, S. (2001). Individual values, social identity, and optimal distinctiveness. In C. Sedikedes & M. B. Brewer (Eds.), *Individual self, relational self, collective self* (pp. 219–237). Philadelphia: Psychology Press.

Bruner, J. S. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Bugental, D. B. (2000). Acquisition of the algorithms of social life: A domain-based approach. *Psychological Bulletin*, *126*, 187–219.

Buss, D. M. (1999). *Evolutionary psychology*. Boston: Allyn & Bacon.

Buss, L. W. (1987). *The evolution of individuality*. Princeton, NJ: Princeton University Press.

Byrne, R. W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence*. Oxford, UK: Clarendon Press.

Campbell, D. T. (1974). "Downward causation" in hierarchically organized biological systems. In F. Ayala & T. Dobzhansky (Eds.), *Studies in the philosophy of biology* (pp. 179–186). London: Macmillan.

Campbell, D. T. (1982). Legal and primary-group social controls. *Journal of Social and Biological Structures*, *5*, 431–438.

Campbell, D. T. (1983). The two distinct routes beyond kin selection to ultrasociality: Implications for the humanities and social sciences. In D. L. Bridgeman (Ed.), *The nature of prosocial development* (pp. 11–41). New York: Academic Press.

Campbell, D. T. (1990a). Asch's moral epistemology for socially shared knowledge. In I. Rock (Ed.), *The legacy of Solomon Asch* (pp. 39–55). Hillsdale, NJ: Erlbaum.

Campbell, D. T. (1990b). Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology. In E. G. Tobach (Ed.), *Scientific methodology in the study of mind: evolutionary epistemology* (pp. 1–15). Hillsdale, NJ: Erlbaum.

Caporael, L. R. (1994). Of myth and science: Origin stories and evolutionary scenarios. *Social Science Information*, *33*, 9–23.

Caporael, L. R. (1995). Sociality: Coordinating bodies, minds and groups. *Psycholoquy [Online serial]*, *6*(1), http://www.cogsci. soton.ac.uk/cgi/psyc/newpsy?6.01.

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*, 276–298.

Caporael, L. R. (2001a). Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology*, *52*, 607–628.

Caporael, L. R. (2001b). Parts and wholes: The evolutionary importance of groups. In C. Sedikides & M. B. Brewer (Eds.), *Individual self, relational self, and collective self* (pp. 241–258). Philadelphia: Psychology Press.

Caporael, L. R. (2003). Repeated assembly. In S. Schur & F. Rauscher (Eds.), *Alternative approaches to evolutionary psychology* (pp. 71–90). Dordrecht, The Netherlands: Kluwer.

Caporael, L. R. (2004). Stones and bones: Selection for sociality. *Journal of Cultural and Evolutionary Psychology*, *2*, 195–211.

Caporael, L. R., & Baron, R. M. (1997). Groups as the mind's natural environment. In J. Simpson & D. Kenrick (Eds.), *Evolutionary social psychology* (pp. 317–343). Hillsdale, NJ: Erlbaum.

Caporael, L. R., & Brewer, M. B. (1995). Hierarchical evolutionary theory: There is an alternative, and it's not creationism. *Psychological Inquiry*, *6*, 31–34.

Caporael, L. R., Dawes, R. M., Orbell, J. M., & van de Kragt, A. J. C. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences Behavioral and Brain Sciences*, *12*, 683–739.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.

Cole, M. (1997). *Culture and cognitive science* [Electronic version]. Retrieved April 1, 2006, from http://lchc.ucsd.edu/People/ Localz/MCole/santabar.html.

Cole, M., Hood, L., & McDermott, R. (1980). Ecological niche picking. In U. Neisser (Ed.), *Memory observed* (pp. 366–373). New York: Freeman.

Crawford, C. B., & Anderson, J. L. (1989). Sociobiology: An envi-

ronmentalist discipline? *American Psychologist American Psychologist*, *44*, 1449–1459.

Darwin, C. (1981). *The descent of man and selection in relation to sex*. Princeton, NJ: Princeton University Press. (Original work published 1871)

Dawkins, R. (1976). *The selfish gene*. New York: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype: The long reach of the gene*. New York: Oxford University Press.

Deaux, K., & Philogène, G. (Eds.). (2001). *Representations of the social: Bridging theoretical traditions*. Oxford, UK: Blackwell.

DeKay, W. T., & Buss, D. M. (1992). Human nature, individual differences, and the importance of context: Perspectives from evolutionary psychology. *Current Directions in Psychological Science*, *1*, 184–189.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1–40.

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, *16*, 681–735.

Durham, W. H. (1976). The adaptive significance of cultural behavior. *Human Ecology Human Ecology*, *4*, 89–121.

Eldredge, N., & Grene, M. (1992). *Interactions: The biological context of social systems*. New York: Columbia University Press.

Endler, J. (1986). *Natural selection in the wild*. Princeton, NJ: Princeton University Press.

Fajen, B. R., & Devaney, M. C. (2006). Learning to control collisions: The role of perceptual attunement and action boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 300–313.

Fajen, B. R., & Turvey, M. T. (2003). Perception, categories, and possibilities for action. *Adaptive Behavior*, *11*(4), 276–278.

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations*. New York: Free Press.

Fiske, A. P. (2000). Complementarity theory: Why human social capacities evolved to require cultural complements. *Personality and Social Psychology Review*, *4*, 76–94.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.

Fiske, S. T. (2003). The discomfort index: How to spot a really good idea whose time has come. *Psychological Inquiry*, *14*, 203–208.

Flinn, M. V., Geary, D. C., & Ward, C. V. (2005). Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence. *Evolution and Human Behavior*, *26*, 10–46.

Foley, R. (1987). *Another unique species: Patterns of human evolutionary ecology*. New York: Longman/Wiley.

Foley, R. A. (1995). *Humans before humanity: An evolutionary perspective*. Oxford, UK: Blackwell.

Geary, D. C. (2005). *The origins of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*, 153–172.

Gottlieb, G. (2002). Developmental–behavioral initiation of evolutionary change. *Psychological Review*, *109*, 211–218.

Gould, S. J. (1980). Is a new and general theory of evolution emerging? *Paleobiology*, *6*, 119–130.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *70*(5), 701–721.

Green, M. C., Strange, J. J., & Brock, T. C. (Eds.). (2002). *Narrative impact: Social and cognitive foundations*. Mahwah, NJ: Erlbaum.

Greenwood, D. (1984). *The taming of evolution: The persistence of nonevolutionary views in the study of humans*. Ithaca, NY: Cornell University Press.

Griffiths, P. E., & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *Journal of Philosophy*, *91*, 277–304.

Gurven, M. (2004). To give and to give not: An evolutionary ecology of human food transfers. *Behavioral and Brain Science*, *27*(4), 543–583.

Hamilton, W. D. (1975). Innate social aptitudes of man. In R. Fox (Ed.), *Biosocial anthropology* (pp. 133–155). New York: Wiley.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition. Vol. 3. The interpersonal context* (pp. 28–84). New York: Guilford Press.

Hassan, F. A. (1981). *Demographic archaeology*. New York: Academic Press.

Hendriks-Jansen, H. (1996). *Catching ourselves in the act*. Cambridge, MA: MIT Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H., et al. (2005). Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*, 795–855.

Higgins, E. T. (2005). Humans as applied motivation scientists: Self-consciousness from "shared reality" and "becoming." In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 157–173). New York: Oxford University Press.

Hodges, B. H., & Baron, R. M. (1992). Values as constraints on affordances: Perceiving and acting properly. *Journal for the Theory of Social Behaviour*, *22*, 263–294.

Hull, D. L. (1988). *Science as a process*. Chicago: University of Chicago Press.

Hutchins, E. (1996). *Cognition in the wild*. Cambridge, MA: MIT Press.

Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral and symbolic variation in the history of life*. Cambridge, MA: MIT Press.

Janicki, M. G. (1998). Evolutionary approaches to culture. In C. Crawford & D. L. Krebs (Eds.), *Handbook of evolutionary psychology* (pp. 163–207). Mahwah, NJ: Erlbaum.

Jarvenpa, R., & Brumbach, H. (1988). Socio-spatial organization and decision-making processes: Observations from the Chipewyan. *American Anthropologist*, *90*, 598–618.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention and memory. *Psychological Review*, *83*, 323–355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.

Kameda, T., & Tindale, R. S. (2006). Groups as adaptive devices: Human docility and group aggregation mechanisms in evolutionary context. In J. A. Simpson, M. Schaller, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 317–341). Philadelphia: Psychology Press.

Kelley, R. L. (1995). *The foraging spectrum: Diversity in hunter–gatherer lifeways*. Washington, DC: Smithsonian Institution Press.

Kitcher, P. (1985). *Vaulting ambition: Sociobiology and the quest for human nature*. Cambridge, MA: MIT Press.

Klein, R. G. (1999). *The human career: Human biological and cultural origins*. Chicago: University of Chicago Press.

Kruglanski, A. W., Pierro, A., Mannetti, L., & De Grada, E. (2006). Groups as epistemic providers: Need for closure and the unfolding of group centrism. *Psychological Review*, *113*, 84–100.

Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin*, *127*, 187–208.

Kusimba, S. B. (2005). What is a hunter–gatherer? Variation in the archaeological record of eastern and southern Africa. *Journal of Archaeological Research*, *13*(4), 337–366.

Lahr, M. M., & Foley, R. A. (1998). Towards a theory of modern human origins: Geography, demography, and diversity in recent human evolution. *Yearbook of Physical Anthropology*, *41*, 137–176.

Landau, M. (1984). Human evolution as narrative. *American Scientist*, *72*, 262–268.

Langer, E. J. (1989). *Mindfulness*. New York: Addison-Wesley.

Lave, J. (1988). *Cognition in practice*. New York: Cambridge University Press.

Leacock, E. B. (1981). *Myths of male dominance*. New York: Monthly Review Press.

Lee, R. B., & DeVore, I. (Eds.). (1968). *Man the hunter*. Chicago: Aldine.

Levins, R., & Lewontin, R. (1985). *The dialectical biologist*. Cambridge, MA: Harvard University Press.

Li, S.-C. (2003). Biocultural orchestration of developmental plasticity across levels: The interplay of biology and culture in shaping the mind and behavior across the life span. *Psychological Bulletin*, *129*, 171–194.

Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin*, *21*, 384–393.

Lickliter, R., & Honeycutt, H. (2003). Developmental dynamics: Toward a biologically plausible evolutionary psychology. *Psychological Bulletin*, *129*, 839–835.

Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, *123*(1), 3–32.

Lloyd, E. A. (1999). Evolutionary psychology: The burdens of proof. *Biology and Philosophy*, *14*, 211–233.

Maynard Smith, J., & Szathmáry, E. (1995). *The major transitions in evolution*. New York: Freeman.

McGrath, J. E., & Kelly, J. R. (1986). *Time and human interaction: Toward a social psychology of time*. New York: Guilford Press.

Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavaviorial and Brain Science*, *29*, 329–383.

Millikan, R. G. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualization situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246–268.

Moser, S. (1998). *Ancestral images: The iconography of human origins*. Ithaca, NY: Cornell University Press.

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science*, *4*, 7–14.

Nelson, K. (2003). Self and social functions: Individual autobiographical memory and collective narrative. *Memory*, *11*, 125–136.

Newbern, D., Dansereau, D. F., & Pitre, U. (1994, June 29–July 3). *Ratings of synchrony in cooperative interaction predict cognitive performance*. Paper presented at the annual convention of the American Psychological Society, Washington, DC.

Norman, D. A. (1990). *The design of everyday things*. New York: Doubleday. (Originally published as *Psychology of Everyday Things*, Basic Books, 1988)

Norris, K. S., & Schilt, C. R. (1988). Cooperative societies in three-dimensional space: On the origins of aggregations, flocks, and schools, with special reference to dolphins and fish. *Ethology and Sociobiology*, *9*, 149–179.

Núñez, R., & Freeman, W. J. (Eds.). (1999). *Reclaiming cognition: The primacy of action, intention and emotion*. Thorverton, UK: Imprint Academic.

Olsen, C. L. (1987). The demography of colony fission from 1878–1970 among the Hutterites of North America. *American Anthropologist*, *89*(4), 823–837.

Oyama, S. (1985). *The ontogeny of information*. New York: Cambridge University Press.

Pollock, G. B. (1989). Suspending disbelief—of Wynne-Edwards and his reception. *Journal of Evolutionary Biology*, *2*, 205–221.

Potts, R. (1996). *Humanity's descent: The consequences of ecological instability*. New York: Morrow.

Potts, R. (1998). Environmental hypothesis of hominin evolution. *Yearbook of Physical Anthropology*, *41*, 93–136.

Pulliam, H. R., & Caraco, T. (1984). Living in groups: Is there an optimal group size? In J. R. Krebs & N. B. Bavies (Eds.), *Behavioural ecology* (pp. 122–147). London: Blackwell.

Rodseth, L., Wrangham, R. W., Harrigan, A. M., & Smuts, B. B. (1991). The human community as a primate society. *Current Anthropology*, *32*, 221–254.

Rogoff, B. (1990). *Apprenticeship in thinking: cognitive development in social context*. New York: Oxford University Press.

Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.

Rosch, E. (1977). Human categorization. In N. Warren (Ed.), *Studies in cross-cultural psychology* (pp. 1–49). New York: Academic Press.

Rose, M. R., & Lauder, G. V. (Eds.). (1996). *Adaptation*. New York: Academic Press.

Ross, L., & Nisbett, R. E. (1993). *The person and the situation: Perspectives of social psychology*. Philadelphia: Temple University Press.

Sale, K. (1980). *Human scale*. New York: Coward, McCann & Geoghegan.

Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 8, pp. 1–86). Hillsdale, NJ: Erlbaum.

Scher, S. J. (2004). A Lego model of the modularity of mind. *Journal of Cultural and Evolutionary Psychology*, *2*, 249–259.

Schmitt, D. P., & Pilcher, J. J. (2004). Evaluating evidence of psychological adaptation: How do we know one when we see one? *Psychological Science*, *15*, 643–649.

Sedikides, C., & Skowronski, J. J. (1997). The symbolic self in evolutionary context. *Personality and Social Psychology Review*, *1*, 80–102.

Simpson, J. A., Schaller, M., & Kenrick, D. T. (Eds.). (2006). *Evolution and social psychology*. Philadelphia: Psychology Press.

Smith, B. H. (2006). Scandalous truth: Science, truth and the human. Durham, NC: Duke University Press.

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Steele, C. M., & Aronson, J. (2003). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Symons, D. (1992). On the use and misuse of Darwinism in the study of human behavior. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 137–159). New York: Oxford University Press.

Szathmáry, E., & Maynard Smith, J. (1995). The major evolutionary transitions. *Nature*, *374*, 227–232.

Tattersall, I. (1995). *The fossil trail*. New York: Oxford University Press.

Tattersall, I. (1999). *Becoming human: Evolution and human uniqueness*. New York: Harvest Books.

Tattersall, I. (2000). Paleoanthropology: The last half-century. *Evolutionary Anthropology*, *9*(1), 56–62.

Taylor, C. (2002). Modern social imaginaries. *Public Culture*, *14*, 91–124.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). New York: Oxford University Press.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York: Springer-Verlag.

Wegner, D. M. (1995). A computer network model of human transactive memory. *Social Cognition*, *13*, 319–339.

Wegner, D. M., Raymond, P., & Erber, R. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, *61*, 923–929.

Williams, G. C. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.

Wilson, D. S., & Sober, E. (1994). Re-introducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, *17*, 585–654.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.

Wilson, E. O. (1978). *On human nature*. Cambridge, MA: Harvard University Press.

Wimsatt, W. C. (1999). Generativity, entrenchment, evolution, and innateness. In V. Hardcastle (Ed.), *Biology meets psychology* (pp. 139–179). Cambridge, MA: MIT Press.

Wimsatt, W. C. (2001). Generative entrenchment and the developmental systems approach to evolutionary processes. In S. Oyama, R. Gray, & P. Griffiths (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 219–237). Cambridge, MA: MIT Press.

Wimsatt, W. C., & Griesemer, J. R. (in press). Re-producing entrenchments to scaffold culture: How to re-develop cultural evolution. In R. Sansom & R. N. Brandon (Eds.), *Integrating evolution and development: From theory to practice*. Cambridge, MA: MIT Press.

Zebrowitz, L. A., & Montepare, J. (2006). The ecological approach to person perception: Evolutionary roots of contemporary offshoots. In J. A. Simpson, M. Schaller, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 81–114). Philadelphia: Psychology Press.

Zhang, J., & Patel, V. L. (2006). Distributed cognition, representation and affordance. *Pragmatics and Cognition*, *14*, 331–341.

Visceral and Somatic Indexes
of Social Psychological Constructs
History, Principles, Propositions, and Case Studies

JIM BLASCOVICH
MARK D. SEERY

It has long been customary for physiologists and behavioral scientists to acknowledge the potential
bearing of one discipline on the other, but only recently has there been a deliberate and
conscientious attempt to merge the two into an expanded science of behavior. There is little
question that the study of behavior in the matrix of general biology is now one of the major,
though fledgling, scientific enterprises of the day. A distinct subdiscipline within this general trend
is represented by research combining social–psychological and physiological research techniques in
the study of human interaction.

—SHAPIRO AND CRIDER (1969, p. 1)

PROPHETS AND PIONEERS

The appearance of Shapiro and Crider's seminal hand-book chapter, "Physiological Approaches in Social Psychology," in 1969 marked optimism on the part of the authors (and editors, Gardner Lindzey and Eliot Aronson, in choosing to invite such a chapter) that the study of human behavior, including social behavior, and biology would soon merge into a major interdisciplinary scientific enterprise. These authors implied that until the late 1960s, physiologists and social psychologists stayed mostly on their own side of an epistemological wall, but one they foresaw crumbling, presumably along with the remnants of Descartes's (1680) centuries old metaphysics of mind–body dualism. This was a bold assertion given that psychophysiology itself, a field attempting to unite the disciplines of general psychology and general physiology, was acknowledged by psychophysiologists themselves to be less than a decade old (Fowles, 1975; Sternbach, 1966). However, the assertion was prescient.

At the time they wrote their chapter, Shapiro and Crider noted that neither physiologists (nor psychophysiologists) on one hand nor social psychologists on the other were particularly well versed in the methodologies—to which we add theory and concepts—of

the other. They argued that a lack of cross-disciplinary expertise led investigators trained in physiological techniques to ignore proven methodological controls for social factors and individual differences. Indeed, these factors had been neatly labeled, even by psychophysiologists, as "situational response stereotypy" and "individual response stereotypy," respectively (Lacey & Lacey, 1958), inadvertently casting social psychologists' major categories of interest as uncontrollable noise in human psychophysiological experiments.

Shapiro and Crider were also quick to point out that social psychologists lacked sophistication in physiology. Their review of the relatively sparse social psychophysiology literature at the time indicated that most social psychological researchers using physiological measures apparently assumed that just about any visceral or autonomically controlled (i.e., visceral) response was interchangeable with any other. The clear conclusion was that social psychologists used unitary autonomic measures, particularly ones thought to be under sympathetic control (e.g., heart rate and skin conductance), interchangeably. A consequence of such naive interchangeability was that social psychologists used increases in sympathetically influenced physiological responses as though they were unambiguous measures of psychological states in

which physiological arousal played a theoretical or defining role (anxiety, tension, motivation, etc.). Hence, social psychological researchers who incorporated physiological measures often failed to define arousal constructs explicitly and, as a consequence, failed to consider both the role of physiological processes and social context restrictions when drawing inferences from physiological measures in their studies (factors contributing to a strong basis for inference when measuring physiological responses are discussed in more detail in subsequent sections).

The rise of neo-Jamesianism in the 1960s reinforced the idea that specification of physiological arousal constructs and the distinctions among physiological responses were unnecessary for social psychological researchers. For example, Schachter and Singer's (1962) classic paper, "Cognitive, Social, and Physiological Determinants of Emotional State," was based on the idea that emotions are not differentiated physiologically but, rather, cognitively and occur when a state of general visceral or autonomic "arousal" is labeled as a function of the social context. Valins (1966) went even further, concluding from his false-feedback studies that only the perception of such arousal was necessary, whether achieved veridically or nonveridically via interoceptive or external feedback. It should be noted that even now the search for "autonomic footprints" of basic-level emotions has proven quite difficult (Cacioppo, Klein, Berntson, & Hatfield, 1993). But the lack of such evidence should not be taken as negative evidence. We do know that at least visceral (e.g., Blascovich & Tomaka, 1996) and somatic (e.g., Cacioppo & Petty, 1981a, 1981b) footprints of superordinate emotional affective states have been identified.

Although the promise of the interplay between social psychology and physiology that Shapiro and Crider foresaw has not been fully realized in the 35 years since publication of their chapter, it has not proven an empty one either. Indeed, their seminal chapter played a role in motivating several social psychologists to become sophisticated in physiology and psychophysiology. These pioneers understood that the promise of social psychophysiology would be realized only if cross-disciplinary training took place. Paramount among these social psychophysiologists was John Cacioppo.

A quarter century after the first social psychophysiology handbook chapter, Cacioppo, a social psychologist by training, Gary Berntson, a biological psychologist, and Lew Tassinary, a cognitive psychologist, collaborated on a second handbook chapter, one appearing in the first edition of this handbook.[1] Cacioppo, Berntson, and Crites (1996) titled their chapter "Social Neuroscience: Principles of Psychophysiological Arousal and Response," in anticipation of the integration of social psychology within the broad new field of neuroscience.

These integrators brought sophistication and empirical substance to the merger foreseen by Shapiro and Crider. The sophistication was epistemological, physiological, social psychological, and educational. The substance included the use of physiological indexes of psychological constructs in empirical studies to advance social psychological theory, studies that incidentally provided model studies for others to emulate. Their educational efforts included an edited volume of social psychophysiological research (Cacioppo & Petty, 1983), which, with Waid's *Sociophysiology* (1984), consisted of the first two such contributions; training workshops[2]; and two edited comprehensive handbooks of social psychophysiology (Cacioppo, Berntson, & Tassinary, 2000; Cacioppo & Tassinary, 1990). Furthermore, it is fair to say that these contributions laid the groundwork for the more encompassing integration of social psychology and biology that today is labeled "social neuroscience."

MIND OR BODY VERSUS MIND AND BODY

Social psychophysiology fits within the context of mind–body interactions. Attempts to describe the relationship between the mind and the body, first explicitly labeled the "mind–body problem" by Descartes (1680), have occupied Eastern and Western philosophers of mind and other scholars for nearly 3 millennia and perhaps longer. In both hemispheres, scholars have proffered dualistic (i.e., mind or body) and monistic (mind and body) views. In Western philosophy, some suggest (e.g., Robinson, 2003) that Plato was the first dualist because he argued in *Phaedo* that the intellect is an immaterial part of the soul. Some (e.g., Ryle, 1949) suggest that Aristotle was the first monist or materialist because he argued that the soul is no more than the form or physical nature of one's body. These distinctions are remarkably similar in ancient Eastern philosophy lending even more credence to the universality of the mind–body problem among scholars.

Although other philosophers and scholars certainly addressed the issue, debates over the mind–body problem accelerated during the 17th century with the writing in 1650 and the publication in 1680 of Rene Descartes's *Six Metaphysical Meditations; Wherein It Is Proved That There Is a God. And That Man's Mind Is Really Distinct from His Body*. Although to this day, Descartes is seen by most (e.g., Damasio, 1994) as a strict dualist, separating the study of the mind from the body, he actually posited a singular connection between the mind and the body, one he placed at the pineal gland (Descartes, 1662/1972). During the 17th and 18th centuries, others attacked Descartes's position, including Spinoza (1677) and Leibniz (1695). The 19th century brought arguments from Carpenter (1874) and Lewes (1877) narrowing the mind–body problem to a "mind–brain" problem, but with mental states (e.g., consciousness) treated as epiphenomenal. William James (1890) attacked the notion of the epiphenomenalism of mental states, especially emotions, and essentially returned psychology to the mind–body connection albeit with the nature of the connections left relatively unspecified. James's work paved the way for work by physiologists such as Walter B. Cannon (1928) and Hans Selye (1950), who provided much of the early theoretical foundation for psychophysiology as a subdiscipline.

Thus, the dénouement of mind-body dualism began to accelerate by the beginning of the 20th century. However, the centuries long separation—the pineal gland notwithstanding—would take nearly 100 years to mend meaningfully and at first only in the scientific psychology and biology communities. Several modern scientific landmark studies forced the issue.

In the mid-1970s, Robert Ader and Nicholas Cohen, members of the psychiatry department at the University of Rochester, conducted a classical conditioning study using rats (Ader & Cohen, 1975). They paired water sweetened with saccharin (conditioned stimulus) with an injection of cyclophosphamide, an immunosuppressive agent (unconditioned stimulus). A few days after the conditioning, the rats were injected with an antigen, sheep erythrocytes, which would normally trigger the production of antibodies. However, Ader and Cohen found that rats that were exposed to saccharin for a second time produced lower levels of antibodies when the antigen was injected than rats that only received the initial saccharin pairing. This demonstrated that the immunosuppressive effects of cyclophosphamide had been conditioned to occur with exposure to otherwise benign saccharin.

Although studies (e.g., Solomon & Moos, 1965) of the relationship between psychological states and immunosuppression appeared prior to this one, there was no strong experimental evidence of a connection between the central nervous system and the immune system. Thus, it was Ader and Cohen's (1975) behavioral study that first provided an experimental basis for inferring a causal connection between the two, something that neither biologists nor other psychologists had achieved. Ader labeled the research area that grew out of his work "psychoneuroimmunology" and inspired many others, most notably Janice Kiecolt-Glaser and Ron Glaser (e.g., Glaser et al., 1987; Kiecolt-Glaser & Glaser, 1995), to study how the mind and nervous, endocrine, and immune systems influence one another.

At about the same time that Ader and Cohen published their paper, Milton Friedman and Ray Rosenman, cardiologists in joint practice, published their book, *Type A Behavior and Your Heart* (1974). On the basis of their own clinical and unobtrusive observations of heart patients as well as more formal study, Friedman and Rosenman concluded that a confluence of traits put certain individuals at increased risk for coronary artery disease. These individuals tended to be hard driving, competitive, time urgent, and hostile, a constellation that they labeled the coronary prone or "Type A" behavior pattern, which increased the risk of such individuals for heart disease. Their book was one of the first to close the gap between centuries of intuition and empirical evidence connecting personality and disease. More important for purposes here, the Type A construct stimulated a marriage between personality and social psychology on the one hand and cardiovascular psychophysiology on the other. Many social psychologists began their forays into social psychophysiology because of their interest in the coronary-prone behavior pattern.

These seminal studies also provided evidence that the merger of psychological and biological disciplines and subdisciplines could bear scientific fruit and advance science in ways unavailable to either parent discipline. By the end of the 20th century, major integrative and theoretical statements had been made including *Descartes Error: Emotion, Reason, and the Human Brain* (Damasio, 1994), *How the Mind Works* (Pinker, 1997), and *The Emotional Brain: The Mysterious Underpinnings of Emotional Life* (LeDoux, 1996). In terms of social psychology and physiology, three "social psychophysiological" chapters appeared in the first edition of this handbook (Cacioppo, Berntson, & Crites, 1996; Davison & Pennebaker, 1996; Zillman & Zillman, 1996), confirming Shapiro and Crider's original vision. These chapters as well as other seminal studies and books have also brought much media attention to the links between mind and body, attention that has increased markedly over the last quarter century.

Implications

The retreat from a Cartesian metaphysics of mind–body dualism to a more integrated, monistic view of mind–body relationships has several implications for combining psychological and biological disciplines. First, it requires real and practical acceptance of a background, metaphysical assumption sometimes labeled the "identity thesis" (e.g., Cacioppo & Tassinary, 1990). The identity thesis assumes that all mental and, hence, psychological, states and processes are embodied corporeally. Second, the monistic view implies that understanding bodily responses can inform our understanding of mental states and processes, and that understanding mental states and processes can inform our understanding of bodily responses. Hence, the monistic view implies that psychological (including social psychological) and biological (including physiological) levels of analysis cannot be separated if the mind–body relationship is to be understood. Third, the monistic view argues against biological reductionism as well as pure psychologism in all its forms (e.g., automatonism, cognitivism, and social constructionism). Fourth, it suggests that state-of-the-art training in social psychology should necessarily include training in biological methods generally and physiology and neuroscience more specifically, without abandoning its connections to other disciplines and subdisciplines such as sociology, cognition, and so on.

MIND SCIENCE AS AN OVERARCHING ORGANIZING CONSTRUCT

As Shapiro and Crider envisioned, social psychology and physiology would merge to form one subdiscipline, social psychophysiology, within the more general merger of psychology and biology. At the time, one could have imagined several such binary combinations within the more general merger. Yet, any binary combination of psychological and biological subdisciplines is not exclusive of any other. And combinations need not be bi-

nary. Multiple combinations (e.g., "social psychoneuro-immunology" and "social psychoendocrinology") can result from all sorts of multidisciplinary activities. The psychological–biological combinations have been subsumed somewhat under the label of "life science" for a relatively long time. Furthermore, combinations of psychological and biological subdisciplines have never really been limited to each other, though some of these extrapsychological and extrabiological linkages are newer (e.g., computational neuroscience) than others (e.g., psychosomatic medicine). The number of the linkages between psychology and biology and other disciplines has grown nearly exponentially (e.g., to any number of mathematics and science, engineering, social sciences, and humanities disciplines and subdisciplines).

An overarching label that has emerged for multidisciplinary approaches involving psychology, biology, and the many possible other disciplines is "mind science." Though 100 or even 50 years ago, many would have scoffed at the label and considered it an oxymoron (and some still do). Whether or not the descriptive label "mind science" takes hold, its mere advancement suggests the study of the mind is not limited to the behavioral, social sciences, and humanities but to the natural life sciences as well. Not unexpectedly, some territoriality seems to be involved among scientists claiming their subdiscipline is the "science of the mind" (e.g., Buss, 2004; Robinson-Riegler & Robinson-Riegler, 2004), but most recognize that some value to an overarching label and "mind science," or something akin to it, is likely to stick.

"Mind science" is not a new label. Indeed, the Mind Science Foundation, a nonprofit organization, has been supporting appropriate interdisciplinary research for many decades. However, trying to describe the exact nature of the interdisciplinary relationships underlying mind science is probably an exercise in futility because the underlying multidisciplinary structure is changing and growing too fast. However, the study of the mind has become more of a global multidisciplinary and interdisciplinary enterprise advancing what historically has been a theological, philosophical, and scientific question at least since Plato and probably earlier.

Mind science can be defined as the study of the antecedents, processes, and consequences of mental function in the context of the identity thesis. Mind science may well become the major scientific frontier for the 21st century. Scientific theory and technology have increased in sophistication so that models and tools have become available that allow investigators to move from the more hypothetical to the more truly scientific level of analysis (Harré, 1967) with regard to the mind.

Social Psychology and Mind Science

Multiple linkages exist among social psychology and other disciplines and subdisciplines within the context of mind science. Even the combinations between and among social psychology and biological disciplines are multiple. However, historically, the connection between social psychology and physiology is probably the oldest and most enduring one of a biological nature. Although social psychology and physiology can be relatively easily defined, neither can be as easily described.

Allport's (1935) classic definition of social psychology as the scientific investigation of "how the thoughts, feelings and behaviors of individuals are influenced by the actual, imagined or implied presence of others" (p. 800) remains accurate today. For didactic purposes, the subject matter of social psychological inquiry can be divided into intraindividual and interindividual processes. Historically, general categories of intraindividual processes have included the affective, cognitive, and behavioral ones noted in Allport's definition, while interindividual processes have included social interactional (e.g., group dynamics and prejudice) and social influence (e.g., persuasion) processes.

Clearly, the distinction between intra- and interindividual processes does not represent a realistic dichotomy as the so-called intraindividual processes do not take place in an interindividual vacuum, nor do interindividual processes exist apart from individuals. Consequently, despite the didactic value of these distinctions, limiting research domains to one or another category or subcategory of social psychological processes runs the risk of epiphenomenalism and lack of generalizability. For example, a consideration of motivational processes without integrating affect and cognition will necessarily lack external validity.

As a field, physiology has a much longer history than psychology and can be defined generally as the study of the functions or vital processes of living things, whether animal or plant, whole (e.g., living beings) or part (e.g., visceral organs and cells). Physiology is one (anatomy another) of the oldest of the biological subdisciplines. Given the reign of a dualistic perspective on mind and body that remained stronger in the biological than the psychological sciences, not surprisingly, subdisciplines such as psychophysiology and physiological or biological psychology evolved more within psychology than biology.

Historically, psychophysiology and physiological or biological psychology have been distinguished from each other in two ways. First, psychophysiology involved independent manipulation of psychological variables with measurement of physiological responses as dependent measures whereas physiological psychology involved independent manipulation of physiological variables (e.g., drug infusions and brain lesions) with behavioral responses as dependent measures. Second, psychophysiologists worked predominantly with human participants whereas physiological psychologists worked primary with animals. Thus, it not surprising that social psychologists would have more affinity with psychophysiologists than with physiological psychologists, although this preference is not as strong as it once was.

The subject matter of human physiology pertinent to psychophysiology can be and has been divided into control and operational systems for didactic purposes. In this vein, the major control systems include the central nervous system (CNS) and the endocrine system. Operational systems include the visceral system, the skeletal–
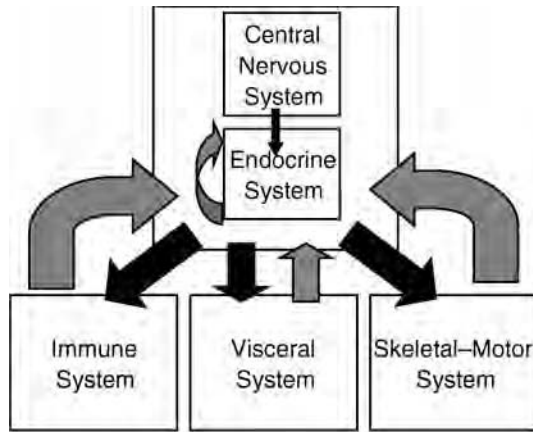
FIGURE 2.1. Cybernetic structure of physiological systems.

motor or somatic system, and the immune system. Clearly, the distinction between control and operational systems is a gross one as even the tiniest of physiological organs (i.e., cells), whether found in the CNS or the visceral system, have both control and operational functions.

The control–operational distinction subscribed to by physiologists and psychophysiologists is based on a cybernetic or feedback model (see Figure 2.1), allowing the continuous operation of bodily processes to deviate from and return to set homeostatic levels. A control system, such as the CNS, operates as a cybernetic model by transmitting operating instructions via efferent neurons and monitoring operational systems via afferent neurons. Similarly, the endocrine system transmits instructions and monitors operational systems via circulating hormones.

The major research questions guiding social psychophysiology, as well as social neuroscience more generally, involve the interaction of intra- and interindividual processes with physiological control and operating systems. Thus, the function or value of a social psychophysiological or neuroscience approach is twofold: understanding social psychological impacts on physiological systems and using physiological measures to understand social psychological processes. Most social psychologists have been more interested in the latter than the former. However, the view endorsed here is that the first function, understanding social psychological impacts on physiological systems, is necessary for the second; that is, to be able to use physiological responses to understand social psychological processes. Here, we focus on such understanding in terms of the interaction among social psychological, control (CNS, endocrine) and operational (i.e., visceral and skeletal–motor) physiological processes.

INFERRING PSYCHOLOGICAL CONSTRUCTS FROM EMBODIMENTS

We advocate a systematic approach to developing, validating, and using peripheral physiological markers of so-

cial psychological constructs. This approach includes a set of what we label and describe in some detail as "first principles" and a set of propositions for increasing the strength of inference drawn from physiological measurement of social psychological constructs. Following a relatively content-free discussion of first principles and propositions and a short discussion of general peripheral physiological measurement issues, we illustrate our approach content-wise by presenting two case studies of the development and validation of physiological markers based on these principles and propositions. One involves visceral and the other involves somatic physiological responses.

First Principles

To measure embodiments of social psychological processes, it is necessary to (1) specify the nature of particular social psychological constructs as precisely as possible; (2) specify the candidate physiological responses used to measure or index the constructs as precisely as possible; and (3) specify a valid or plausible theoretical physiological rationale for linking the constructs and physiological indexes as precisely as possible. Indeed, this logic can and should be used for indexes of any social psychological construct, whether those indexes are subjective (e.g., self-report) or objective (e.g., behavioral or physiological). Furthermore, it should be noted that the order of implementation of the second and third principles is often reversed.

Principle 1: Specifying The Nature of the Social Psychological Construct

Given the categories of processes that social psychologists study (e.g., relationships, emotion, and intergroup interactions), many social psychological concepts or constructs have proven "fuzzy" or difficult to define in social psychology. Part of the definitional difficulty is due to the seemingly ubiquitous use of common language labels for constructs in social psychology (e.g., attitudes, self-esteem, risk taking, compassion, racial prejudice, stigma, love, and liking). Such common language labels can lead researchers and investigators to rely on or be heavily influenced by common language connotations of the construct labels themselves rather than to rely on more precise conceptual analyses and explicit definitions of the constructs (Blascovich & Ginsburg, 1978).

Another conceptual difficulty derives from a tendency to define constructs with a focus on a single domain (e.g., affective, cognitive, or behavioral), sometimes leading to conceptual criteria for indexing these constructs in ways that fail to generalize across domains. For example, attitudes have been defined relatively narrowly depending on whether investigators are concerned with affect (i.e., defined in terms of liking), cognitions (i.e., defined in terms of associations of attitude objects with evaluations in memory), or behavior (i.e., propensity to respond). Hence, a set of physiological responses such as overt or covert facial electromyographic responses based on the muscles controlling smiling and frowning might have a

strong rationale if attitudes are defined as liking but little or no relationship if they are defined in terms of cognitions or behavior.

Such conceptual obfuscation creates several problems. First, it prevents the increasing clarity that comes with debate and refinement of explicit statements defining social psychological constructs. Second, without explicitly defined conceptual criteria, there is no conceptual or face validation check for manipulation of construct relevant independent variables. Third, such obfuscation promotes definitional operationism (Campbell, 1988b), wherein constructs are defined in terms of their measures (e.g., intelligence is what intelligence tests measure, or stress is a particular physiological response), a dangerous road inferentially (i.e., affirmation of the consequent) and consequently theoretically.

Precise specification of what an investigator means by a concept or construct avoids these issues, though it does not guarantee conceptual accuracy and acceptance. However, at least such specification does not finesse the problem of definition in the way that relying on the connotations of a common language label does. Rather, precise specification creates a point of departure for conceptual arguments, refinements, and, ultimately and ideally, conceptual clarity. As long as the conceptual criteria for a construct are stated, labels become less important. For example, one can define "risk taking" as behavior that puts something at stake for the actor, including both the possibility of gain and the possibility of loss, and for which there is a point in time when what is at stake cannot be withdrawn (Blascovich & Ginsburg, 1978). One can choose another label for the construct so defined such as "gambling" or even "X" without loss or change of meaning on the scientific level. And, one can argue the conceptual criteria explicitly.

Principle 2: Specification of Candidate Physiological Indexes

For human neuroscience, including social psychophysiology, we assume that the critical connection between mind and body occurs at the level of the CNS. However, this metaphysical assumption does not reduce the value of physiological responses controlled via peripheral visceral and somatic branches of the peripheral nervous system for understanding social psychological processes. Peripheral embodiments of psychological processes are driven by CNS embodiments of those processes as the latter controls the former. Hence, in theory, any psychological process that can be indexed via peripheral physiological responses should be distinguishable via CNS activity.

However, in many cases, the use of peripheral embodiments holds some practical advantages over CNS embodiments and, at least currently, some technological advantages. First, the costs of recording and scoring peripheral measures are much less than costs for recording and scoring central measures, especially those based on positron emission topography (PET) and functional magnetic resonance imaging (fMRI). Second, peripheral measures are typically less intrusive, though neither need be invasive. Third, compared to the technology for CNS measures (e.g., fMRI), the technology for peripheral measures, even ones requiring relatively sophisticated technology, does not restrict postures and movements, including speech, facial expressions, and so on, making even ambulatory measurement possible.

Fourth, at present, the predicted peripheral physiological patterns associated with certain psychological constructs are often open to precise measurement whereas the predicted CNS patterns associated with their underlying control processes are not. For example, specific physiological patterns of challenge and threat states (see below) can be distinguished peripherally via cardiovascular patterns. However, given the current state of technology, the hypothalamic differentiation underlying these peripheral patterns cannot be distinguished via brain imaging. Thus, functional brain imaging data may currently be able to demonstrate activation of the hypothalamus but not to distinguish between the oppositional motivational states resulting from that activation. Peripheral indexes are currently necessary for the latter.

Here, we are concerned with indexes derived from physiological responses controlled by the CNS via the peripheral nervous system. The major branches of the peripheral nervous system include the autonomic and somatic ones. The former provides neurological pathways allowing the brain and spinal cord to control the visceral organs via its sympathetic and parasympathetic branches. The latter provides neurological pathways allowing the brain and spinal cord to control skeletal muscles.

Many peripheral physiological responses manifest themselves in ways that can be assessed relatively easily, including pulse rate, respiration rate, and overt facial muscle movements, which can be assessed with little if any specialized technology. Not surprisingly, then, many investigators, buoyed by the false assumption that any autonomic response is as good as any other (see above), chose physiological responses as candidate indexes of psychological constructs based on convenience and on ease of measurement rather than specificity and, hence, validity.

With more advanced technology, many more peripheral physiological responses can be measured. Polygraphs advanced peripheral electrophysiological measures of biopotential signals—for example, those produced by cardiac activity via the electrocardiogram (ECG), covert skeletal muscle activity via the electromyogram (EMG), and gastric activity via the electrogastrogram (EGG). The advent of various electromechanical and electronic transducers made measurement possible for peripheral responses lacking endogenous biopotential signals such as eccrine sweat gland activity via skin conductance responses, stroke volume, and cardiac output via cardiac impedance responses, and so forth. In sum, the plethora of peripherally controlled physiological measures that present themselves as potential indexes of social psychological constructs have presented many opportunities for social psychophysiologists.

Principle 3: Specification of Theoretical Physiological Linkages between Social Psychological Constructs and Candidate Physiological Measures

Precisely specifying both the social psychological construct and the candidate physiological responses is a necessary but not sufficient step toward establishing valid physiological indexes of those constructs. Investigators must also specify a valid theoretical rationale for linking the two. Unlike many self-report and behavioral measures, a rationale of simple face validity is not sufficient to link physiological measures with psychological constructs. The absence of a plausible theoretical rationale limits a physiological index to speculation based purely on empirical correlation rather than demonstrable or putative causal connections. The strength of a physiological index rests on the validity of the rationale. If the rationale is questionable or purely speculative, the index is cast in doubt.

Generally speaking, there is a positive relationship between the complexity of the set of candidate physiological responses and the complexity of the theoretical rationale required for linking physiological responses to psychological constructs. In addition, given the increase in costs associated with the measurement of larger sets of physiological responses, it is not surprising that social psychologists and others have historically shown a propensity to index psychological constructs with unitary peripheral physiological measures. However, this trade-off makes it extremely difficult to avoid the problem of many-to-many relationships between constructs and unitary physiological indexes (see propositions earlier) even if the construct and the candidate physiological response are highly specific and the appropriate domain or context has been limited. Hence, the strength of inference is likely to be compromised by an approach to an index based on a unitary physiological response (Blascovich, 2000; Blascovich & Kelsey, 1990; Cacioppo & Tassinary, 1990).

For example, although reviewers (e.g., Landis, 1930) have long understood the lack of a one-to-one relationship between electrodermal activity and psychological constructs, pursuit of electrodermal activity as a more or less informative index of psychological and social psychological constructs has persisted long since. One can theoretically link the magnitude of a single electrodermal measure (e.g., skin conductance level) to a psychological construct (e.g., fear) based on an up-to-date model of autonomic activation of electrodermal activity. For example, Dawson, Schell, and Filion (2000) posit that electrodermal activity is elicited by affective stimuli that trigger amygdalar activation of the sympathetic branch of the autonomic nervous system.

Hence, theoretically, an increase in skin conductance response level can index fear. However, without ruling out alternative causes of other influences on electrodermal activity also posited by Dawson and colleagues (2000) (e.g., activation of the reticular formation due to increased muscle tone, hypothalamic activation due to thermoregulation, premotor cortex activity due to fine muscle control, and prefrontal cortical activity due to orienting and attention), one cannot guarantee a one-to-one relationship between the construct (i.e., fear) and the response (i.e., skin conductance level). One conceivably could restrict the psychosocial context to lessen possible influences of increased muscle tone, fine motor control, thermoregulation, and orienting and attention, thereby strengthening the inference provided by the electrodermal measure. However, even after ruling out these alternative causes, one would also have to rule out the possibility of alternative emotional constructs because amygdalar activation can be triggered by either positive or negative affect or emotions. All one could really infer is that an affective response occurred perhaps with an intensity indexed by the electrodermal response, hardly a strong basis for concluding that a psychological fear response had occurred.

It is important to note here that the physiological theory (i.e., Dawson et al., 2000) is not the cause of the problem of inference in this example—quite the opposite. Nor is it a problem of measurement technology. Rather, the problem is a failure to fully account for the complexity of the theoretical linkage by limiting the set of candidate physiological responses to a unitary one while ignoring other measures (e.g., electromyographic recordings of peripheral skeletal–motor activity, electrocortical measures, and blood flow patterns in the brain) that would help form a pattern of responses that together with electrodermal response would provide a stronger basis for inferring the operation of the construct (i.e., fear) when the specified pattern held.

Additional Inferential Issues

A perfect physiological index would have a completely transitive or invariant one-to-one relationship with the psychological construct so that one would not occur without the other (Cacioppo & Tassinary, 1990). Unfortunately, such completely transitive relationships are relatively uncommon in psychology, especially when physiological responses are posited as indexes of psychological constructs, including social psychological ones. However, steps can be taken to increase the likelihood of a one-to-one or invariant relationship.

Even specific physiological systems (e.g., cardiovascular and electrodermal) generally respond to multiple inputs (e.g., psychological and metabolic demands) in overlapping ways. For example, both motivational task engagement and physical exercise can increase cardiac responses in similar ways (e.g., heart rate and ventricular contractility). Such many-to-many relationships make it difficult to disentangle specific psychosocial inputs from specific physiological effects and, hence, to establish specific construct-to-physiological response relationships that can be used as the basis for invariant physiological indexes. Hence, the task of establishing the desirable one-to-one relationships involves being able to discriminate constructs on the one hand and patterns of physiological responses on the other.

Generally, the more one limits the social psychological construct and expands the set of physiological measures indexing it, the closer the construct and index can come to a one-to-one or "invariant index" (Blascovich, 2000; Cacioppo & Tassinary, 1990). Cacioppo and colleagues (2000) described a taxonomy of possible relationships between psychological constructs and physiological responses that helps clarify the nature of these relationships. Their scheme (see Figure 2.2) distinguishes categories of construct-measure relationships as a function of discriminant value or specificity and generality.

Cacioppo and Tassinary (1990) argue that limiting the psychosocial context helps increase the likelihood that a physiological index will discriminate between target and alternative constructs. For example, an investigator can be more certain that increases in heart rate and ventricular contractility are driven by, and therefore index, non-aerobic task engagement rather than aerobic task engagement (i.e., metabolic demands) if the context in which these responses are assessed is bound in such a way that does not permit physical exertion. Cacioppo and Tassinary also suggest ways in which investigators can decrease the plethora of possible many-to-one relationships—thereby increasing the possibility of one-to-one relationships between constructs and physiological responses—via fractionation of multiple physiological responses over time (see discussion below). Hence, the greater the psychosocial specificity and the construct-response specificity, the more likely a physiological index is to approach a one-to-one relationship with a construct.

An implication of this scheme is that social psychologists desirous of physiologically indexing constructs should aim for a bound context and a set of multiple physiological measures—thereby a "marker." In essence, markers are context-bound invariants. To the extent that the context or domain is limited, resulting use of the physiological markers will increase internal validity of experiments and be theoretically useful. To the extent that the bound context corresponds to a large category of psychosocial situations, the physiological markers will have external validity and generalizability in addition to internal validity. Blascovich (2000) incorporated and expanded on Cacioppo and Tassinary's (1990) suggestions in the form of propositions for increasing the discriminant validity of markers. In revised form here, these propositions include (1) *divergent validation*, (2) *patterns of overlapping multiple responses*, and (3) *assessment over time*. None of these propositions is unique to psychophysiological measurement. Indeed, they are propositions borrowed largely from the social psychological methodology literature and applied to psychophysiological measurement.

Proposition 1: Divergent Validation

Divergent validation is well-known in social psychology generally via, especially, the work of Donald Campbell (1988a), and Proposition 1 is an application of his argument. Hence, our proposition of *divergent validation* states that valid markers are more likely to be useful inferentially if physiological responses are shown to diverge between target and oppositional constructs rather than between the presence and absence of the psychological input or processes circumscribed by a singular construct. For example, the inference drawn from a physiological response, such as the amplitude of a skin conductance response, that occurs during an instance of a stimulus that induces avoidance, such as a snake, is more likely and convincing if the response does not occur during an instance of a stimulus that induces approach, such as a teddy bear, than if the comparison is between a stimulus-inducing avoidance and a neutral one (e.g., a book). It follows, then, that validation of physiological markers will be more likely for use in indexing constructs of a more rather than less superordinate nature. For example, attempting to develop discriminant physiological indexes for positive and negative affect should be easier than attempting to develop distinctive physiological indexes for happiness and joy, or fear and anger.

Proposition 2: Patterns of Overlapping Multiple Physiological Responses

The proposition of *patterns of overlapping multiple physiological responses* states that valid physiological markers are more likely to be achieved if the set of physiological responses targeted is more rather than less inclusive, and if the same set is used to index contrasting oppositional constructs; that is, if the set of responses differs in form but not substance. As illustrated previously, increased heart rate and ventricular contractility do not necessarily distinguish nonaerobic task engagement from aerobic task engagement. Both should cause increases in heart rate and ventricular contractility. However, oxygen demand and oxygen consumption should increase during exercise but not during a nonmetabolically demanding performance. Hence, expanding the set of responses in this example to include oxygen consumption provides a way of distinguishing nonaerobic from aerobic task en-
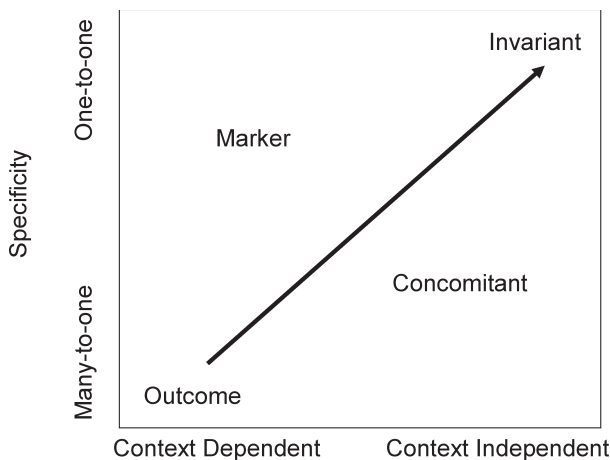

FIGURE 2.2. Categories of construct–measure relationships (Cacioppo et al., 2000).

gagement. Indeed, this has been shown to be the case (Rousselle, Blascovich, & Kelsey, 1995).

Proposition 3: Assessment over Time

The proposition of *assessment over time* states that valid physiological markers are more likely to be achieved if the set of physiological responses targeted is assessed continuously over time. For example, if one is using facial electromyography to assess positive affect, the intensity of zygomaticus major (i.e., the "smile" muscles in the cheeks) activity may not differentiate a genuine or automatic smile from an intentional forced one, whereas the time course of the activity may.

In sum, many-to-many relationships between social psychological constructs and indexes based on physiological responses appear to be the rule rather than the exception. However, investigators can increase the likelihood of achieving one-to-one relationships between their constructs and physiological indexes by carefully limiting the context within which they apply, by using a divergent validational approach, and by constructing indexes based on overlapping multiple responses over time. However, it must be stated that none of these strategies is logically necessary so that single physiological response indexes that provide a strong basis for inference, though less likely, are not logically impossible. Conversely, use of all these strategies does not guarantee valid indexes. Rather, the "devil is in the details" of the specification of psychological constructs, candidate physiological responses, and valid rationales linking the two.

Summary

Social psychologists can find much of value in psychophysiological and physiological theory on which to select candidate physiological response sets as indexes of theoretical constructs (see examples below). The pursuit of indexes based on unitary physiological measures, although appealing from an economic point of view, has not proven very successful in our view. However, neither past history nor logic absolutely rules out the establishment of valid unitary physiological indexes of social psychological constructs.

PERIPHERAL PHYSIOLOGICAL INDEXES OF SOCIAL PSYCHOLOGICAL CONSTRUCTS

General Issues

Physiological indexes of the operation of both intra-individual and interindividual social psychological processes can only be measured at the individual level. And, like other subjective (i.e., self-report) and behavioral measurement, physiological measurement is subject to individual differences, measurement error, and artifact. However, the measurement of physiological responses has some advantages over self-report and behavioral responses in this regard.

One advantage is due to the fidelity of physiological recording technology. Modern digital physiological measurement technology permits nearly noise-free, highly precise recording of physiological responses. An additional advantage stems from the relatively automatic and covert nature of physiological responses. Compared to self-report and behavioral measures, physiological responses enjoy relative immunity to demand characteristics and participant impression management concerns.

This immunity is illustrated in a review of self-report, behavioral, and physiological measures taken during perceiver interactions with physically and racially stigmatized others across several experiments by Blascovich, Mendes, and Seery (2002). They concluded that the theoretically predictable threat experienced by nonstigmatized participants interacting with stigmatized others (Crocker, Major, & Steele, 1998) was evidenced by physiological response patterns indicative of threat. Not only did corresponding self-report responses fail to indicate threat, but compared to participants interacting with nonstigmatized others, participants interacting with stigmatized individuals reported them as more intelligent, attractive, and so on, suggesting sensitivity on participants' part to impression formation concerns. Dambrun, Depres, and Guimond (2003) found similar fractionation between physiological and self-report responses to participants' contact with outgroup members.

In addition, advantages accrue to physiological measures by virtue of the relative ease with which they can be recorded continuously and online. Continuous recording not only permits increased reliability generally but also makes one-to-one relationships between physiological responses and psychological constructs more likely because continuous measurement is necessary for assessing physiological response patterns over time in line with the third proposition discussed previously. Online recording also avoids problems associated with prospective and retrospective self-reports because physiological responses are recorded conterminously with the operation of the psychological processes that they index without interrupting the flow of an experiment.

Peripheral Physiological Indexes of Basic Psychological Processes Underlying Social Psychological Processes

Social psychological phenomena and processes rest on and involve more rudimentary psychological processes, especially low-level perceptual and cognitive ones. Thus, psychophysiological indexes of these more basic processes (attention, perception, information processing, memory, etc.) can be valuable to social psychologists, as illustrated by the emergence and importance of dual-process theoretical models such as Wegner's (1994) model of ironic processing, Devine's (1989) automatic and control processes of prejudice, and Wilson, Lindzey, and Schooler's (2000) model of dual attitudes. When social psychologists are only interested in whether or not some stimuli or set of stimuli is perceptually or

cognitively processed at all rather than exactly how it is processed, even a unitary response can provide a reasonable basis for inference under proper conditions.

Perceptual Recognition

An interesting illustration involves perceptual (i.e., visual) aspects of person memory—specifically, whether or not an individual can respond to a previously known face even without any memory of it. Studies by Bauer (1984) and later Tranel and Damasio (1985) demonstrated that individuals who have lost the ability to recognize faces consciously (i.e., prosopagnosics) exhibited more frequent and larger skin conductance responses to pictures of familiar faces (e.g., family members) than to unfamiliar faces, even though they showed no evidence of conscious recognition. In this case, investigators can infer a basic or low-level recognition effect even though there is not a one-to-one correspondence between electrodermal activity and facial recognition.

Cognitive Processing

Another illustration involves pupillary responses, primarily pupillary dilations, as an index of cognitive processing load. Beatty and Lucero-Wagoner (2000) argue that autonomically controlled task-evoked pupillary responses (TEPRs) index brain processes associated with the intensity of human cognitive processing. They based this claim on research conducted as long ago as the 19th century (e.g., Heinrich, 1896; Schiff, 1875) as well as more recent research (e.g., Ahern & Beatty, 1981; Hess & Polt, 1964) showing that task difficulty is correlated with the intensity of TEPRs. Again, just as with the use of electrodermal activity as a unitary index of recognition, investigators can infer heightened cognitive processing from coterminous TEPRs even though there is not a one-to-one correspondence between pupillary responses and cognitive processing.

Summary

The types of inferences that can be drawn from physiological indexes based on unitary peripheral physiological correlates of basic perceptual or cognitive processes can be important to social psychologists. For example, in terms of dual-process theories, the strength of an electrodermal response as an index of perceptual recognition can provide the basis for the assumption that a low-level or automatic response exists to be controlled. Moreover, the strength of task-evoked pupillary responses can index cognitive load, which in turn would provide an inferential basis for whether or not cognitive control mechanisms were activated. Neither index, of course, would be very specific, about either the exact nature of the perceptual recognition in the case of the skin conductance index or the exact nature of the type of cognitive processing involved in control. To the best of our knowledge, a combination of such indexes has not been yet employed by social psychologists.

Peripheral Physiological Indexes of Social Psychological Constructs and Processes

The value of peripheral physiological indexes for social psychologists is not limited to indexes of basic perceptual and conceptual processes based on a simple notion of the identity thesis as described earlier. Rather, social psychologists have relied on more sophisticated physiological theoretical notions to validate and employ peripheral physiological markers of the operation of social psychological processes themselves. Several physiological models provide possible theoretical linkages for establishing peripheral physiological indexes of these higher-level social psychological constructs. These models include but are not limited to Cacioppo and Petty's (1981a, 1981b) linkages between affect and facial muscle movements, Obrist's (1981) general model of cardiovascular psychophysiology, Dienstbier's (1989) model of physiological toughness, Berntson, Cacioppo, and Quigley's (1991) doctrine of autonomic space, Porges's (1995) polyvagal theory, and Wright and Kirby's (2001) effort determination theory.

Page limitations for this chapter preclude an exhaustive general review of the use of peripheral physiological indexes in social psychology. Excellent general coverage can be found elsewhere (e.g., Gardner, Gabriel, & Diekman, 2000). Here, the focus is on two prototype examples or case studies illustrating, in depth, the use of peripheral physiological indexes according to the principles and propositions described earlier. In accord with the theme of this chapter, the first involves peripheral autonomic indexes of motivational processes and the other peripheral somatic indexes of affective processes.

PERIPHERAL AUTONOMIC INDEXES: A CASE STUDY OF CHALLENGE AND THREAT

Researchers have investigated motivation from a social psychological perspective for many years (see early review by Berkowitz, 1969). More than a decade ago, Blascovich and colleagues began to develop a biopsychosocial model of motivational processes integrating cognitive, social psychological, and physiological processes (Blascovich & Mendes, 2000; Blascovich & Tomaka, 1996). Their work fits the prototype criteria for physiologically indexing psychological constructs as discussed previously, including explicit definition of key constructs and specification of physiological response measures linked to the constructs via established physiological theory. Their work also illustrates the value of limiting the domain or context in which physiological indexes are applied and adherence to the propositions (i.e., divergent validation, patterns of overlapping multiple physiological responses, and assessment over time) discussed previously for increasing the likelihood of a one-to-one relationship between physiological indexes and psychological constructs.

Constructs, Candidate Physiological Indexes, and Theoretical Rationale

Principle 1: Construct Specification

The biopsychosocial model focuses on performance situations that are goal relevant to the individual. Challenge and threat processes involve deliberate and/or automatic influences on evaluations of resources (e.g., skills, knowledge, and abilities; various dispositions; and social support) relative to evaluations of demands (e.g., danger, uncertainty, required effort, various dispositions, and social evaluation). Challenge results from an overall assessment that one's resources meet or exceed the demands of the situation, and threat results from an overall assessment that the demands of the situation exceed one's resources. According to the biopsychosocial model, though seemingly categorically defined, challenge and threat actually represent anchors on a bipolar motivational continuum (Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003).

As Cacioppo and Tassinary (1990) and Blascovich (2000) have argued, finding one-to-one (i.e., invariant, or general) physiological indexes of social psychological constructs that apply across all possible contexts is very difficult and, hence, unlikely. They suggest that this difficulty can be decreased and the likelihood of finding one-to-one correspondence increased by limiting the context in which the physiological index applies to the social psychological constructs. Consequently, any proposed physiological indexes are likely to enjoy a one-to-one relationship only in the empirically limited context even though the theory may be more generally applicable. Such a context limitation is not a limitation of the theory but rather an empirical limitation.

Blascovich and colleagues have limited the empirical context in which they typically have investigated challenge and threat in two ways. First, they restrict the context to non-metabolically demanding goal-relevant situations. This eliminates physically demanding situations, which would interfere with a one-to-one relationship between the candidate cardiovascular indexes (see later) and challenge and threat processes. The accompanying metabolic demands driven by large muscle activity would elicit cardiovascular responses that could obfuscate the relationship of the cardiovascular physiological indexes to the psychological demands that contribute to challenge and threat motivation.

Second, they limit the non-metabolically demanding goal-relevant situation to situations requiring active instrumental cognitive responses on the part of the individual (e.g., test taking, speech giving, game playing, negotiations, and singing). They term such situations "motivated performance situations" and contrast them to passive performance situations in which instrumental cognitive responses are not necessarily required (e.g., emotionally coping with the death of a loved one or watching a scary movie). Based on the work of Obrist (1981) as well as their own (e.g., Tomaka, 1993), Blascovich and colleagues concluded that their candidate cardiovascular indexes of challenge and threat do not hold in passive situations. As with the first, this limitation is an empirical one. It does not limit resource/demand assessments and, hence, challenge and threat to motivated performance situations. It only limits the validity of the physiological indexes of challenge and threat to motivated performance situations.

Principle 2: Specification of Candidate Physiological Responses

Blascovich and his colleagues focused on cardiovascular responses as candidate measures or markers of challenge and threat. Their initial consideration of cardiovascular measures was based on intuitions drawn from previous literature linking increases in various cardiovascular responses (e.g., heart rate reactivity and blood pressure reactivity) during stressful or threatening situations. Such cardiovascular reactivity research and theory came to the fore in personality and social psychology during the 1970s and 1980s as many investigators sought to find a link between the coronary-prone personality type and heart disease (see Blascovich, in press, for a more thorough discussion).

Principle 3: Specification of a Plausible Theoretical Physiological Rationale for Linking Candidate Physiological Measures to Social Psychological Constructs

Paul Obrist (1981) categorized cardiovascular reactivity patterns as relatively benign or malignant according to their antecedents. He argued that when cardiac responses are driven by somatic demands during aerobic exercise or when cardiac responses occur in passive coping situations, a benign pattern that he labeled "cardiac–somatic coupling" emerges. He argued that cardiac responses are not driven by somatic demands in active coping situations and that a potentially pathophysiological pattern that he labeled "cardiac–somatic uncoupling" emerges. Subsequently and not surprisingly, many cardiovascular reactivity researchers began limiting the experimental contexts in which they used cardiovascular reactivity indexes to investigate links between personality constructs and cardiovascular disease to active coping situations.

Less than a decade later, Richard Dienstbier (1989) published his *Psychological Review* article on physiological toughness and weakness. Relying nearly exclusively on investigations of cardiovascular responses of rodents, his work suggested that further clarification was necessary to Obrist's distinctions. Specifically, Dienstbier's work provided the basis for distinguishing benign from malignant patterns of cardiovascular reactivity within active coping situations. A year later, Kasprowicz, Manuck, Malkoff, and Krantz (1990) found evidence that distinguished "cardiac" from "vascular" responders during active coping situations. A few years later Manuck, Kamarck, Kasprowicz, and Waldstein (1993) suggested that vascular reactivity was the more malignant (with regard to cardiovascular disease) of the two.

Dienstbier (1989) concluded that physiologically tough animals, ones that survive and even thrive in potentially dangerous or threatening situations, exhibit a pattern of cardiovascular responses in such situations that is driven primarily by activation of the sympathetic–adrenomedullary (SAM) axis. According to Dienstbier, SAM axis activation mobilizes energy reserves in a benign way such that sympathetic neural stimulation of the myocardium increases heart rate and ventricular contractility while adrenal medullary stimulation quickly increases circulating catecholamines, including epinephrine and norepinephrine. The circulating catecholamines induce systemic vasodilation, thereby decreasing total peripheral resistance (TPR). This combination of cardiac and vascular stimulation increases blood flow (i.e., cardiac output; CO). Typically, such SAM activation has a quick onset and offset. Physiologically weakened animals, on the other hand, exhibit marked hypothalamic–pituitary–adrenal cortical (HPA) axis[3] activity in addition to SAM axis activation. The HPA axis activity increases the production of cortisol and inhibits the adrenal medullary release of catecholamines, thereby reducing the vasodilatory effects of SAM axis activation and, in turn, decreasing CO. Typically, HPA activation is slower to habituate. Figure 2.3 depicts both patterns.

Validation of Cardiovascular Physiological Indexes of Challenge and Threat

Dienstbier's model of physiological toughness provided the theoretical basis for the cardiovascular response patterns that Blascovich and colleagues would later establish as markers of challenge and threat processes in nonmetabolically demanding motivated performance situations. Within such a limited context, the specific responses (i.e., heart rate, ventricular contractility, CO, and TPR) could be used in accordance with the *propositions of divergent validation, patterns of overlapping multiple physiological responses, and assessment over time*, thereby increasing the likelihood of achieving one-to-one relationships between the indexes and the primary constructs underlying the biopsychosocial model of challenge and threat.

Correlational Validation

Tomaka, Blascovich, Kelsey, and Leitten (1993) conducted a series of correlational experiments to ascertain the relationship between self-reported evaluations of perceived demands and resources prior to active coping performance tasks and patterns of cardiovascular responses during subsequent performance. They hypothesized that an evaluation of resources exceeding demands (i.e., challenge) would be accompanied by the pattern of physiological toughness specified by Dienstbier (i.e., increases in heart rate, ventricular contractility, and CO, and a decrease in TPR; see Figure 2.3), and that an evaluation of demands exceeding resources (i.e., threat) would be accompanied by a pattern of physiological weakness specified by Dienstbier (increases in heart rate and vasocontractility, but little change in CO and little change or even increases in TPR; see Figure 2.3). These hypotheses were confirmed.

A within-subjects study encompassing four motivated performance tasks in succession was reported by Quigley, Barrett, and Weinstein (2002). Using hierarchical linear modeling (HLM) analytic techniques, these investigators demonstrated that within-subjects changes in pretask appraisals of demands and resources across tasks were statistically reliably related to changes in cardiovascular challenge and threat patterns during task performance. In other words, as pretask appraisals changed in the direction of challenge from threat, the patterns of cardiovascular responses became more challenge-like.

Experimental Validation

To strengthen the inferential basis for a causal connection between challenge and threat processes and the candidate cardiovascular indexes, experimental studies were
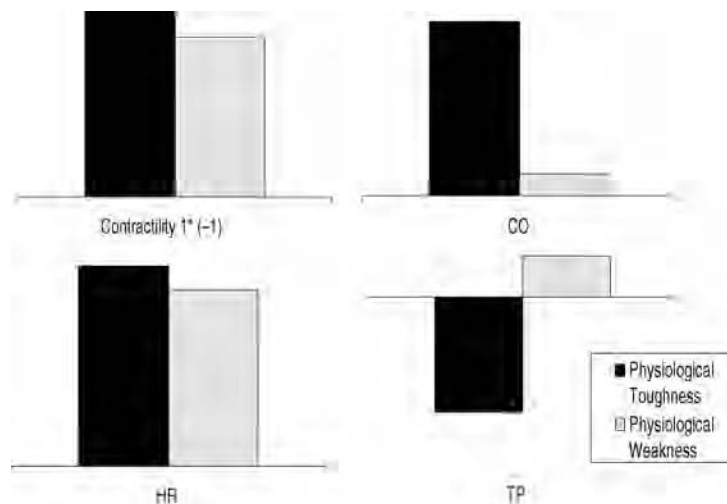

FIGURE 2.3. Dienstbier's (1989) cardiovascular patterns of physiological toughness/weakness.

conducted. For example, Tomaka, Blascovich, Kibler and Ernst (1997, experiment 1) manipulated instructional set in a manner likely to influence challenge and threat evaluations and, hence, patterns of cardiovascular responses via Dienstbier's model. Specifically, manipulations of the content of the task instructions; that is, practice (e.g., "This is just a practice run.") versus criterion (e.g., "This is an important run for you to do well.") coupled with vocal tone (i.e., friendly vs. hostile) during task instructions resulted in the hypothesized cardiovascular patterns of challenge and threat reactivity. Hunter (2001) also manipulated instructions such that participants were told that because they would be reading or singing the U.S. national anthem later during the experiment they needed to practice reading or singing the first couple stanzas. The predicted challenge and threat patterns of cardiovascular responses occurred such that significantly more of the individuals who sang evidenced the threat pattern than individuals who anticipated reading, and significantly more of the individuals who read evidenced the challenge pattern than individuals who anticipated singing.

Manipulated Physiological Studies

To determine whether interoceptions associated with the challenge and threat patterns of cardiovascular responses influenced the evaluations of demands and resources, Tomaka and colleagues (1997, experiments 2 and 3) manipulated participants' cardiovascular patterns mimicking the challenge and threat patterns physically (while engaged in moderate aerobic exercise, mimicking challenge in one experiment; while experiencing a warm pressor mimicking challenge in the other). During these manipulations, participants' self-reported evaluations of demands and resources for an upcoming serial subtraction task did not differ as a function of condition. These results suggested that overall challenge/threat evaluations were not likely caused by the physiological patterns themselves, but, rather, that the physiological patterns resulted from the evaluation or assessment mechanisms as postulated by the biopsychosocial model.

Predictive Validation

Recently, Blascovich, Seery, Mugridge, Weisbuch, and Norris (2004) conducted a study to determine the predictive validity of the cardiovascular indexes of challenge and threat patterns. Varsity baseball and softball players (position players other than pitchers; i.e., hitters) served as participants. Participants each gave two 3-minute speeches, a performance-irrelevant (control) speech (why I am a good friend) and a performance-relevant (predictor) speech (how I would approach a critical hitting situation). A unitary cardiovascular index derived from the set of multiple cardiovascular measures used to index challenge and threat was derived for each speech task. Controlling for the indexed cardiovascular responses during the performance-irrelevant speech, the cardiovascular challenge/threat index that occurred during the performance-relevant speech reliably predicted

major outcome measures of offensive baseball performance (the runs created index, batting averages, etc.).

Summary of Validational Studies

In aggregate, the aforementioned studies provide strong validation that challenge and threat are marked by specifiable and divergent patterns of cardiovascular responses across the same set of specific measures. They also attest to the value of grounding such candidate indexes in physiological theory (e.g., Dienstbier, 1989). Based on the validational studies alone, these markers appear to be robust.

The validational studies have also produced results that are remarkably consistent in at least two respects. First, no gender differences have ever been found that suggest that men and women evidence different cardiovascular patterns under either challenge or threat. Second, with one exception, better performance results were found under challenge rather than threat conditions. The one exception occurred in the study by Hunter (2001). She found that participant performance on a vigilance task was better under threat than challenge motivation. Indeed, one might expect better vigilance during threat as functional.

Application of Cardiovascular Indexes of Challenge and Threat to Theoretical Issues in Social Psychology

The validation of the cardiovascular indexes of challenge and threat was not the final goal of the proponents of the biopsychosocial model of challenge and threat. Rather, the goal was to be able to use such indexes to test social psychological theory. A corpus of work using these indexes to test various social psychological theories has paralleled the validational work. The basic rationale for this theory-testing work is based on a single argument. To the extent that challenge and threat motivational states can be predicted by a personality or social psychological theory, and the predictions can be tested in a nonmetabolically demanding motivated performance context, the cardiovascular indexes of challenge and threat can be used to test the theory. The conditions specified by this argument have proven remarkably attainable permitting tests of hypotheses drawn from a plethora of personality and social psychological theories.

Attitude Functionality

Allport (1935), Fazio (1989), Katz (1960), and others maintain that attitudes are functional in the sense that they facilitate decision making. Blascovich and colleagues (1993) predicted that during a motivated performance situation requiring rapid preference decisions, individuals with preexisting task-relevant attitudes are more likely to be challenged than threatened. For example, in one of the reported studies (Blascovich et al., 1993, Study 2), the investigators manipulated preexisting attitudes by having participants repeatedly rehearse attitudes toward abstract paintings in one of two 15-item subsets of abstract paintings in the first phase of an ex-

periment. In the second phase, they required participants to rapidly express preferences for one of two paintings in each of 34 slides of randomly paired paintings entirely chosen from either the rehearsed or the nonrehearsed subsets of paintings. During the second phase, the cardiovascular response measures underlying the challenge and threat patterns were continuously recorded. The results indicated that participants in the rehearsed paintings condition exhibited the challenge pattern of cardiovascular response, and participants in the nonrehearsed paintings condition exhibited the threat pattern.

Dispositions

The cardiovascular indexes of challenge and threat have enabled researchers to investigate dispositional influences on task performance in motivated performance situations. For example, Lerner (1980) and others (e.g., Lazarus & Folkman, 1984) hypothesized that dispositional belief in a just world allows individuals to adapt better to the demands of everyday life. Tomaka and Blascovich (1994) tested this notion in a correlational study. As expected, participants high in belief in a just world exhibited the challenge pattern of cardiovascular responses, whereas participants low in belief in a just world exhibited the threat pattern during the performance a mental arithmetic task in a motivated performance situation.

In a somewhat more complex set of studies, Seery, Blascovich, Weisbuch, and Vick (2004) utilized the cardiovascular indexes of challenge and threat to test predictions based on Kernis's theory (Kernis, 1993; Kernis & Waschull, 1995) that individuals with unstable high self-esteem are plagued by self-doubt and, hence, are likely to have low resource evaluation. The results indicated that individuals with unstable high self-esteem were threatened by failure feedback in a motivated performance situation and that individuals with stable high self-esteem remained challenged.

Social Facilitation

Social facilitation effects have often been described as the oldest experimental social psychological finding and, not surprisingly, have attracted the attention of theorists for more than a century. The basic findings are that the presence of others, whether coactors or members of an audience, facilitates individual performance on a well-learned task and inhibits performance on an unlearned or novel task. Based on the biopsychosocial model of challenge and threat, Blascovich, Mendes, Hunter, and Salomon (1999) hypothesized that the presence of others increases the goal relevance of task performance regardless of previous task experience. Furthermore, they hypothesized that, based on a relative lack of resources for performance, individuals attempting performance on a novel or unlearned task should be threatened during task performance, and individuals attempting performance on a well-learned task should be challenged. These hypotheses were confirmed by cardiovascular indexes of challenge and threat.

Social Comparison

Mendes, Blascovich, Major, and Seery (2001) examined challenge and threat processes during upward and downward social comparisons. They hypothesized that participants interacting with upward-comparison others should experience threat, whereas participants interacting with downward-comparison others should experience challenge. Their experiments provided a conservative test of these hypotheses because the investigators chose a motivated performance situation that involved a cooperative rather than a competitive task. In their first experiment, as predicted, participants cooperating with upward-comparison partners exhibited the cardiovascular pattern indicative of threat, and participants cooperating with downward-comparison partners exhibited the cardiovascular pattern indicative of challenge. In their second experiment, as predicted, attitudinal dissimilarity with their partners exacerbated this pattern of responses.

Stigma

Many theorists support the hypothesis that nonstigmatized individuals experience threat, based on a sense of uncertainty, discomfort, anxiety, and even danger, when interacting with members of stigmatized groups (Crocker et al., 1998; Goffman, 1963; Jones et al., 1984). However, until recently relatively little research has directly demonstrated this widespread assumption. Blascovich, Mendes, Hunter, Lickel, and Kowai-Bell (2001) and Mendes, Blascovich, Lickel, and Hunter (2002) conducted several experiments in which the cardiovascular indexes of challenge and threat confirmed this hypothesis. The effects held whether the interactant was stigmatized by virtue of disfigurement (i.e., facial birthmark), race, or socioeconomic status.

PERIPHERAL SOMATIC INDEXES: A CASE STUDY OF POSITIVE AND NEGATIVE AFFECT

Researchers have investigated affect and emotion from a social psychological perspective since William James's classic discussion of emotions (Bradley, 2000). Much work has focused on both the universality of basic emotional experiences (e.g., Ekman & Friesen, 1986) and possible autonomic footprints (i.e., physiological indexes) of the so-called basic-level emotions. Regarding the latter, there seems to be little agreement that such indexes exist (Cacioppo et al., 1993; Stemmler, 1989). However, there has been progress in the development of physiological indexes of the relevant superordinate states, positive and negative affect.

Constructs, Candidate Physiological Indexes, and Theoretical Rationale

Principle 1: Construct Specification

As discussed earlier, one cannot establish a meaningful one-to-one relationship between a social psychological construct and a physiological response without explicitly

defining the construct that is to be indexed. In the case of positive and negative affect, several definitions have been proffered. Most appear consistent with Shaver, Schwartz, Kirson, and O'Conner's (1987) definition that positive and negative affect are superordinate categories of feeling states with positive and negative valence, respectively. Positive affect encompasses specific states such as happiness and excitement, whereas negative affect encompasses states such as sadness and fear; the specific states under each category possess unique aspects, but they also share a common positive or negative element.

Principle 2: Specification of Candidate Physiological Indexes

Cacioppo and colleagues (e.g., Cacioppo & Petty, 1981a, 1981b; Cacioppo, Petty, Losch, & Kim, 1986) promoted the use of electromyography (EMG) as a way to quantify facial expressions, which they reasoned should vary with intrapersonal affective experience. Although not the only approach to identifying candidate physiological or even somatic indexes of affect, facial muscle movements underlying facial expressions have proven one of the most fruitful. In particular, Cacioppo and colleagues focused on muscles primarily involved with smiling (zygomaticus majori) and frowning (corrugator supercilii), reasoning that smiling and frowning obviously are associated with positive and negative affect, respectively.

Principle 3: Specification of a Plausible Theoretical Physiological Rationale for Linking Candidate Physiological Measures to Social Psychological Constructs

The theoretical roots supporting the use of somatic markers of positive and negative affect rest on Darwin's (1872/1965) arguments regarding the evolutionary significance of facial expressions. In the last several decades, a substantial body of literature has emerged that has illustrated the critical role of facial expression in social interaction and the expression of affect and emotion (e.g., Ekman, 1993).

Validation of Somatic Physiological Indexes of Positive and Negative Affect

A potentially straightforward way to quantify facial expressions is to ask observers to rate them; indeed, much research has established the value of this approach, ranging from the collection of overall impressions of untrained observers to the use of video recording techniques to evaluate expressions that appear for only a fraction of a second (e.g., Ekman & Friesen, 1978). However, it is possible (and arguably common) for the muscle movements that create facial expressions to be so covert that no visible expression emerges to be coded (e.g., Cacioppo & Petty, 1979). Observational methods have little value in these situations, but somatic psychophysiological measures (i.e., EMG) can detect even minute muscle activity. This holds true even when individuals are motivated to consciously control their facial expressions (e.g., intergroup interactions); in spite of such efforts,
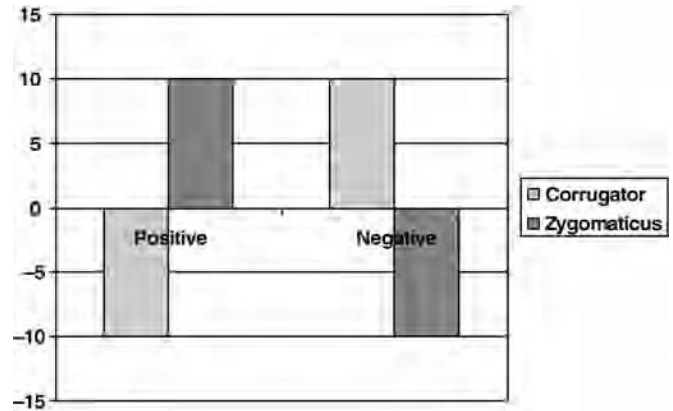
FIGURE 2.4. Facial electromyographic markers of positive and negative affect.

subtle leakage occurs that is accessible to EMG measurement.

For facial EMG, electrodes sensitive to muscle action potentials are placed over specific facial muscles (for procedural details and a more comprehensive review of EMG literature, see Cacioppo, Tassinary, & Fridlund, 1990; Tassinary & Cacioppo, 2000). Changes in muscle activity that occur with the onset of a relevant event (e.g., stimulus presentation, imagination, and social interaction)—regardless of whether or not visible facial expressions appear—are measured and used to index affective experience. Several elements of muscle activity can be recorded, including duration and intensity. Cacioppo and colleagues (1986) validated the use of the two candidate muscles to assess both valence and intensity of affect: zygomaticus majori ("smile muscles" in the cheeks) and corrugator supercilii ("frown muscles" in the brow). As depicted in Figure 2.4, during positive affect, zygomaticus activity increases and corrugator activity decreases, whereas during negative affect, zygomaticus activity decreases and corrugator activity increases. The theoretically based patterns of these somatic responses fit the propositions of *divergent validation, and patterns of overlapping multiple physiological responses*, thereby increasing the likelihood of achieving one-to-one relationships between the indexes and the primary constructs underlying the biopsychosocial model of challenge and threat.

In their study, Cacioppo and colleagues (1986) presented positively and negatively valenced images to participants and recorded participants' self-reported liking for each image along with EMG activity. Results confirmed the expected fractionation of zygomaticus and corrugator activity, which was also consistent with self-reported liking. In contrast, activity in unrelated muscles did not differ by image valence.

Application of Somatic Indexes of Positive and Negative Affect to Theoretical Issues in Social Psychology

Facial EMG has emerged as a valuable methodological tool in a variety of contexts. Recent examples in which

EMG was used to test social psychological theory regarding the role of affect in two research domains are described below.

Ease of Stimulus Processing

The mere exposure effect (Bornstein, 1989; Zajonc, 1968) is a robust finding in which increasing familiarity with a neutral stimulus (e.g., yearbook photos, Chinese ideographs, and polygons) breeds a more positive attitude toward it; in other words, merely being exposed repeatedly to a stimulus increases liking for it. Explanations for the mechanism that underlies the mere exposure effect fall into two general categories: cognitive explanations and affective explanations.

Cognitive explanations include the nonspecific activation model (Mandler, Nakamura, & Van Zandt, 1987) and the perceptual fluency/attributional model (Bornstein & D'Agostino, 1994). The nonspecific activation model holds that exposure to a stimulus creates a representation of it in memory, which is further activated by repeated exposure. When an individual is asked to evaluate the stimulus on a given dimension, the heightened activation associated with the stimulus's representation is applied to that dimension as long as it is relevant to the stimulus. This causes the individual to be more likely to judge that the stimulus possesses the given dimension. For example, Mandler and colleagues (1987) found that participants evaluated familiar black polygons as brighter, darker, or more likable than unfamiliar black polygons, depending only on which of the questions happened to be asked. According to the perceptual fluency/attributional model, a familiar stimulus is easier to process, which creates an affectively neutral experience of perceptual fluency. This fluency is then attributed to the most likely cause suggested by contextual cues, such as those provided by the items on a self-report questionnaire. Thus, in both cognitive models, the mere exposure effect does not depend on or include an affective component. Instead, greater liking after repeated exposure is merely an artifact of the questions posed to participants.

Affective models, in contrast, maintain that affect plays a critical role in the link between familiarity and liking. Biological and evolutionary theories suggest that an unfamiliar stimulus evokes fear innately, which is then habituated with repeated exposure (Zajonc, 1968). Hence, lack of familiarity is associated with danger and negative affect because it would be adaptive to avoid stimuli that have not demonstrated their benign nature over repeated interactions (Bornstein, 1989). Alternatively, the hedonic fluency model (Winkielman & Cacioppo, 2001) holds that the experience of fluency from cognitive models is necessarily positively valenced because it indicates that the individual possesses the skills to successfully reach a goal in the environment.

Harmon-Jones and Allen (2001) used facial EMG to pit cognitive and affective explanations against each other. After recording baseline zygomaticus and corrugator activity levels, 10 photos of women were repeatedly displayed to female participants. Participants were then shown these 10 familiar photos along with 10 unfamiliar photos of different women in a random order (the two sets of photos were counterbalanced, such that for half of the participants, the first set was made familiar, and for the other half, the second set was made familiar). Facial EMG responses were recorded during presentation and participants rated their degree of liking for each photo after its presentation.

Results revealed the typical mere exposure effect: Participants expressed greater self-reported liking for the familiar faces than for the unfamiliar ones. More important, participants exhibited greater zygomaticus activity while viewing familiar faces than while viewing unfamiliar ones, consistent with greater positive affect. Comparing these responses to baseline levels demonstrated that zygomaticus activity increased during exposure to familiar faces relative to baseline, whereas it did not differ from baseline during exposure to unfamiliar faces. There were no effects of familiarity on corrugator activity. These findings indicate that familiar stimuli engender a positive affective response that (1) is not an artifact of agreeing with self-report questions regardless of the stimulus dimension they assess and (2) does not reflect a relative difference driven by unfamiliar stimuli engendering negative affective responses. This supports affective explanations of the mere exposure effect over purely cognitive ones.

The mere exposure effect can be viewed as one example of a more general effect of stimulus processing ease: The easier a stimulus is to process, the more it will be liked, regardless of the specific presentation characteristics that facilitate processing. Winkielman and Cacioppo (2001) conducted an investigation similar to Harmon-Jones and Allen's (2001), manipulating ease of stimulus processing rather than stimulus familiarity.

In Study 1, Winkielman and Cacioppo (2001) presented participants with 20 black-and-white line drawings of common, neutrally valenced objects. Just prior to presentation of each stimulus, participants were subliminally exposed to a visual contour, created by removing all but the outer edge of a line drawing and then further degrading the image. Previous research has demonstrated that exposure to a contour that matches a subsequent target image facilitates processing of the target (e.g., Reber, Winkielman, & Schwarz, 1998); thus, for each participant, half of the targets were preceded by a matching contour and half were preceded by a mismatching contour (matching and mismatching stimuli were counterbalanced across participants). After each target presentation, half of the participants rated how positive their reaction was to the image and half rated how negative their reaction was. Facial EMG responses were recorded in the prepresentation period for each trial (baseline) and during presentation.

Results demonstrated that for participants who made the positively framed ratings, both zygomaticus activity and self-reported ratings were higher after matching contours than mismatching contours. This difference in zygomaticus activity was only significant immediately after stimulus presentation, not several seconds later. No effects emerged for corrugator activity or in the negatively framed condition.

In Study 2, Winkielman and Cacioppo (2001) conducted a conceptual replication, manipulating stimulus exposure duration instead of contour prime, with the logic that longer exposure time facilitates processing (Mackworth, 1963). The same stimuli from Study 1 were used in Study 2. Exposure duration condition (all less than 1 second) was manipulated within subjects. Instead of rating either how positive or how negative each stimulus was, all participants made a single bipolar judgment. Consistent with Study 1, results revealed that longer exposure duration resulted in both greater zygomaticus activity immediately after stimulus presentation (but not later) and more positive self-report ratings. Corrugator activity did not significantly differ by condition.

In sum, Winkielman and Cacioppo's (2001) studies demonstrate that manipulations that increase processing ease result in sudden and short-lived positive affect that is not limited to possibly artifactually self-reported responses. Together with Harmon-Jones and Allen's (2001) findings, the overall results suggest that purely cognitive explanations of the effects of processing ease are inadequate. Instead, positive affect in particular seems to play an important role. It is critical to note that this insight was made possible only with the theory-based application of facial EMG techniques, highlighting the utility of social psychophysiological methodology in general.

Prejudice and Discrimination

Facial EMG has also proven its value in the domain of prejudice and discrimination. Arguably, prejudice involves negative affect in the form of attitudinal bias toward a group of people, whereas discrimination refers to bias in behavior. A long-standing problem for researchers in this area is that people are reluctant to admit their own prejudices in the face of a social climate that values tolerance (McConahay, Hardee, & Batts, 1981). As discussed previously, physiological measures including facial EMG can avoid such self-presentation problems.

The work of Vanman and colleagues provides two illustrations of effective use of EMG in this context. In two scenario studies with white participants, Vanman, Paul, Ito, and Miller (1997) found that participants self-reported more positive affect for imagined black than white partners but exhibited EMG activity indicative of more negative affect for blacks than whites. In a third study with white participants, Vanman et al. administered the Modern Racism Scale (McConahay et al., 1981) and divided participants into high- and low-prejudice groups based on a median split. Participants then viewed a series of stimulus photos of black and white individuals, after each of which they rated the target individual's friendliness. Facial EMG was recorded during presentation of the photos. Consistent with the first two studies, results revealed that participants' self-reported friendliness ratings were higher for black targets than white targets, regardless of prejudice level. However, high-prejudice participants exhibited lower zygomaticus and higher corrugator activity during presentation of black target photos than white target photos, indicating lower positive affect and higher negative affect; in contrast,

low-prejudice participants did not significantly differ in EMG activity between target photo groups. These three studies demonstrate the value of facial EMG in measuring implicit affect underlying attitudes (i.e., prejudice), particularly when self-reported responses are likely to reflect self-presentational concerns.

Vanman, Saltz, Nathan, and Warren (2004) took an additional step, linking implicit negative affect assessed via facial EMG to discriminatory behavior. White participants first decided which of three applicants (two white, one black) should be selected for a prestigious teaching fellowship. Participants then completed the Implicit Associations Test (IAT; Greenwald, McGhee, & Schwartz, 1998), which is designed to assess implicit attitudes about race by comparing reaction times to various pairings of black/white-related words and pleasant/unpleasant-related words. Several weeks later, in an ostensibly unrelated study, participants rated the friendliness of black and white target photos while EMG activity was recorded. Vanman and colleagues computed bias scores for both zygomaticus and corrugator activity by subtracting values during exposure to photos of blacks from values during exposure to photos of whites.

Results indicated that higher zygomaticus bias (i.e., greater activity for white targets than black targets) was associated with selecting a white applicant over a black applicant, even after accounting for the higher probability of selecting a white applicant by chance. Corrugator activity and IAT score failed to predict discrimination. In combination, the findings from Vanman and colleagues (2004) not only establish facial EMG as an implicit measure of prejudice that reveals affect potentially obscured in self-reports but also support the notion that such affect may play a role in carrying out biases in behavior.

SUMMARY

The integration of social psychology and biology foreseen by Shapiro and Crider (1969) has increased markedly over time, particularly in the last decade. Furthermore, this integration has broadened from the connection between social psychology and peripheral physiology on which Shapiro and Crider focused to connections with CNS physiology and endocrine processes as well. In our view, social neuroscience including social psychophysiology is only now in its adolescent period. In our view, its further development will lead to major advances to both the psychological and biological sciences.

The connection between social psychology and peripheral physiology in particular remains viable and strong. The robustness of this connection derives from advances in many areas. The rejection of mind–body dualism and the resulting rise in a more monistic view have eliminated a major epistemological barrier between social psychology and physiology. A logical basis for drawing strong inferences from physiological indexes of social psychological constructs has been elucidated and continues to be refined. Additions to and improvements in existing technologies have enabled investigators to assess multiple peripheral physiological responses over time.

Plausible theories provide the justification for identification of peripheral physiological response patterns that reliably index social psychological constructs and processes. Finally, programmatic lines of research incorporating peripheral physiological indexes of constructs and processes have demonstrated the value of using such measures to test social psychological theories.

NOTES

1. Neither the third nor the fourth edition of the *Handbook of Social Psychology* included a social psychophysiology chapter, attesting to the perception of social psychophysiology as something less than mainstream by many in the field until recently.
2. Approximately 60 mostly social psychologists participated in these intensive, National Science Foundation supported, monthlong training sessions between 1985 and 1990.
3. Dienstbier actually used the pituitary–adrenal–cortical (PAC) axis rather than hypothalamic–pituitary–adrenal (HPA) axis nomenclature in his paper. Here we follow what we perceive to be the more modern nomenclature though we do not wish to imply anything more than PAC activity.

REFERENCES

Ader, R., & Cohen, N. (1975). Behaviorally conditioned immunosuppression. *Psychosomatic Medicine*, *37*, 333–340.

Ahern, S. K., & Beatty, J. (1981). Physiological evidence that demand for processing capacity varies with intelligence. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 121–128). New York: Plenum Press.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 788–894). Worcester, MA: Clark University Press.

Bauer, R. N. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia*, *22*, 457–469.

Beatty, J., & Lucero-Wagoner, B. (2000). The papillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge, UK: Cambridge University Press.

Berkowitz, L. (1969). Social motivation. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 3, pp. 50–135). Reading, MA: Addison-Wesley.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space and the laws of autonomic constraint. *Psychological Review*, *98*, 459–487.

Blascovich, J. (2000). Psychophysiological methods. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 117–137). Cambridge, UK: Cambridge University Press.

Blascovich, J. (in press). Challenge, threat, and health. In J. Shah & W. Gardner (Eds.), *Handbook of motivation science.* New York: Guilford Press.

Blascovich, J., Ernst, J. M., Tomaka, J. Kelsey, R. M., Salomon, K. A., & Fazio, R. H. (1993). Attitude as a moderator of autonomic reactivity. *Journal of Personality and Social Psychology*, *64*, 165–176.

Blascovich, J., & Ginsburg, G. P. (1978). Conceptual analysis of risk taking in "risky shift" research. *Journal for the Theory of Social Behavior*, *8*, 217–230.

Blascovich, J., & Kelsey, R. M. (1990). Using cardiovascular and electrodermal measures of arousal in social psychological research. *Review of Personality and Social Psychology*, *11*, 45–73.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In A. J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 59–82). Paris: Cambridge University Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., & Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology*, *80*, 253–267.

Blascovich, J., Mendes, W., Hunter, S., & Salomon, K. (1999). Social facilitation, challenge, and threat. *Journal of Personality and Social Psychology*, *77*, 68–77.

Blascovich, J., Mendes, W. B., & Seery, M. (2002). Intergroup encounters and threat: A multi-method approach. In D. Mackie & E. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 89–110). New York: Psychology Press.

Blascovich, J., Mendes, W. B., Tomaka, J., Salomon, K., & Seery, M. (2003). The robust nature of challenge and threat: A reply to Wright and Kirby. *Personality and Social Psychology Review*, *7*, 234–243.

Blascovich, J., Seery, M., Mugridge, C., Weisbuch, M., & Norris, K. (2004). Predicting athletic performance from cardiovascular indicators of challenge and threat. *Journal of Experimental Social Psychology*, *40*, 683–688.

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 1–51). New York: Academic Press.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency. *Social Cognition*, *12*, 103–128.

Bradley, M. M. (2000). Emotion and motivation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 602–642). Cambridge, UK: Cambridge University Press.

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind* (2nd ed.). Boston: Allyn & Bacon.

Cacioppo, J. T., Berntson, G. G., & Crites, S. L., Jr. (1996). Social neuroscience: Principles of psychophysiological arousal and response. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 72–101). New York: Guilford Press.

Cacioppo, J. T., Bernston, G. G., & Tassinary, L. G. (2000). *Handbook of psychophysiology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Cacioppo, J. T., Klein, D. J., Bernston, G. G., & Hatfield, E. (1993). The psychophysiology of emotion. In M. L. & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 67–83). New York: Guilford Press.

Cacioppo, J. T., & Petty, R. E. (1979). Attitudes and cognitive response: An electrophysiological approach. *Journal of Personality and Social Psychology*, *37*, 2181–2199.

Cacioppo, J. T., & Petty, R. E. (1981a). Electromyograms as measures of extent and affectivity of processing. *American Psychologist*, *36*, 441–456.

Cacioppo, J. T., & Petty, R. E. (1981b). Electromyographic specificity during covert information processing. *Psychophysiology*, *35*, 441–456.

Cacioppo, J. T., & Petty, R. E. (Eds.). (1983). *Social psychophysiology: A sourcebook*. New York: Guilford Press.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*, 260–268.

Cacioppo, J. T., & Tassinary, L. G. (Eds.). (1990). *Handbook of psychophysiology*. Cambridge, UK: Cambridge University Press.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2000). Psychophysiological inference. In J. T. Cacioppo, L. G. Tassinary, & G. Bernston (Eds.), *Handbook of psychophysiology*

(2nd ed., pp. 3–23). Cambridge, UK: Cambridge University Press.

Cacioppo, J. T., Tassinary, L. G., & Fridlund, A. J. (1990). The skeletomotor system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 325–384). New York: Cambridge University Press.

Campbell, D. T. (1988a). Convergent and discriminant validation by the multitrait–multimethod matrix. In E. S. Overman (Ed.), *Methodology and epistemology for social science: Selected papers–Donald. T. Campbell* (pp. 37–62). Chicago: University of Chicago Press.

Campbell, D. T. (1988b). Definitional vs. multiple operationism. In E. S. Overman (Ed.), *Methodology and epistemology for social science: Selected papers–Donald Campbell* (pp. 31–36). Chicago: University of Chicago Press.

Cannon, W. B. (1928). *Bodily changes in pain, hunger, fear, and rage* (2nd ed.). New York: Appleton-Century-Crofts.

Carpenter, W. B. (1874). *Principles of mental physiology, with their applications to the training and discipline of the mind, and the study of its morbid conditions.* London: Henry S. King & Co.

Crocker, J., Major, B. N., & Steele, C. (1998). Social stigma. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 504–553). Boston: McGraw-Hill.

Damasio, A. (1994). *Descartes error: Emotion, reason, and the human brain.* New York: Avon Books.

Dambrun, M., Despres, G., & Guimond, S. (2003). On the multifaceted nature of prejudice: Psychophysiological responses to ingroup and outgroup ethnic stimuli. *Current Research in Social Psychology*, *8*, 200–204.

Darwin, C. (1965). *The expression of the emotions in man and animals.* Chicago: University of Chicago Press. (Original work published 1872)

Dawson, M. E., Schell, A. M., & Filion, D. L. (2000). The electrodermal system. Interpersonal processes. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 200–223). Cambridge, UK: Cambridge University Press.

Davison, K. P., & Pennebaker, J. W. (1996). Social psychosomatics. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 102–129). New York: Guilford Press.

Descartes, R. (1972). *Treatise of man* (T. Steele Hall, Trans. & Commentator). Cambridge, MA: Harvard University Press. (Original work published 1662)

Descartes, R. (1680). *Six metaphysical meditations; Wherein it is proved that there is a God. And that man's mind is really distinct from his body* (W. Melyneux, Trans.). London: Printed by B.G. for Benj. Tooke.

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review*, *96*, 84–100.

Ekman, P. (1993). Facial expression of emotion. *American Psychologist*, *48*, 384–392.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system (FACS): A technique for the measurement of facial actions.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion*, *10*, 159–168.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*, 1208–1210.

Fazio, R. H. (1989). On the power and functionality of attitudes: The role of attitude accessibility. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 153–179). Hillsdale, NJ: Erlbaum.

Fowles, D. C. (1975). *Clinical applications of psychophysiology*. New York: Columbia University Press.

Friedman, M., & Rosenman, R. H. (1974). *Type A behavior and your heart.* New York: Knopf.

Gardner, W. L., Gabriel, S., & Diekman, A. B. (2000). Interpersonal processes. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 643–664). Cambridge, UK: Cambridge University Press.

Glaser, R., Rice, J., Sheridan, J., Fertel, R., Stout, J., Speicher, C., et al. (1987). Stress-related immune suppression: Health implications. *Brain, Behavior, and Immunity*, *1*, 7–20.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association task. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin*, *27*, 889–898.

Harré, R. (1967). *An introduction to the logic of the sciences*. London: Macmillan.

Heinrich, W. (1896). Die aufmerksamkeit und die Funktion der Sinnesorgane. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, *9*, 342–388.

Hess, E. H., & Polt, J. M. (1964). Pupil size as related to interest value of visual stimuli. *Science*, *143*, 1190–1192.

Hunter, S. B. (2001). *Performance under pressure: The impact of challenge and threat states on information processing*. Unpublished doctoral dissertation, University of California, Santa Barbara.

James, W. (1890). *The principles of psychology*. New York: Holt.

Jones, E. E., Farina, A., Hastorf, A. H., Markus, H. Miller, D. T., & Scott, R. A. (1984). *Social stigma: The psychology of marked relationships*. New York: Freeman.

Kasprowicz, A. L., Manuck, S. B., Malkoff, S. B., & Krantz, D. S. (1990). Individual differences in behaviorally evoked cardiovascular response: Temporal stability and hemodynamic patterning. *Psychophysiology*, *27*, 605–619.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*, 163–204.

Kernis, M. H. (1993). The roles of stability and level of self-esteem in psychological functioning. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 167–182). New York: Plenum Press.

Kernis, M. H., & Waschull, S. B. (1995). The interactive roles of stability and level of self-esteem: Research and theory. *Advances in Experimental Social Psychology*, *27*, 93–141.

Kiecolt-Glaser, J. K., & Glaser, R. (1995). Psychoneuroimmunology and health consequences: Data and shared mechanisms. *Psychosomatic Medicine*, *57*, 269–274.

Lacey, J. L., & Lacey, B. I. (1958). Verification and extension of the principle of autonomic response–stereotypy. *American Journal of Psychology*, *71*, 50–73.

Landis, C. (1930). Psychology and the psychogalvanic reflex. *Psychological Review*, *37*, 381–398.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon & Schuster.

Leibniz, G. (1695). The new system of the nature of substances. In R S. Woolhouse & R. Franks (Eds.), *Leibniz's "new system"* (pp. 7–37). Oxford, UK: Oxford University Press.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.

Lewes, G. H. (1877). The physical basis of mind. London: Trübner & Co.

Mackworth, J. F. (1963). The duration of the visual image. *Canadian Journal of Psychology*, *17*, 62–81.

Mandler, G., Nakamura, Y., & Van Zandt, B. J. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 646–648.

Manuck, S. B., Kamarck, T. W., Kasprowicz, A. L., & Waldstein, S. R. (1993). Stability and patterning of behaviorally evoked cardio-

vascular reactivity. In J. Blascovich & E. S. Katkin (Eds.), *Cardiovascular reactivity to psychological stress and disease: An examination of the evidence* (pp. 83–108). Washington, DC: American Psychological Association.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution*, *25*, 563–579.

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during interactions with white and black men. *Personality and Social Psychology Bulletin*, *28*, 939–952.

Mendes, W. B., Blascovich, J., Major, B., & Seery, M. (2001). Effects of social comparisons on challenge and threat reactivity. *European Journal of Social Psychology*, *31*, 477–479.

Mendes, W. B., Reis, H. T., Seery, M., & Blascovich, J. (2003). Cardiovascular correlates of emotional disclosure and suppression: Do content and gender context matter. *Journal of Personality and Social Psychology*, *84*, 771–792.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective*. New York: Plenum Press.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Porges, S. W. (1995). Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. *Psychophysiology*, *32*, 301–318.

Quigley, K. S., Barrett, L. F., & Weinstein, S. (2002). Cardiovascular patterns associated with threat and challenge appraisals: A within-subjects analysis. *Psychophysiology*, *39*, 292–302.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency of affective judgments. *Psychological Science*, *9*, 45–48.

Robinson, H. (2003, Fall). Dualism In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Available: *http://plato.stanford.edu/archives/fall2003/entries/dualism/*

Robinson-Riegler, G. L., & Robinson-Riegler, B. (2004). *Cognitive psychology: Applying the science of the mind*. New York: Allyn & Bacon.

Rousselle, J. G., Blascovich, J., & Kelsey, R. M. (1995). Cardiorespiratory response under combined psychological and exercise stress. *International Journal of Psychophysiology*, *20*, 49–58.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*, 379–399.

Schiff, J. M. (1875). *La pupille considere comme esthesiometre* (R. G. de Choisity, Trans.). Paris: J.B. Baeilliere.

Seery, M., Blascovich, J., Weisbuch, M., & Vick, S. B. (2004). The relationship between self-esteem, self-esteem stability, and cardiovascular reactions to performance feedback. *Journal of Personality and Social Psychology*, *87*, 133–145.

Selye, H. (1950). *The physiology and pathology of exposure to stress: A treatise based on the concepts of the general adaptation syndrome and the diseases of adaptation*. Montreal: Acta.

Shapiro, D., & Crider, A. (1969). Psychophysiological approaches in social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 3, pp. 19–49). Reading, MA: Addison-Wesley.

Shaver, P. R., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further explorations of a prototype approach. *Journal of Personality and Social Psychology*, *52*, 1061–1086.

Solomon, G. F., & Moos, R. H. (1965). The relationship of personality to the presence of rheumatoid factor in asymptomatic rela-

tives of patients with rheumatoid arthritis. *Psychiatric Medicine*, *27*, 350–360.

Spinoza, B. (1677). *Opera posthuma, quorum series post praefationem exhibetur*. Amstelodami: J. Rieuwertsz.

Stemmler, G. (1989). The autonomic differentiation of emotions revisited: Convergent and discriminant validation. *Psychophysiology*, *26*, 617–632.

Sternbach, (1966). *Principles of psychophysiology*. New York: Academic Press.

Tassinary, L. G., & Cacioppo, J. T. (2000). The skeletomotor system: Surface electromyography. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 163–199). Cambridge, UK: Cambridge University Press.

Tomaka, J. (1993). *Effects of cognitive appraisal on autonomic, affective, and behavioral reactions to stress*. Unpublished doctoral dissertation, State University of New York at Buffalo.

Tomaka, J., & Blascovich, J. (1994). Effects of justice beliefs on cognitive appraisal of and subjective, physiological, and behavioral responses to potential stress. *Journal of Personality and Social Psychology*, *67*, 732–740.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *65*, 248–260.

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*, 63–72.

Tranel, D., & Damasio, H. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science*, *228*, 1453–1454.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology*, *4*, 400–408.

Vanman, E. J., Paul, B. Y., Ito, T. A., & Miller, N. (1997). The modern face of prejudice and structural features that moderate the effect of cooperation on affect. *Journal of Personality and Social Psychology*, *73*, 941–959.

Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial discrimination by low-prejudiced whites: Facial movements as implicit measures of attitudes related to behavior. *Psychological Science*, *15*, 711–714.

Waid, W. M. (1984). *Sociophysiology*. New York: Springer-Verlag.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Wilson, T. D., Lindzey, S., & Schooler, T. (2000). A model of dual attitudes. *Psychological Review*, *107*, 1–26.

Wright, R. A., & Kirby, L. D. (2001). Effort determination of cardiovascular response: An integrative analysis with applications in social psychology. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 255–307). San Diego, CA: Academic Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs*, *9*(2, Pt. 2), 1–27.

Zillmann, D., & Zillmann, M. (1996). Psychoneuroendocrinology of social behavior. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 39–71). New York: Guilford Press.

CHAPTER 3

Social Cognitive Neuroscience
Historical Development, Core Principles, and Future Promise

KEVIN N. OCHSNER

The supermarket checkout line provides interesting lessons about the human psyche. Beside the social norms that dictate that we stand in line and politely pay our bill, and before we encounter the disaffected teenage checker, there is the point-of-purchase magazine gambit. Strategically placed above the checkout conveyor belt are rows of magazines heralding the latest pop-culture happenings. Typically, two types of cover stories clamor for attention: those about the relationship hijinks of high-profile people and those about new businesses and technology that could "change the way we live and work." Since its emergence at the beginning of the 21st century, social cognitive neuroscience (SCN) has been purported to be both. At turns, SCN has been billed either as a hot new power coupling of social psychology and cognitive neuroscience or as a fast-growing research startup looking for investment capital.

The fact that we use relationship metaphors to describe SCN may not be that surprising. Indeed, the supermarket checkout line is a regular reminder that humans are fundamentally social beings (Fiske, 1991; Fiske, Kitayama, Markus, & Nisbett, 1998). Magazines capitalize on the facts that we care a great deal about our personal and our professional relationships and that stories and metaphors help us understand them (Lakoff, 1987, 1993). The beauty of metaphors is that when correctly applied they can help us intuitively grasp what something means. The danger of metaphors is that when incorrectly applied they can lead us to overlook important differences between the objects of comparison.

In this regard, using different relationship metaphors to describe SCN invites different types of questions about its nature. On one hand, a personal metaphor might lead to questions about how parent disciplines begat SCN, what they were thinking when they did so, and how SCN will develop and mature. On the other hand, a professional metaphor might lead to questions about SCN's business plan, principles for effective production, and growth potential.

The goal of this chapter is to address the nature of SCN. Toward that end, it uses the personal and professional relationship metaphors to organize discussion of questions about its practice, its principles, and its promise. The first section of the chapter takes the personal metaphor as a starting point for describing SCN's historical roots and developmental progression. The second section uses the professional metaphor to springboard consideration of core principles that govern its practice. The third and final section highlights promising directions for future work.

The goal of this chapter is not to describe basic psychological and neural principles that underlie a specific type of socioemotional ability. Rather, the goal is to describe the development and nature of basic principles that underlie SCN *as an approach* that can be applied to investigating any number of topics (Blakemore & Frith, 2004; Blakemore, Winston, & Frith, 2004; Ochsner & Lieberman, 2001). For principles related to specific types of ability, the reader is referred to other chapters in this volume, and to SCN work on social cognition and theory

of mind (Gallagher & Frith, 2003; Keysers & Perrett, 2004; Lieberman, 2003; Puce & Perrett, 2003; Saxe, Carey, & Kanwisher, 2004), self-reflection and self-perception (Heatherton, Macrae, & Kelley, 2004; Lieberman & Pfeifer, 2005), perception of faces and other nonverbal cues (Allison, Puce, & McCarthy, 2000; Calder et al., 2002; Haxby, Hoffman, & Gobbini, 2002), and emotion and self-regulation (Beer, Shimamura, & Knight, 2004; Ochsner & Feldman Barrett, 2001; Ochsner & Gross, 2004, 2005; Ochsner & Schacter, 2000).

WHAT'S IN A NAME?: TRACING THE HISTORICAL DEVELOPMENT OF SOCIAL COGNITIVE NEUROSCIENCE

The German psychologist Hermann Ebbinghaus (1908) once wrote, "Psychology has a long past, but a short history," by which he meant that the roots of modern psychological inquiry run long and deep, but the field operates as if current psychological research is a recent development. The same might be said of SCN: Amid attention to its recent growth and development, lessons from its long past may be overlooked. In this section, we chronicle some of the past and present debates and disciplines that have contributed to SCN's gene pool and some developmental milestones that have marked its growth. The goal is to understand the interdisciplinary origins of SCN, its relationships to allied fields, and how and why SCN has a distinct identity.

The Genealogical Tree: From Animal Models to Phineas Gage

As illustrated in Figure 3.1, SCN is an interdisciplinary field that seeks to explain social phenomena at three levels of analysis: the social level of behavior and experience, the cognitive level of mental representations and processes, and the neuroscience level of brain systems. Although articles, laboratories, and conferences bearing the name "social cognitive neuroscience" have only appeared since the turn of the century, SCN has a long and important ancestry.
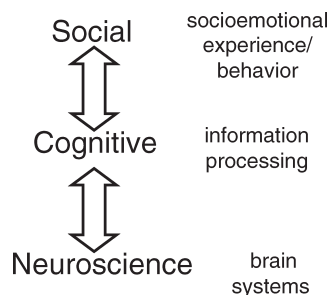
FIGURE 3.1. The three levels of analysis for social cognitive neuroscience.

SCN's Close Relatives

During the past century there have been a number of important findings in psychology and neuroscience that can be seen as precursors to SCN. The majority of this work has been done with animal models, with interest in human work growing during the past quarter century.

A LONG LINE OF ANIMAL MODELS

A great deal of work on the brain bases of social emotional behavior has employed animal models. In general, this work has followed two threads. The first concerns the study of prosocial behavior. Some of the earliest and best-known examples of this work came from Harry Harlow, whose studies of the effects of maternal deprivation on the development of social function in young primates is still cited an introductory textbooks today (Harlow, Dodsworth, & Harlow, 1965; Harlow & Harlow, 1962). In a similar vein, the classic studies of Kluver and Bucy purported to show that damage to the inferior temporal lobes resulted in a severe disruption of sexual, feeding, social, and maternal behavior (Kluver & Bucy, 1939). Modern work, however, has demonstrated that these effects are not wholly reliable. For example, Amaral and others have shown that the specific deficits observed in social function (e.g., showing either social disinhibition or fear of conspecifics) may depend on the specific location of cortical lesion in the temporal lobe, and in particular on the nature and extent of the damage to the amygdala (Amaral et al., 2003; Bauman, Lavenex, Mason, Capitanio, & Amaral, 2004a, 2004b). Consistent associations between impaired primate social behavior and damage to the amygdala and orbitofrontal cortex (among other regions) led Brothers (1990) to hypothesize that these systems together comprise a "social brain." Complementary work in rodents has less often concerned the amygdala, but more often other subcortical nuclei and neurohormonal factors important for mating, pair bonding, or maternal care. For example, Insel and colleagues have shown that oxytocin and vasopressin are important for promoting memory for and bonds with conspecifics (Insel & Fernald, 2004), and Meaney and colleagues have shown that early pup experiences of maternal licking and grooming set a threshold for subsequent responses to stress manifested at multiple levels of the neuroaxis (Meaney, 2001).

The second thread in animal research has emphasized the motivational and emotional, rather than the social, aspects of behavior. Some of the earliest work of this kind appeared during the first half of the century when numerous brain stimulation and lesion studies identified subcortical nuclei essential for the manifestation of various types of species-specific aggressive, fearful, and sexual behaviors (e.g., Davey, Kaada, & Fulton, 1949; Fangel & Kaada, 1960; Kaada, Andersen, & Jansen, 1954). This work was the impetus for some of the first "neural circuit" theories of emotion (e.g., Cannon, 1987; Papez, 1958). Building on this work, Paul MacLean later advanced the concept of the triune brain, alluded to above,

in which impulses for both prosocial (e.g., play and maternal care) and defensive (e.g., aggression and flight) behavior depended on evolutionarily old brain systems humans share with other mammals (MacLean, 1969). Higher neocortical systems were thought to control the expression of these innate drives, motivations, and emotions. Following MacLean's lead, Panksepp (1998) later described multiple subcortical neural systems dedicated to distinct motivated and/or emotional behaviors, ranging from play to maternal care to aggression and love. Then, in the late 1980s, emotion came to the fore in neuroscience with LeDoux's (2000) seminal work on the role of the amygdala in conditioned fear.

RECENT REEMERGENCE OF HUMAN RESEARCH

Interestingly, in comparison to animal work that has seen decades of concerted effort, human research on the neural systems involved in social and emotional behavior has been slow to develop. This has not been due to a lack of early interest in these questions, however, either from social psychologists or from neuroscientists. As is recognized increasingly today (e.g., Heatherton, 2004), social psychologists early on recognized the importance of understanding the brain bases of social behavior. In one of the first social psychology textbooks, Floyd Allport (1924) wrote:

> The chief contributions of the cortex to social behavior may be summarized as follows: It underlies all solutions of human problems, which are also social problems, and makes possible their preservation in language, customs, institutions, and inventions. It enables each new generation to profit by the experience of others in learning this transmitted lore of civilization. It establishes habits of response in the individual for social as well as for individual ends, inhibiting and modifying primitive self-seeking reflexes into activities which adjust the individual to the social as well as to the non-social environment. Socialized behavior is thus the supreme achievement of the cortex. (p. 31)

Similar early emphasis can be found in one of the most famous neuropsychological cases of all time, which concerned disruptions of socioemotional behavior. In the late 1800s the "mysterious" case of Phineas Gage was well documented and generated great interest (Goldenberg, 2004; Macmillan, 2000). After damage to his orbitofrontal cortex in a freak railroad construction accident, Gage was described as being "no longer Gage." Gage had intact cognitive faculties but apparently diminished ability to conform to social norms (making lewd comments, inappropriate jokes, displaying inappropriate affect, etc.).

Despite this early interest, human research on the social or emotional brain was largely absent until the 1980s. The reasons for this dry spell are likely fourfold. First, it was during the first half of the 20th century that radical behaviorism was the dominant force in psychology. Because behaviorism was decidedly and antagonistically nonmentalistic, and social and emotional behaviors have a strong experiential component, the study of such behaviors was left out of the research mix.[1] Second, as described earlier, socioemotional behaviors often have been conceived as more primitive and animalistic than are our so-called higher cognitive faculties. In a sense, affect was noise in the cognitive signal, and the effects of affect were to be eliminated. Thus, when the yoke of behaviorism was cast aside and the cognitive revolution reintroduced mentalistic concepts to psychology, emotion again was left out of the cognitive science research mix (Gardner, 1985). Third, both in behaviorism and in cognitive science (as well as their descendants—behavioral neuroscience and cognitive neuroscience) there has been an emphasis on identifying species-general principles that govern behavior. This emphasis on the general leaves out social and emotional factors that vary by individuals and by contexts. Fourth, and last, psychologists and neuroscientists simply lacked the tools to easily and precisely study the brain bases of socioemotional phenomena in humans. Animal researchers could stimulate or ablate brain systems, but human researchers were left to the study the consequences of uncommon brain-damaging accidents of nature (such as stroke).

An early challenge to this status quo came from social psychology in the 1980s, when the use of peripheral psychophysiological measures to index autonomic nervous system (ANS) activity seemed to promise a means for studying the linkage between psychological and biological mechanisms governing social behavior. Unfortunately, such measures turned out to have only limited value for this purpose because (1) they do not directly measure the operation of the brain systems that implement psychological processes and (2) the measures themselves often show little differentiation across qualitatively different task contexts. As a consequence, correlations often are weak between ANS measures and either self-report or performance measures of behavior (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000), which renders these measures suitable for studying the *physiological consequences* of particular types of emotion or thought but less well-suited for drawing inferences about *information-processing mechanisms* (Blascovich, Mendes, Tomaka, Salomon, & Seery, 2003; Tomaka, Blascovich, Kibler, & Ernst, 1997; Wright & Kirby, 2003).

Be that as it may, interest in human neuropsychological research on the brain bases of socioemotional behavior was growing slowly, and appeared on multiple fronts in the 1980s. Examples include, but are not limited to, research on the involvement of the right hemisphere in nonliteral aspects of language, such as humor and metaphor (Brownell, Simpson, Bihrle, Potter, & Gardner, 1990; Kaplan, Brownell, Jacobs, & Gardner, 1990; Winner, Brownell, Happe, Blum, & Pincus, 1998); studies of the psychological mechanisms underlying face perception and their breakdown in prosopagnosia (Ellis, 1992; Farah, 1990; Young & Ellis, 1989); descriptions of Capgras syndrome, which involves delusions that loved ones have been replaced by exact replicas (Ellis & Lewis, 2001); demonstrations that amnesics who lack explicit memory for melodies or encounters with people nonetheless can acquire preferences for them (Johnson, Kim,

& Risse, 1985); conceptually sophisticated psychophysiological work demonstrating distinct patterns for appraisals of threat versus challenge (Blascovich et al., 1992; Tomaka et al., 1997); scalp electrophysiological studies identifying cortical correlates of attitudes and evaluations (Cacioppo, Crites, Gardner, & Berntson, 1994; Crites, Cacioppo, Gardner, & Berntson, 1995); and a resurgence of interest in the effects on decision making and social behavior of orbitofrontal lesions like those suffered by Phineas Gage (Bechara, Damasio, & Damasio, 2000; Beer, Heerey, Keltner, Scabini, & Knight, 2003; Damasio, 1994).

This work set the stage for increased availability and common usage of functional imaging techniques such as positron emotion tomography in the mid-1980s, and functional MRI in the early 1990s, which enabled researchers to study the cortical and subcortical brain bases of phenomena in healthy normal individuals. The first topics to be studied using functional imaging were classic cognitive psychological phenomena involving language, attention, memory, and vision. This was due in large part to the fact that many of the best cognitive psychologists quickly became cognitive neuroscientists, and imaging work on their topics of interest has dominated cognitive neuroscience since its inception. But in the late 1990s, something changed and a new field emerged that was devoted specifically to the use of neuroscience methods to study the brain bases of socioemotional phenomena.

From Zeitgeist to Distinct Identity

What was it that changed? The preceding review points to the development of a Zeitgeist with three crucial elements. First, salient animal (e.g., work of LeDoux and Panksepp) and human (e.g., work of Damasio) studies of social and emotional behavior had achieved a great deal of notoriety, in part because they represented modern approaches to classic problems in both biological and social science that were discussed earlier. Second, functional imaging had become highly accessible in many research institutions, and high-profile imaging publications regularly received a great deal of attention. For researchers across many disciplines, these two factors made salient the questions and methodologies that when combined later would form the basis of social cognitive neuroscience.

Of course, as described previously, researchers have been using neuroscience methods to study questions about socioemotional phenomena for quite a long time. But researchers had not yet realized that the seemingly disparate strands of research listed earlier could be woven into a coherent whole. The third element of the Zeitgeist helped spark this realization. This element was a research climate very favorable to interdisciplinary research in which the past two decades had seen numerous new terms coined to describe distinct interdisciplinary fields. For example, the term "social cognition" came into common usage in the early 1980s to refer to the use of cognitive psychological theories and methods to study phenomena typically of interest to social psychologists. Then in the late 1980s, the term "cognitive neurosci-

ence" was coined to refer to the use of neuroscience methods to study the brain bases of phenomena typically studied by cognitive psychologists. In the early 1990s, the term "social neuroscience" was coined to refer quite broadly to any research that linked the biological and social levels of analysis. This move broadened the use of an earlier term, "social psychophysiology," which had been used to describe the initially promising but ultimately limited movement in social psychology (described earlier) toward using peripheral autonomic measures as indices of underlying psychological processes. And finally, in the 1990s, the term "affective neuroscience" gave a name to the growing area of research (also described earlier) aimed at discovering the affective/emotional functions of specific brain systems.

In this context, social psychologists and cognitive neuroscientists in a number of locations began to use the term "social cognitive neuroscience" to refer to the use of cognitive neuroscience methods to study socioemotional phenomena. The first papers using the term described SCN as a marriage between social cognition on the one hand and cognitive neuroscience on the other (Lieberman, 2000; Ochsner & Lieberman, 2001; Ochsner & Schacter, 2000). Two motivations prompted the use of this term. First, it was thought to provide an accurate label for a new kind of interdisciplinary research that capitalized on what its parent disciplines have in common, at the same time making good use of their unique strengths. Both social cognition and cognitive neuroscience are concerned with information-processing mechanisms: whereas social cognition links the study of particular kinds of intra- and interpersonal experiences and behaviors to information-processing models of psychological mechanisms, cognitive neuroscience links these models to their neural substrates using neuropsychological, electrophysiological, and functional imaging methodologies. SCN puts it all together. The second motivation was pragmatic. It was hoped that the term "social cognitive neuroscience" might be intuitively appealing to both social cognition and cognitive neuroscience researchers who would see the name of their field in the new term and might therefore be encouraged to participate in it. Despite these principled hopes, however, in actual practice the nature of a field is defined by those who work within it. And as usage of the term "SCN" began to grow, important questions arose about the boundaries of the field.

Staking a Claim to a Research Domain

Questions about the domain and scope of a new research domain are common and important to address during its formative years. In the case of SCN, at least two important questions concerning its scope and boundaries need to be addressed.

The first question is simply whether it is useful to try to define a distinct new area of research. Or, in other words: With all those other interdisciplinary fields and subfields already out there, why coin a term for a new one? The answer, of course, depends on whether the new field is truly distinct, and whether the new term proves useful for

guiding, promoting, and drawing new researchers into conducting a specific and potentially new kind of research. In this regard, it is useful to consider SCN's relationship to its closest neighbors: *social neuroscience* (SN)*, affective neuroscience* (AN), and *cognitive neuroscience* (CN)*. Each of these terms and the fields they define have been around for a decade or more.

In the case of SN, although the term originally was intended to be quite broad, within psychology, "SN" initially was used to describe human studies that linked social variables to psychophysiological, endocrine, and immunological measures (Cacioppo, 1994). Animal researchers also began using the term "SN" to describe their research linking neuroendocrine and subcortical brain systems to affiliate and bonding behaviors (Insel & Fernald, 2004). In this context SCN was a newer term that appealed to social cognition and cognitive neuroscience researchers who did not identify themselves with the types of research SN had been used to describe previously. Similarly, the term "affective neuroscience" had been used most often by animal and clinical researchers studying the cortical and subcortical bases of so-called basic emotions and their role in affective disorders (Davidson, Jackson, & Kalin, 2000; Panksepp, 1998). Researchers interested in SCN also were interested in affect but construed more broadly the range of affect-related phenomena they wished to address. For SCN, this range included phenomena not typically of interest to affective neuroscientists such as attitudes, stereotyping, person perception, self-reflection, and decision making. An interest in such topics also differentiates SCN from CN, whose domain is typically conceived as the study of so-called basic mental abilities, such as memory and attention, that may be deployed in any number of social or nonsocial contexts. Thus, the term "SCN" appealed to researchers who (1) were interested in using cognitive neuroscience methods to study a wide array of socio-emotional phenomena, (2) wanted to use this combined methodology to elucidate the information processing level of analysis, and (3) did not identify with the types of research questions and content areas previously associated with related fields, such as SN, AN, and CN.

That being said, the kinds of research to which a given term refers are somewhat fluid, and certainly evolve over time. For example, some researchers use the terms "SN" and "SCN" interchangeably, whereas others see them as distinct but interrelated. Perhaps the most useful way to think about this issue is in terms of a part–whole relationship. As originally intended, SN can be used broadly to describe many types of research that link social phenomena to their biological substrates described at any one of many levels of analysis, ranging from the cortical region to the neurotransmitter system. By contrast, SCN refers to an important subset of this larger domain, where researchers specifically integrate social cognitive and cognitive neuroscientific methods (see Figure 3.6, and the section "Mapping a Road Toward the Future").

This brings us to the second question facing an emerging field of research: What are its boundaries? Or in the case of SCN, what is *social* about SCN?[2] A partial answer to this question was provided by examining historical boundaries between disciplines that help define the relationship of SCN to its neighbors. For the rest of the answer one must understand that what is *social* about SCN is determined in large part by the proclivities of researchers who call themselves *social cognitive neuroscientists*.

Primarily two types of researchers have rallied around the SCN flag. Many are cognitive neuroscientists who bring with them numerous habits and assumptions about the way in which any type of phenomenon should be studied using neuroscience methodologies. As discussed later, this has led them to favor the use of memory, perception, and attention paradigms to study neural responses to visual perceptual stimuli that have social signal value (such as faces). For these researchers, what is *social* about SCN is that the purview of cognitive neuroscience has been broadened to include the processing of "basic" social stimuli.

The other main group drawn to SCN are social cognition researchers who also bring with them some important assumptions. In particular, they are interested in studying a much wider range of phenomena. Indeed, social psychology's purview includes the study of a wide variety of interpersonal phenomena, ranging from nonverbal perception to persuasion, as well as many intrapersonal phenomena, such as self-perception and self-regulation. For a social cognition researcher, what is *social* about SCN is that involves unpacking what is special about the way people—with all their motivations, goals, and contexts—process stimuli, and what happens when the stimuli are themselves social. As is discussed in the section, "Principles Governing the Practice of SCN," incorporates both of these perspectives in its core principles concerning the types of *contexts* and *content* with which the field is concerned.

Following the Parent's Lead: Taking One Step Down to Take Many Steps Forward

SCN is in the interesting position of viewing itself either as *social cognition* plus *neuroscience* methods or as *cognitive neuroscience* plus *social* content. However one views it, the new addition can be seen as broadening its scope, explanatory power, and conceptual breadth beyond that of its parent discipline. The addition of a new level of analysis to an existing area of research is old hat for social cognition and cognitive neuroscience, for similar transitions were responsible for the inception of each of these parent disciplines.

In the case of social psychology, a shift down one level of analysis from the social to the information-processing level marked the birth of social cognition. This shift has been credited as providing an answer to the, "crisis in social psychology" that happened in the late 1970s (Taylor, 1998). At that time social psychology lacked a unifying conceptual framework to describe the similarities and differences between different phenomena. Researchers working on seemingly similar topics came up with their own individualized lists of factors that predicted long lists of dependent variables. The information-processing language of cognitive psychology offered a way out: by

appealing to concepts such as automatic and controlled processing, accessibility and spreading activation, selective attention, and schemas and scripts, social psychologists could start providing theoretical explanations that described the processes linking these lists together, which gave their theories greater coherence and staying power.

The case of cognitive neuroscience is quite parallel: Its birth was marked by a shift down one level of analysis from the information-processing level to the level of neural substrates. This move was prompted by a number of factors, not the least of which was a desire to use new methodologies to obtain new kinds of data that can help constrain information-processing models of cognitive phenomena (Ochsner & Kosslyn, 1999; Posner & DiGirolamo, 2000). As has been described in greater detail elsewhere, neuroscience has been used to provide converging evidence concerning the existence of multiple memory systems, the nature of mental representations underlying visual mental imagery, and the fractionation of attention into multiple interacting subsystems (e.g., Kosslyn, Ganis, & Thompson, 2001; Posner & Petersen, 1990; Schacter, 1997). For cognitive psychologists, the power of neuroscience methods is that patterns of brain activations or neuropsychological deficits can be used to draw inferences about the number and nature of underlying psychological mechanisms (Kosslyn, 1999; Ochsner & Lieberman, 2001).

By taking a step down, social psychology became social cognition,[3] and cognitive psychology became cognitive neuroscience. Thereafter, each field took many conceptual and empirical steps forward. SCN's emergence can be construed as either other a step down for social cognition or a step up for cognitive neuroscience. Either way, SCN is following in its parent's footsteps.

From Conception to Coherence: Milestones on the Road to Maturity

Once the term "SCN" began being used and gained currency in the research world, investigators scattered across numerous disciplines began to feel that they might share an identity. One of the key events in the crystallization of a singular identity for SCN was the first stand-alone meeting dedicated to the topic held in April 2001 at UCLA. This 2½-day meeting was organized not just by social psychologists and cognitive neuroscientists, but also by developmental psychologists, anthropologists, and political scientists. The meeting included attendees from all these disciplines and more, ranging from health psychology to behavioral neuroscience. The makeup of the organizers and attendees of this meeting is significant because it signaled that from the get-go, SCN offered a banner under which scientists interested in studying socioemotional phenomena at multiple levels of analysis could rally and find like-minded individuals whose work would be very relevant to their own.

In this regard, the role of developmental psychology in the growth of SCN is particularly important. Developmental psychologists had used the term "theory of mind" to describe the social and emotional deficits suffered by children with autism. Such children do not understand that other humans are agentic beings guided by internal mental states that describe goals, feelings, wants, and desires. Rather, they perceive other humans to be "sacks of flesh" that move unpredictably (Baron-Cohen, 1995). Psychologists studying autism had already been developing models of the disorder that cut across the social, information-processing, and neural levels of analysis. This multilevel approach was exemplified by the 1995 book *Mindblindness*, by Simon Baron-Cohen, which presented an empirical and theoretical account of autism that interpreted behavioral experiments in terms of hypothetical neural substrates. This book can be seen as one of the earliest examples of a social cognitive neuroscience analysis of a phenomenon, even though it predated the coalescence of the field. As described below, from the outset one of the strongest research programs within SCN concerned the neural bases of the social cognitive processes that support theory of mind and related abilities.

Since the UCLA conference, there have been numerous markers of growth in SCN research. Small private conferences dedicated solely to SCN work have been held at institutions such as Dartmouth, the University of Chicago, and Princeton. On the national level, starting in 2004 SCN preconferences were held prior to the annual meetings of the Society for Personality and Social Psychology and the Cognitive Neuroscience Society. The first jobs specifically advertising for positions with a focus on SCN were listed for Dartmouth in 2000 and Columbia in 2002. By 2006, postings for social psychology jobs with an SCN focus had become common. Undergraduate and graduate courses in SCN and related topics mushroomed, with growth in graduate programs offering training in SCN keeping pace. SCN's growth also has been apparent in the numerous special issues devoted to the topic that appeared in both psychology and neuroscience journals, including *Neuropsychologia*, *Journal of Personality and Social Psychology*, *Political Psychology*, *Journal of Cognitive Neuroscience*, *Neuroimage*, and *Cognitive Brain Research*.

But perhaps the most important developmental milestone for SCN has been the availability of funding from national agencies. In 1999 and 2000 the National Science Foundation (NSF) awarded Small Grants for Emerging Research (SGER, or "sugar" grants) to numerous researchers seeking to establish SCN research programs on numerous topics, ranging from stereotyping and person perception to self reflection and emotion regulation. These grants were awarded at the discretion of the director of behavioral science at NSF, Steven Breckler, who sought to provide seed money that would enable researchers to acquire pilot data for future grant applications. Then in 2001 the National Institute of Mental Health issued a Request for Applications (RFA) in social neuroscience, which provided the first opportunity for SCN researchers to apply for funding to a program specifically designed to meet the needs of interdisciplinary work linking social behavior to its neural bases. By 2005, other agencies, such as the National Institute on Drug Abuse and the National Institute on Aging, had issued similar RFAs.

The importance of funding for launching a new field cannot be underestimated. As an illustration of this point, consider the rapid growth of cognitive neuroscience in the early 1990s. Throughout the 1990s the private McDonnell-Pew Foundation provided substantial grants to researchers seeking to develop CN research programs, to train postdoctoral fellows, and to hold conferences. Other private foundations also provided money for small meetings that helped establish core groups of scientists whose work would exemplify the CN approach (M. S. Gazzaniga, personal communication). Although other factors played important roles, including a summer training institute at Dartmouth as well as an annual meeting and society, the availability of money, and the vote of confidence it implies, is essential for the development of any field. In this regard, SCN has been recognized as a distinct field by federal funding agencies. In the final section of this chapter, we revisit the topic of funding in the context of translational research that connects basic findings in SCN to clinical disorders characterized by socioemotional deficits.

PRINCIPLES GOVERNING THE PRACTICE OF SOCIAL COGNITIVE NEUROSCIENCE

In an American twist on the sentiments Ebbinghaus expressed earlier, Yale psychologist Neil Miller once quipped that "In most disciplines the scientists of today stand on the shoulders of the great scientists that have come before them. In Psychology, we step on their face" (S. Kosslyn, personal communication). The observation that psychologists don't just forget the past but actively may try to erase it may be an academic reflection of the market-driven mentality of Western culture. In the capitalist marketplace businesses compete for consumer dollars by marketing products as if no similar products ever had been offered before. Cars, cameras, and cookies are all the newest, most unique, and most satisfying. To the extent that product placements appeal to history, it is to emphasize that a particular product is the newest exemplar of a long line of products that always have been the most unique and most satisfying.

Miller is suggesting that psychologists are no different. Psychologists are essentially academic business-people hocking their theories and results in the marketplace of ideas. The consequence is that researchers often present their work in an historical vacuum that emphasizes their unique contributions at the expense of relating it to prior work.[4] Importantly, the tendency to ignore research peers is present not just at the level of the individual scientist but at the level of the scientific discipline as well. It is easy for researchers in either the traditionally "hard," biological and physical sciences or the traditionally "soft," social sciences, to believe that their cross-disciplinary colleagues can—and perhaps should—be ignored. Whether it is because their colleagues across the research fence are perceived to ask fundamentally different types of questions, to use fundamentally different kinds of methods, or to offer fundamentally uninteresting answers,

it is clear that many biological and social scientists do not buy what their colleagues have to offer.

This all-too-human tendency to value one's kin, one's comrades, and one's research culture over those of others is as much a danger for SCN as it is for any discipline. The costs of such disciplinary myopia could be especially acute, however, for SCN. Because the field requires incorporation of the methods and theories of different disciplines, researchers who "go it alone" risk making conceptual and methodological mistakes both naive and serious. As illustrated in this section, whereas in *principle* SCN work prioritizes interdisciplinary collaboration, integrative methodology, and multilevel theory, in *practice* this does not always turn out to be the case. This section highlights core principles that govern the psychology–neuroscience partnership that lies at the heart of SCN by illustrating the partnership with examples of problems that may arise if it is not honored. These examples are then used to distill four core principles for SCN.

Specifying the Goals of SCN Research

There are two types of goals that motivate SCN research in particular, and more generally, any research seeking to link psychological and neural levels of analysis (Ochsner & Lieberman, 2001; Sarter, Berntson, & Cacioppo, 1996). The first goal is sometimes referred to as brain mapping, which means conducting experiments that link the involvement of specific brain regions to specific types of behavior and experience. The emphasis here is on drawing *functional inferences* about brain systems by carefully manipulating task demands and observing corresponded changes in the recruitment of brain regions. Research conducted in this mode is necessary to draw *functional inference* about the processes associated with specific brain systems. Drawing *functional inferences* has been the primary goal guiding CN and AN research, which has produced detailed models of the neural systems involved in both high- and low-level visual cognition, implicit and explicit memory, visual mental imagery, fear conditioning, and numerous other phenomena (see Gazzaniga, 1995, 2000, 2004, for reviews).

These models of brain function provide the foundation for work guided by the second goal of SCN research, which is to use information about brain function to draw inferences about the psychological processes underlying a particular phenomena. The emphasis here is on drawing *psychological inferences* about the processes underlying a given behavior or experience by using the activation of particular brain systems as markers for the occurrence of particular kinds of psychological processes. The strength of these *psychological inferences* depends on the reliability with which particular functions can be ascribed to particular brain systems. For example, in the case of some psychological processes that have received a great deal of empirical attention, such as the encoding and retrieval of explicit memories, the reliability of these inferences is fairly strong: One can be reasonably certain that, for example, activation of the hippocampus reflects recruitment of a process used to encode configurations of activated perceptual inputs and stored representations that

together comprise an explicitly addressable memory for a life episode (Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998; Davachi, Mitchell, & Wagner, 2003; Schacter, Alpert, Savage, Rauch, & Albert, 1996). By contrast, the neural bases of many emotional and social cognitive processes of particular interest to SCN have just begun to be investigated. This means that one must be careful when drawing psychological inferences based on patterns of brain activity whose association with specific brain regions has yet to be solidified.

The reciprocal interplay of these two goals in guiding research design and inference is illustrated in Figure 3.2. In a sense, every experiment is serving both of these goals by (1) providing additional information about psychological processes that elicit activation in particular regions of the brain and (2) requiring that we place these results in the context of previous research to draw inferences about the psychological processes that observed activations represent. Decades of neuropsychological and neurological study combined with 15–20 years of functional imaging research has provided a reasonable set of methodological tools and theoretical models for understanding the function of brain systems implicated in memory, attention, language, and other "classically cognitive" processes (Gazzaniga, 1995, 2000, 2004). Work on core self-referential, social cognitive, and emotional processes builds on this foundation.

In Practice, How Are These Goals Achieved?

Whichever goal guides a specific experiment, one must be careful that (1) the experiment is designed to maximize the potential for drawing the strongest inferences possible about either brain function, psychological processes, or both; (2) one appropriately interprets the meaning of one's results in the context of other studies. As is argued below, the probability that these two constraints are satisfied may vary as a function of whether one approaches SCN research as an extension of one's existing field (i.e., adding neuroscience data to social psychology or social phenomena to cognitive neuroscience) or whether one treats SCN as a true interdisciplinary partnership that draws on the theories and methods of both fields simultaneously.

Potential Pitfalls of Market Expansion

The former approach is analogous to market expansion in business: Just as a corporation might decide to expand from supplying automobile engines to making the entire automobile, social psychologists and cognitive neuroscientists might decide to expand their domain of inquiry to include the neural or social levels of analysis. As discussed in the first section of this chapter, the development of SCN can be described as involving exactly this type of market expansion on the part of SCN's parent disciplines. However, for this expansion to be successful, one must acquire the expertise necessary to succeed in the new market. An engine maker turned automobile manufacturer would surely fail if it did not incorporate principles of ergonomics when designing a car interior, principles of materials science when selecting the rubber for tires, and so on. In the context of SCN, a failure to acquire either the necessary neuroscience or social psychological expertise can result in experiments that either draw improper functional inferences, that lack an organizing meta-theory, or both.

Psychological Inference

Infer operation of psychological process from patterns of brain activation

Brain System

Psychological Process

Functional Inference

Infer psychological function of brain system from covariance with behavior and experience

FIGURE 3.2. SCN seeks to draw two types of inferences concerning the relationship between psychological and brain processes. The first are functional inferences about patterns of experience and behavior to the operation of specific brain systems. The second are psychological inferences concerning the mental processes underlying a given experience behavior. In the case of functional imaging experiments, drawing these inferences depends on being able to reliably associate specific functions with specific brain systems based on prior research. For detail and explanation, see the section "Specifying the Goals of SCN Research."

THE FAULTS AND FOIBLES OF FUNCTIONAL INFERENCE

To illustrate one important danger of market expansion without acquiring proper expertise, consider the case of a social psychologist who wishes to use functional imaging to deepen her understanding of the way in which we make predictions about our emotional states. This kind of prediction is known as an *affective forecast*, and behavioral studies have suggested that there are numerous ways in which we overpredict the duration of our negative emotions and underestimate our ability to cope with adversity (Gilbert & Wilson, 2000). Let us imagine that this researcher would like to use functional magnetic resonance imaging (fMRI) to draw inferences about the psychological processes supporting affective forecasts about the near and far future. Her hypothesis is that judgments about the near future might evoke strong emotional responses whereas judgments about the far future might be made in a cold, abstract and propositional fashion (Eyal, Liberman, Trope, & Walther, 2004; Trope & Liberman,

2003). To test this hypothesis she records brain activity while participants make judgments about how they would feel about a hypothetical event (e.g., winning the lottery) if it took place in the near future (tomorrow) as compared to the far future (in a year). She knows that analysis of imaging data is similar to analysis of the behavioral data in that it involves contrasts of values on dependent measures in two conditions of interest. Thus, to identify brain regions more strongly associated with near forecasts, she subtracts activation on *far* trials from activation on *near* trials. To identify regions associated with far forecasts, she does the reverse, subtracting activation on *near* trials from activation on *far* trials. As illustrated in all four panels of Figure 3.3, the *near > far* contrast reveals greater activation in a region of medial prefrontal cortex (MPFC) identified by the crosshatched square.

What inference can she draw about the relationship of MPFC activation to her initial hypothesis? To address this question, she turns to the imaging literature on emotion and finds 48 different studies that show activation in MPFC when participants are perceiving, remembering, or experiencing emotion. As can be seen in Figure 3.3A–3.3D, the region identified in her *near > far* forecasts contrast lies in the middle of these emotion-related activations. She concludes that her initial hypothesis is supported—that near forecasts are more emotional than far forecasts—and proceeds to write up her results and

sends them to a well-known cognitive neuroscience journal for publication.

What might her reviewers say about this paper? Imagine that the paper receives three reviews. Reviewer A is an affective neuroscience researcher whose work has suggested that emotion activates MPFC (see Figure 3.3A). Reviewer A therefore writes a positive review suggesting that this is an innovative application of imaging research to address social psychological questions. Reviewer B, however, does not write a positive review. Reviewer B is a cognitive neuroscience researcher who has found that self-referential judgments activate MPFC (see Figure 3.3B), and Reviewer B suggests that the experiment has failed to rule out an alternate hypothesis: when making *near* forecasts participants might be more likely to think about the personal implications of their choices (e.g., If I win the lottery, will I get greedy? Will it change me?). Reviewer B reasons that engagement of self-reflective judgments recruits MPFC and could be the reason the social psychologist observed MPFC activation in her experiment. Reviewer C also writes a negative review. Reviewer C is a cognitive neuroscience researcher who has found that theory-of-mind tasks activate MPFC (see Figure 3.3C), and C suggests that the experiment has failed to rule out another alternate hypothesis: When making *near* forecasts individuals are more likely to think about the thoughts, beliefs, and feelings of other people (e.g., If I
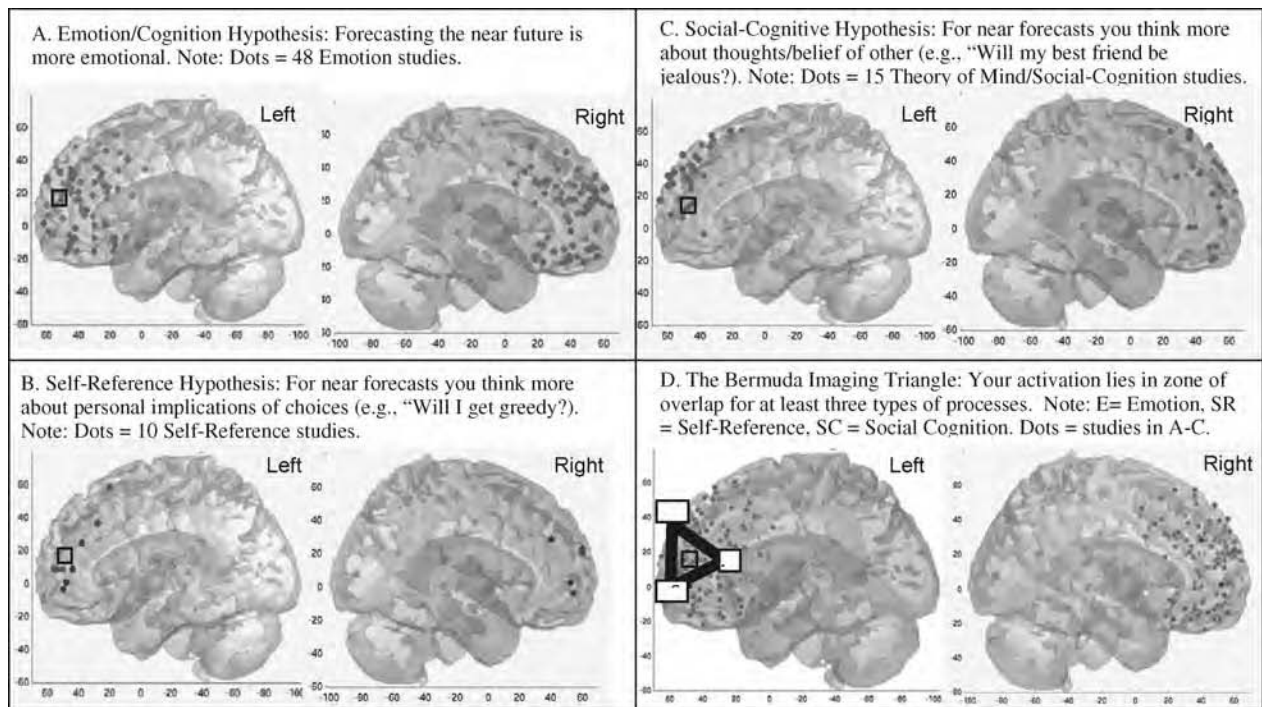
FIGURE 3.3. A diagrammatic illustration of the difficulty of inferring the operation of specific psychological processes given the one has observed activation in a specific brain region, such as medial prefrontal cortex (MPFC). Activation for a hypothetical affective forecasting task is illustrated by the crosshatched square superimposed over the left MPFC. Across panels A–C, MPFC activation is superimposed on images showing activation associated with emotion (A), self-referential judgments (B), and social cognitive attributions (C). For detail and explanation, see the section "The Faults and Foibles of Functional Inference" for details.

win the lottery, will my best friend be jealous? What will my wife think? Will my coworkers think I'm cool?). Reviewer C reasons that engagement of social-cognitive judgments recruit MPFC and could be the reason the social psychologist observed MPFC activation in her experiment.

At this point, the poor social psychologist might feel that she has been caught in what could be called the *Bermuda Imaging Triangle* (Figure 3.3D) from which her paper cannot escape. Because her focus of activation lies in a zone of overlap for studies that involve three (or more) putatively different types of psychological processes, it is difficult to draw post hoc inferences about what her pattern of activation means. As a consequence, the social psychologist could be perceived as naive by neuroscientists, her paper could be rejected, and she may feel confused and lose confidence that imaging can prove useful as a tool in her research program. An alternative, but in certain ways equally unfortunate, review outcome would be for her paper to receive only positive reviews from researchers that are favorable to her initial hypothesis (like Reviewer A). In this case, publication of her paper could be greeted by skepticism and disdain by neuroscience colleagues who study self-referential processing and social cognition who might perceive her as naive for thinking MPFC was only involved in emotion. Either way—in the review process or when the paper reaches the literature—evaluation of her work might suffer because she has not yet grasped the relevant neuroscience literature and thereby taken into account alternative explanations for the phenomenon question.

Given the multiple functions that could be associated with MPFC, what can the social psychologist infer about the meaning of MPFC activation in her study? And how can these inferences help her achieve a more positive publication result? At least three factors will determine the strength of her inferences. First and foremost, as is the case for all forms of psychological experiment, the design of her study is critical in determining what psychological inferences she can draw about the meaning of her observed brain activations. For example, she would be on firmer footing in making the claim that participants were more emotional for *near* than for *far* forecasts if her experiment had included a self-report or autonomic measure that indicated this was the case. In general, one must always include behavioral measures in experiments that index the psychological constructs in question. As described in the next section, many early imaging experiments on emotion failed to do this and as a consequence, are ambiguous with respect to why particular patterns of activation were observed.

Second, as highlighted by Figure 3.3, the extent to which converging evidence from other domains of research implicate clear functions for a given region will also constrain the inferences one might draw. In the case of MPFC, one might question whether the overlap of MPFC regions involved in emotion, self-reference, and theory of mind is more apparent than it is real. Cross-study comparisons of brain activation foci are clouded by various factors that could lead to apparently similar or dissimilar patterns of activation, including: individual dif-ferences in functional brain anatomy, differences in methodology and operationalization of psychological constructs, and the facts that different researchers spatially normalize and analyze their data in different ways. This makes it is difficult to say whether two studies with apparently similar locations of MPFC activation (or activation of any other brain region) truly are recruiting identical regions.

One way to address this question is by designing studies that include conditions that allow one to test alternative accounts of the processes underlying the phenomenon question. In the forecasting study, the social psychologist either could have included separate additional tasks that involve self-reference or theory of mind or could have included conditions in her forecasting task that vary the extent to which one is engaged in self-referential or theory-of-mind processing. If greater MPFC activation was found for judgments that elicit stronger emotion, are more self-referential, and involve judging others' mental states, then the social psychologist could infer that a computation common to all these judgments has been recruited. By contrast, if all three types of judgment recruited distinct regions of MPFC— and *near* forecasts are associated with just one of them— then the social psychologist could more clearly infer that one type of process was involved.

Third, leverage for drawing psychological inferences about patterns of brain activation can be gained by performing meta-analyses that identify patterns of association that are reliable across large numbers of studies. Meta-analyses for studies of emotion (Phan, Wager, Taylor, & Liberzon, 2002; Wager & Feldman Barrett, 2004; Wager, Phan, Liberzon, & Taylor, 2003) as well as various higher cognitive processes (e.g., working memory and attention switching (Wager, Jonides, & Reading, 2004; Wager & Smith, 2003) have begun to identify patterns of activation that can be associated reliably with specific types of stimuli (e.g., memories or images), emotional states (e.g., sadness vs. anger), and individual differences (e.g., gender). The promise of meta-analyses is that they may one day be able to provide probabilities that activations observed in any specific region reflect different types of psychological processes (T. Wager, personal communication, January 2005). MPFC, for example, may turn out to consist of multiple overlapping subregions and, depending on where one's observed activation falls, could be associated with self reflection with a high probability (0.7), emotion with a moderate probability (.4), and theory-of-mind attributions with a low probability (.1). Of course, meta-analyses are only as good as the studies they comprise, which should be designed to test specific hypotheses about the phenomenon in question. These hypotheses can be informed by prior results, by meta-analyses, or, as discussed in the following section, by metatheoretical perspectives that motivate experimental designs.

METATHEORETICAL MISSES, MISSTEPS, AND MISTAKES

The social psychologist of the preceding section was described as somewhat naive about neuroscience data and

theory concerning MPFC function. More broadly, she could be described as being naive about the metatheoretical perspective that guides cognitive neuroscience research. A metatheory describes the relationship between dependent and independent variables at a level of abstraction removed from the particulars of a given phenomenon, and can be said to coherently describe core elements of theories within a field.

The metatheory guiding cognitive neuroscience research was first described by the late vision scientist David Marr (1982). Marr described three levels of analysis at which a given behavior or experience could be explained. At the highest *computational* level one provides a precise description of the phenomena to be explained. Marr termed this the "computational level" because it provides a description of the computational output to be produced by mechanisms described at lower levels of analysis. At the middle *algorithmic* level one provides a description of the information-processing mechanisms that give rise to the phenomena described at the computational level. At the lowest *implementation* level one provides a description of the neural hardware that instantiates information processing mechanisms described at the algorithmic level. Marr believed that these levels were independent, positing the functionalist view that any computation could be produced by numerous algorithms, each of which could be implemented in any type of hardware. Cognitive neuroscience, however, treats the levels as interdependent and interacting, because it recognizes that (1) not all algorithms can be implemented in all hardware (see example of mathematical operations permitted by Roman vs. Arabic numerals from Kosslyn & Maljkovic, 1990), and (2) although there may be numerous ways in which a given computational output could be produced by different algorithms, cognitive neuroscience is concerned with the ones that the human brain actually implements. The goal of cognitive neuroscience research is to construct a *theory of a functional architecture* that describes a phenomenon at these three levels of analysis (Ochsner & Kosslyn, 1999). This metatheoretical perspective finds expression in theories of memory, mental imagery, language, attention, motor control, and various other phenomena (for reviews, see Gazzaniga, 1995, 2000, 2004).

Our social psychologist's lesson in market expansion is that knowledge about the brain systems that could be involved in the processes under investigation is necessary at the outset. This would allow her to formulate a priori hypotheses about the brain regions that could be involved in affective forecasting, which would comprise the social psychologist's first-pass theory of the functional architecture underlying forecasts (Ochsner & Kosslyn, 1999). This theory should include a description of the specific kinds of forecasts under investigation, the psychological processes that are engaged for each type of forecast, and the hypothetical neural substrates for each type of psychological process. The psychological processes and neural systems could include those implicated in emotion as well as other types of processes, such as self-reflection or theory of mind. This theory then could be used to design appropriate control conditions for her

experiment that could rule out alternative hypotheses about the mechanisms underlying forecasting. If these criteria are met, she will be in a much stronger position to infer that predicted patterns of brain activation (if observed) support her hypotheses about the processes underlying *near* as opposed to *far* affective forecasts.

Although our examples thus far have focused on what a social psychologist might not know about cognitive neuroscience research, it is no less important for cognitive neuroscientists interested in SCN to recognize their own potential for naiveté about the metatheory guiding social psychological experimentation. Indeed, because cognitive neuroscientists thus far comprise the largest number of scientists interested in SCN, it is perhaps more important that they become aware of the metatheoretical stance taken by many social psychologists. Although aspects of this metatheory have been described in many ways, a consensual account would include the interaction of two elements: (1) a person, with all of his or her dispositions, chronically accessible mental constructs, concerns, temperaments, moods, and so on; and (2) a situation, which may include other individuals and their thoughts and feelings, and various cues that set goals or are the triggers for specific behaviors, including prescriptions for socially normative behaviors, the desire to make a good impression, and so on. This *person × situation* interaction posits that a person's behavior at any given moment in time is a product of the interaction between who that person is and what behaviors are available, permissible, and possible in a given context (Ross & Nisbett, 1991). Critically, it implies that the meaning we ascribe to a stimulus is not inherent in the stimulus itself but, rather, is a flexible product of our interpretation or *construal* of its meaning according to our current goals, which in turn may be a function of the current context. Individuals often fail to realize that their judgments, impressions, emotions, memories, and experiences are the product of *construals*, which can explain the numerous self-serving and self-enhancing biases the people exhibit (Robins & Beer, 2001; Taylor, Kemeny, Reed, Bower, & Gruenewald, 2000). This view has been described as *naive realism*, a term that captures the implicit assumption that one's perceptions reflect a realistic picture of the world that accurately and directly conveys its true nature (Robinson, Keltner, Ward, & Ross, 1995; Ross & Ward, 1996). This metatheoretical perspective finds expression in theories of personality, attribution, self-awareness, stereotyping, impression formation, and various other phenomena (Higgins, 1997; Mendoza-Denton, Ayduk, Mischel, Shoda, & Testa, 2001; Mischel & Shoda, 1995; Mischel, Shoda, & Mendoza-Denton, 2002; Ross & Nisbett, 1991).

Lack of awareness of this metatheoretical perspective could lead a cognitive neuroscientist interested in SCN to miss the richness of socioemotional phenomena that is captured by the person × situation interaction. To illustrate this possibility, consider early cognitive neuroscience studies of emotion. Most of these studies examined either uninstructed perception, recognition/identification, or memory for stimuli with ostensible emotional value. The most commonly used stimuli were facial ex-

pressions of emotion, followed by photographs of real-world scenes that elicited different types of emotional responses, emotionally evocative film clips, emotion-laden autobiographical memories, and auditory (e.g., screams and music), tactile (e.g., feather or shock), olfactory (e.g., sour or sweet odors), or gustatory (e.g., chocolate) sensory stimuli that elicited pleasant or unpleasant feelings. Meta-analyses of these studies have revealed that emotion-related activations do not array themselves simply as a function of modality of input, type of task, or type of "basic" emotion involved (Feldman Barrett, Ochsner, & Gross, in press; Phan et al., 2002; Wager et al., 2003).

Why the inconsistent results? From the perspective of a social psychologist, these studies treated emotion as a stimulus property like shape, size, or color rather than a context-dependent *appraisal* process that interprets the emotional value the stimulus in the context of an individual's current goals, wants, and needs (Figure 3.4) (Ochsner & Feldman Barrett, 2001; Ochsner & Gross, 2004; Scherer, Schorr, & Johnstone, 2001). As a consequence, these studies fail to manipulate the way in which individuals construed, or *appraised*, the meaning of stimuli, which leaves participants free to appraise stimuli in numerous different ways. Classic studies in social psychology, including the seminal work of Schacter and Singer and Lazarus, have shown that how we cognitively appraise the meaning of an event may determine how we respond emotionally to it (Lazarus, 1991; Schachter & Singer, 1962), including the regulation of emotion via deliberate or spontaneous appraisal of emotional stimuli as neutral (Erber, 1996). In part, lack of control over appraisal seems to have been a holdover from animal models of emotion that were derived from studies employing stimuli with primary reinforcing properties, such as pleasant tastes or electric shocks. Some cognitive neuroscientists have suggested that human emotions are essentially responses to linked with differing degrees of complexity to reinforcers of this type (Rolls, 1999). While it is true that emotions are adaptive responses to situations of relevance to current goals, and that some may involve evolutionarily conserved responses to reinforcing

stimuli, from a social psychological perspective human emotions may be elicited by both stimulus–response (S-R) and schematic mental representations; many involve experiential, behavioral, and autonomic components; and may involve various types of cognitive processing (for examples, see Scherer et al., 2001). The failure to appreciate this perspective also led to the failure of many studies to provide independent behavioral confirmation that emotional responses were, in fact, elicited. In the absence of independent behavioral verification that emotions were elicited—in the form of self-reports of experience, indices of autonomic arousal, or evidence of emotional behaviors such as facial expression—it is difficult to determine why or why not activation in a given brain structure has occurred.

The theory motivating most of these studies was that stimuli with emotional properties should activate classically "limbic," lower-order emotion-processing structures such as the amygdala but should not activate classically higher-order cognitive structures such as prefrontal cortex. Although some studies were consistent with this simple hypothesis (e.g., Buchel, Morris, Dolan, & Friston, 1998; Fredrikson et al., 1998; Irwin et al., 1996; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; Lane et al., 1997; Morris, Ohman, & Dolan, 1999; Shin et al., 1997; Zald, Lee, Fluegel, & Pardo, 1998), numerous other studies were not. Some observed activation in prefrontal cortex but not the amygdala (e.g., Canli, Desmond, Zhao, Glover, & Gabrieli, 1998; Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Kesler-West et al., 2001; Mayberg et al., 1999; Phillips et al., 1998; Taylor et al., 1998), whereas others observed activation in both structures (e.g., Buchel, Dolan, Armony, & Friston, 1999; Crosson et al., 1999; Damasio et al., 2000; Phillips et al., 1997). From the emotion-as-stimulus-property perspective, these results were difficult to explain and seldom received specific commentary. By contrast, from the emotion-as-appraisal perspective, these results can be seen as the product of a relative reliance on top-down cognitive appraisal processes as compared to bottom-up stimulus-driven appraisal processes (see Figure 3.4 and Feldman Barrett

Social Psychology	+	Cognitive Neuroscience	=	Social Cognitive Neuroscience
• Emotion as appraisal process dependent on context, needs, and goals		• Emotion as perception of stimulus property like shape, size, color		• Emotion as interaction of bottom-up (stimulus-driven) and top-down (goal-driven) appraisals, influenced by context
• Examine when, why, how appraisals produce particular emotions		• Find brain bases of recognition, memory, or attention to		• Examine brain bases of interactions between bottom-up and top-down appraisal processes

FIGURE 3.4. Diagram illustrating the differing approaches to explaining (middle row) and studying (bottom row) emotion for social psychology, cognitive neuroscience, and social cognitive science. For detail and explanation, see the section "Metatheoretical Misses, Missteps, and Mistakes."

et al., in press). The cognitive neuroscientist's lesson in market expansion is that the activations observed in different brain structures is not just a function of stimulus content but also a function of context and construal as well. As described earlier, current SCN experiments are investigating the conditions in which specific types of top-down or bottom-up appraisals lead to different types of emotional responses (see also Figure 3.4).

The Metatheory of SCN: Four Core Principles

The metatheory of SCN can be described as a combination of the metatheories guiding cognitive neuroscience and social psychology. SCN aims to describe behavior at three levels of analysis (see Figure 3.1 and Ochsner & Lieberman, 2001). The first is described as the *social* level, which includes a description of the phenomena of interest in terms of the experience and behavior of a person in a given context, as she perceives and interacts with a social target. Importantly, that target could be the someone else, or it could be the perceiver herself, as she reflects on her own traits, states, and goals and attempts to understand and make use of them. The second is the *cognitive* level, which includes a description—in information-processing terms—of the psychological processes that give rise to the experience or behavior of interest. The use of the term "cognitive" here is a holdover from the parent disciplines of social cognition and cognitive neuroscience and is not meant to imply a specific type of processing mechanism that is *cognitive* as opposed to something else. Instead, "cognitive" is a placeholder term for the notion that various types of processes and representations comprise the functional mechanisms that give rise to social level phenomena. Whether we call those mechanisms *cognitive* or *affective* or *motivational* may differ depending on the social level phenomena we are trying to explain, The third is the *neural* level, which includes a description of the neural systems implicated in the psychological processes hypothesized to underlie the social level phenomenon. SCN research uncovers relationships between variables described at these three levels of analysis by conducting studies that provide information about the psychological processes associated with specific brain systems, or uses information about brain systems to inform theories of the psychological processes engaged in social behavior (see Figure 3.2).

At its core, this metatheoretical perspective rests on four principles (see Figure 3.5).

- *Converging evidence:* Data from multiple levels of analysis provides constraints on theory
- *Content:* Processing of and by people is special
- *Construal:* All stimuli are multiply construable
- *Context:* Individual (internal) and situational (external) factors influence how specific types of content are construed

FIGURE 3.5. A list of four principles that form the core of social cognitive neuroscience's meta theory. For detail and explanation, see section "The Metatheory of SCN: Four Core Principles."

Constraints and Convergence

The first principle is that multileveled theories of behavior must be constrained by data collected using multiple methods with variables described at the three levels of interest. The combination of multiple streams of data allows researchers to converge on theoretical explanations that are robust and flexible and are not tied to a single specific experimental methodology. This principle is part of the bedrock of cognitive neuroscience research (Ochsner & Kosslyn, 1999), which employs multiple types of neuroscience methods, including not just functional neuroimaging (which is the emphasis in this chapter) but analysis of behavioral deficits in neuropsychological populations, electrophysiology, and many other related techniques (for reviews, see Kosslyn, 1999). In an analogous fashion, this principle also is part of the bedrock of social psychological research that employs multiple types of behavioral methods to study phenomena of interest. These methods range from self-report to various *measures* of behavior that indirectly indicate operation of a particular psychological process (response times, recall or other memory measures, etc.) to psychophysiological methods indexing sympathetic and parasympathetic arousal, and more (for reviews, see Cacioppo, Tassinary, & Berntson, 2000; Gilbert, Fiske, & Lindzey, 1998). SCN incorporates all these methods (Cacioppo, Berntson, Taylor, & Schacter, 2002).

Content

The second principle is that SCN is concerned with special kinds of content, at all levels of analysis. At the social level the field is concerned with social content, which includes stimulus cues involved in interactions with people. These cues include nonverbal perceptual inputs such as faces, facial expressions, eyes, bodies, and biological motion but also include higher-level inputs such as spoken language, nonverbal behaviors, and other forms of social communication including the beliefs and attitudes expressed by others. As noted earlier, the person with whom we interact is in some cases ourself, as our own behaviors, beliefs, and feelings become the focus of self-awareness, perception, and judgment. At the cognitive level the information-processing mechanisms engaged when processing social cues may be tailored to processing social information. The mental representations supporting perception of social cues, and the processes engaged when making judgments about them, may be different than those involved in processing nonsocial stimuli such as inanimate objects. Furthermore, the processes engaged typically are *hot*, which means that they involve processes that interpret the *emotional significance* of social cues, that *motivate* us to perceive or judge ourselves and others in particular ways, and that *affectively* color our experience while doing so. Most every judgment—social or otherwise—carries with it an evaluative core (Osgood, 1976), which places motivational processes at the core of SCN. At the neural level, there may be specific systems supporting the representations and processes engaged when processing socioemotional con-

tents. A significant portion of current SCN work aims to understand whether the processes and neural systems engaged when processing social stimuli are similar to or different than those engaged when processing nonsocial stimuli (e.g., Mitchell, Heatherton, & Macrae, 2002; Mitchell, Macrae, & Banaji, 2004).

Construal

The third principle is that the meaning of any stimulus is a function of the way in which its meaning is construed. In other words, the impact of a stimulus is determined both by the bottom-up, stimulus-driven processing of its intrinsic perceptual properties (e.g., a pleasant or unpleasant smell or the depiction of a disgusting scene) as well as top-down, goal-driven processing that can control the way in which stimulus meaning is extracted (Feldman Barrett et al., in press). As described previously, the power of construal to determine the meaning of a stimulus is one of the bedrock elements of social psychology's metatheoretical perspective (Ross & Nisbett, 1991).

Context

The fourth principle derives from the fact that any given stimulus can be construed in many different ways, which has been termed "the multiple construal problem." Two kinds of context help determine the way in which perceivers solve the multiple construal problem by interpreting, for example, a stimulus person as a lawyer rather than a black man, or a comment as a joke rather than an insult. The first is an individual's *internal* context, which consists of his or her biological temperament, preexisting beliefs, attitudes, chronically accessible schemas, goals, memories, and biases (Kosslyn et al., 2002). The second is an individual's *external* context, which consists of various aspects of their current situational milieu. The external context can include situational cues that specify specific behavioral goals (e.g., to make a good impression or to determine if someone is deceiving you) and place constraints on normatively acceptable behavior (e.g., not cutting in line at the supermarket or not making lewd comments to a woman in a bar), by activating specific mental representations that comprise an individual's internal context (see Higgins, Chapter 19, this volume) for a detailed discussion of how person and situation factors influence activation of different types of knowledge, and in turn, construals).

A Metaperspective on SCN's Metatheory: Specificity versus Universality

If SCN's metatheory describes the kinds of theoretical accounts the field aims to develop, how specific should these theories be? Traditionally, a primary goal of much physical behavioral and biological science has been the derivation of species-general, universal laws that govern behavior irrespective of context, content, and construal. The search for such universal laws of behavior may represent an attempt to mimic the physical sciences, which aim to formulate "fundamental" laws of matter and energy that are widely applicable. The laws that govern behavioral and biological science may be fundamentally different, however (Kagan, 1998). Over the course of evolutionary time, natural selection exerts the greatest pressure for an organism to retain a heterogeneous set of mechanisms supporting behaviors appropriate for specific survival contexts that in turn depend on specific biological adaptations. The key is that behavioral and biological science fundamentally is about understanding the mechanisms that govern an interaction between an organism's biological endowment and its ecological environment (i.e., its *situation*). In keeping with this notion, SCN places relatively less emphasis on abstract universal principles and relatively greater emphasis on discovering contextually sensitive laws that govern human behavior. These laws specify underlying mechanisms at multiple levels of analysis. For example, following social cognitive models of dual processing, we might specify how and when automatic and controlled processes come into play as a function of situational factors such as motivational involvement and information-processing capacity, and following cognitive neuroscience models of control, we might specify the prefrontal dynamics underlying these processes.

THE FUTURE PROMISE OF SOCIAL COGNITIVE NEUROSCIENCE

This chapter began by drawing an analogy between personal and professional relationships on one hand and the nature of SCN on the other hand. From the personal perspective, social psychology and cognitive neuroscience can be seen as parent disciplines whose progeny has grown into an independent discipline. From the professional perspective, social psychology and cognitive neuroscience can be seen as business partners who need each other's expertise in order to expand their research market.

Which metaphor is correct? The answer is that metaphors are not mutually exclusive. Each one highlights different aspects of SCN that are important for understanding its development and its practice. Indeed, understanding the nature of any field presents its own multiple construal problem. In the case of SCN, the field can be construed as the offspring of successful parents with venerable research bloodlines, as an entrepreneurial startup with research potential, or a bit of both. The choice of construal may depend on the question one asks. If one wants to understand SCN's historical antecedents, the personal metaphor may be most useful. If one wants to understand what social psychologists and cognitive neuroscientists need to know about what the other discipline has to offer, the partnership perspective may be most useful.

Whichever metaphor guides one's construal of SCN, the question arises as to what the future may hold for the ongoing development of the field. Its rapid growth and the proliferation of conferences, publications, and funding opportunities suggests that SCN is not merely a flash in the pan. That being said, it is important to consider

what questions currently are central to SCN research and how current research may translate into promising directions for future work.

Predicting the Future: Three Examples from Current Research

An old aphorism states that past behavior is the best predictor of future behavior. This suggests that a good way of predicting what the future may hold for SCN is to consider the state of its current understanding of a few key questions. In this section, we illustrate the SCN approach with examples of current research that highlight the difference between approaching topics from an integrative SCN perspective as opposed to approaching them from a social psychological or cognitive neuroscience perspective alone.

Person Perception

"Person perception" is an umbrella term referring to the various ways in which we first perceive social cues, judge social targets, and subsequently form impressions of social actors. SCN research on person perception has followed two major threads. The first has been heavily influenced by the cognitive neuroscience emphasis on bottom-up, stimulus-driven processing of perceptual cues. Research following this thread has examined the neural systems responsive to face as compared to nonface objects, facial expressions of various kinds, facial features such as the direction of eye gaze, bodies as compared to other objects, and biological as compared to nonbiological motion (Allison et al., 2000; Puce & Perrett, 2003). Much of this work is concerned with the question of whether or not social cues enjoy privileged status in the brain. For example, one hotly contested debate asks whether faces possess unique features and are processed by a dedicated cortical module (the fusiform face area, or FFA) or whether faces are one example of a stimulus for which we have gained great expertise and are processed by cortical regions tuned to support recognition of any stimulus for which we are recognition experts (Gauthier, Curran, Curby, & Collins, 2003; Gauthier & Tarr, 1997; Gauthier et al., 2000; Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997; Kanwisher, Stanley, & Harris, 1999; Yovel & Kanwisher, 2004). It is noteworthy that traditional social psychological approaches to person perception take it for granted that social cues are recognized accurately and concern themselves with subsequent stages of judgment and impression formation. By contrast, SCN models unpack this initial step into a suite of neural systems dedicated for processing different types of social cues. In so doing, SCN models deepen the understanding of person perception offered by social psychological models that are concerned less with the question of how nonverbal cues are recognized.

The second thread in SCN research on person perception concerns the way in which we infer or understand the intentions of other individuals. This research thread also has been heavily influenced by the cognitive neuro-

science emphasis on stimulus-driven processing as exemplified by research on "mirror neurons" (Rizzolatti & Craighero, 2004), Mirror neurons were first described in primate studies of motor control. Some prefrontal cortical neurons would fire when a monkey would execute a hand motion and also would fire when a monkey would observe a different actor's hand executing the same action. Subsequent studies suggested that these neurons were sensitive to the goal of an action and not just its superficial gestural characteristics (Rizzolatti & Craighero, 2004). Human neuroimaging research picked up this thread by demonstrating common regions of primary motor, left inferior prefrontal, and parietal cortex that seem to be similarly responsive to the perception and execution of a motor action (Buccino, Lui, et al., 2004; Buccino, Vogt, et al., 2004; Decety, Chaminade, Grezes, & Meltzoff, 2002; Hari et al., 1998; Iacoboni et al., 1999; Ruby & Decety, 2001). Other studies have suggested that common neural systems may be involved when individuals perceive and pose facial expressions of emotion (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003) or both personally experience and watch others experiencing pain or disgust (Calder, Keane, Manes, Antoun, & Young, 2000; Jackson, Meltzoff, & Decety, 2005; Morrison, Lloyd, di Pellegrino, & Roberts, 2004; Singer et al., 2004; Wicker et al., 2003). These findings have led some to propose that "shared representations" underlying the perception and execution/experience of an action form the foundation of our ability to understand others' mental states (Gallese, Keysers, & Rizzolatti, 2004; Meltzoff & Decety, 2003). Thus, shared representations presumably provide an automatic internal simulation of what it would be like for the perceiver to perform an observed action and this simulation, or perception–action "resonance" enables the perceiver to understand what someone is doing and why.

There are three significant problems with this account. The first is that, to date, the activation of motor mirror neurons has never been shown to predict the ability to understand actions—whether social or nonsocial—of another individual in either a laboratory or a real-world context. Studies involving the experience and perception of pain have shown that common regions of activation covary with individual differences in empathic ability, but they have not yet demonstrated that the tendency to coactivate a given region in both perceptual and experiential contexts predicts the ability to accurately judge what another person is feeling or thinking. If the mirror neuron account was correct, such evidence should be found. The second problem is that the shared representations that have been identified to date are informationally sparse and are unlikely to support the inferences necessary to understand another person's intentions. Both developmental and social psychological models of intentional inference suggest that individuals draw on both semantic and episodic memories to help guide judgments about what a given individual is thinking or feeling in a given context, as well as higher-order capacities for reasoning and judgment. The shared motor and affect representations identified therefore are unlikely to contain this information. They might support vicarious

learning about the causes and consequences of particular actions, but they likely do not support understanding that a person is trying to deceive us, that he or she might have incorrect beliefs about what we believe, and so on. The third, and perhaps most significant, problem is that mirror neurons/shared representations theories of person perception really are not theories at all. They are descriptions of data and provide neither precise social-level descriptions of the full range of phenomena to be explained nor information-processing descriptions of the steps necessary to achieve specific kinds of interpersonal understanding. By contrast, social psychological models of person perception provide both kinds of descriptions, which enables them to explain the situational and motivational factors that determine when and how we are likely to imitate others (Chartrand & Bargh, 1999; van Baaren, Horgan, Chartrand, & Dijkmans, 2004; van Baaren, Maddux, Chartrand, de Bouter, & van Knippenberg, 2003) and more generally what determines whether we are accurate or inaccurate empathizers (Hodges & Wegner, 1997; Ickes, 1997) or social judges (Gilbert, 1998; Krueger, 2003). Our errors of person perception are particularly revealing because they lay bare the egocentric biases, heuristics, and implicit theories that guide our judgments of others (Epley, Savitsky, & Gilovich, 2002; Saxe, 2005; Wilson & Brekke, 1994). Mirror neuron accounts of social cognition do not speak to these errors.

Current SCN work is moving beyond simple shared-representation models of person perception toward linking the systems used to encode perceptual cues to those important for the high-level processes and memory representations used to draw social inferences (Gallagher & Frith, 2003; Lieberman, Gaunt, Gilbert, & Trope, 2002; Mason, Banfield, & Macrae, 2004; Ochsner et al., 2005) and is just beginning to understand how these representations may bias the person perception process (Mitchell, Macrae, & Banaji, 2005). Two of the strengths of this work are its strong foundation in developmental psychological models of intentional inference and its breakdown in autism (Frith, 2001; Frith & Frith, 2003; Saxe et al., 2004) and its potential to help fractionate the person perception process into component parts by identifying different common and distinct patterns of brain activation associated with different types of judgment (Kosslyn, 1999; Ochsner & Lieberman, 2001). Thus, imaging experiments may be able to identify systems involved in the dynamic interplay of bottom-up and top-down processes during person perception, as well as systems involved in biased as compared to accurate judgments.

Self-Perception

Another major theme of current SCN research concerns the neural correlates of self-perception. As was the case for research on person perception, work on self-perception has transitioned from being strongly influenced by the CN emphasis on perception and recognition to a broader emphasis influenced by social psychological models of self-knowledge.

An initial attempt to organize neuroscience research on self-perception suggested that right-hemisphere advantages for various self-related tasks supports the theory that the right hemisphere plays a special role in self-recognition (Keenan, Nelson, O'Connor, & Pascual-Leone, 2001). These tasks included recognizing photographs of the self as compared to others (Keenan, Freund, Hamilton, Ganis, & Pascual-Leone, 2000; Keenan et al., 1999), retrieving autobiographical memories (Tulving, Kapur, Craik, Moscovitch, & Houle, 1994), and maintaining a coherent map of the body and its location in space, which is commonly disrupted by right parietal lesions that produce deficits of body and spatial awareness (Ramachandran, 1995). The major problem with this view is that it lacks a coherent theory or metatheory to explain what it is that links all these behaviors together. In what sense is viewing a photograph of one's self as compared to a stranger a core social cognitive process that is critical for social functioning? In what sense does this have anything to do with perceiving one's personality attributes and one's qualities? How does retrieving an autobiographical memory or being unaware that one has a neuropsychological deficit involve processes similar to or different than those involved in recognizing yourself in a photo? What criteria determine whether a task does or does not involve these processes? These questions have not been addressed by this account.

A second line of research on self-perception has used functional imaging to study the neural correlates of the self-reference effect in memory (Symons & Johnson, 1997), which refers to an advantage in memory for trait words (e.g., friendly) encoded by judging how well each trait describes oneself as compared to encoding them by judging some other semantic or nonsemantic attribute (e.g., number of syllables). Linking words to the complex organizational structure of self-knowledge is thought to enhance memory to levels difficult to match without some other sufficiently organized and elaborate method of study (Symons & Johnson, 1997). Initial neuroimaging studies of this effect suggested a special role for MPFC in judging the relevance of trait words to the self as compared to judging their relevance to famous but not personally known individuals, such as President George W. Bush (Craik et al., 1999; Fossati et al., 2003; Kelley et al., 2002; Macrae, Moran, Heatherton, Banfield, & Kelley, 2004). Subsequent studies have suggested that MPFC activation may be elicited by various other kinds of self-referential judgments as well, including assessing one's emotional state (Ochsner, Knierim, et al., 2004), preferences (Zysset, Huber, Ferstl, & von Cramon, 2002; Zysset, Huber, Samson, Ferstl, & von Cramon, 2003), abilities (Johnson et al., 2002), and attitudes (Cunningham & Johnson, 2007). Furthermore, some studies have suggested that similar regions of MPFC are involved in judging one's own feelings or attributes and the feelings and attributes of others (Ochsner et al., 2005; Ochsner, Knierim, et al., 2004; Schmitz, Kawahara-Baccus, & Johnson, 2004). These findings suggest that the process of reflecting on what others think about us is very similar to

the process of thinking about ourselves directly, which bears on theories of the origin and nature of self-knowledge.

The fact that MPFC has been implicated in various types of self-perception, as well as mental state attribution more generally, is striking for least two reasons. The first is that it suggests that seemingly disparate social phenomena may share common underlying information processing and neural mechanisms. These data could not have been obtained using behavioral studies alone. This is important because social psychologists typically study some of these phenomena—such as attitudes and judgments of personality—independently from one another. The fact that they may share common neural mechanisms sheds light on their potential similarities in terms of common psychological processes that depend on those brain systems. The second reason was highlighted earlier, in the section "Principles Governing the Practice of SCN," which considered the difficulty of drawing specific inferences about the nature of these neural mechanisms given the fact that MPFC has been implicated in multiple different behaviors. An important direction for future research will be performing within-study comparisons contrasting different types of self-referential judgment as well as different types of judgments about the mental state of others. The goal is to determine the functional organization of MPFC, which could help clarify the similarities and dissimilarities on the different types of judgment associated with MPFC activation.

In this regard, it may be important to distinguish theoretically between different senses of self (cf. Gillihan & Farah, 2005), a topic that has been the focus of much social psychological research (Baumeister, 1998). One distinction that might be useful contrasts the first-person sense of ownership over one's actions and perceptions and the third-person sense of being the object of one's introspection. William James referred to these two senses of self as the *I* and the *Me* (James, 1890). The experience of the *I* is immediate and direct and accompanies both the stream of conscious sense of experiencing one's perceptions as well as the agentic sense of controlling one's behaviors. The experience of the *me* is metacognitive and indirect and accompanies the sense of reflecting on one's attributes, abilities, states, and body. It has been hypothesized that ventral portions of MPFC may be more strongly associated with the *I*, whereas dorsal portions of MPFC may be more strongly associated with the *me* (Ochsner & Gross, 2005). It remains for future work to test this hypothesis (for other potentially useful distinctions related to types of self-knowledge, see also (Higgins, 1996b; Lieberman, Jarcho, & Satpute, 2004).

Self-Control

The ability to control the content of one's thoughts, the nature of one's feelings, and the expression of one's actions is commonly referred to as self-regulation. In contrast to topics such as person perception and self-perception, this topic has seen perhaps the greatest amount of independent research from each of SCN's parent disciplines. Within CN, self-regulation first was studied under the rubric of executive function, which referred to a variety of abilities impaired by frontal lobe function, such as planning and problem solving, that were measured by neuropsychological task batteries. Global concepts of a central executive generated by this work have given way to focused models of specific forms of cognitive control, including working memory, selective attention, and response selection, all of which can be measured by sensitive speeded response-time tasks. Within self-perception, "self-regulation" has been an umbrella term referring to any number of different instances in which an individual needs to curb an impulse (e.g., to diet), alter an emotion or mood, or salve a blow to one's self-esteem or social relationships.

The key is that within both disciplines, the need to inhibit or transform prepotent responses has been a major focus of research. The long history of self-regulation research in both CN and self-perception has provided a firm foundation for interdisciplinary bridges to be built in the form of collaborative SCN research. This research has taken the form of using CN models of cognitive control and emotion to help elucidate the dynamics underlying various forms of control over person perception, affect, and emotion, traditionally of interest to social psychologists (Ochsner, in press; Ochsner & Gross, 2005).

For example, a number of studies have manipulated the level of attention paid to briefly presented faces that express emotion or are exemplars of racial ingroups or outgroups. These studies have tested the hypotheses that (1) emotionally relevant social stimuli should activate the amygdala, and (2) following the logic of behavioral experiments these responses could be considered automatic to the extent that they do not vary as a function of attention (Öhman, Flykt, & Lundqvist, 2000). In general, results have been mixed. Some studies have found results consistent with these hypotheses (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Öhman, 2002; Vuilleumier et al., 2002; Vuilleumier, Armony, Driver, & Dolan, 2001), whereas others have found that amygdala responses diminish as a function of attentional load (Pessoa, Kastner, & Ungerleider, 2002; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002) or exposure (Hart et al., 2000).

Although there may be numerous reasons for these discrepant results, one possibility is salient in light of earlier discussion. Much CN research on emotion has treated it as a stimulus property, has lacked a meta-theoretical perspective, and as a consequence has failed to measure or manipulate the way in which participants appraise the meaning of stimuli with affective relevance. Thus, it is possible that qualitatively different processes—including those involved in cognitive control—become engaged when participants devote their full attention to the encoding of emotionally evocative stimuli as compared to when few attentional resources are available.

This possibility is supported by research suggesting that when attentional resources are available and directed toward individuating (Wheeler & Fiske, 2005) or thinking verbally (Hariri, Bookheimer, & Mazziotta,

2000; Hariri, Mattay, Tessitore, Fera, & Weinberger, 2003; Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005) about socioemotional stimuli, that amygdala responses may be reduced. In part, this may be due to heightened awareness of ambivalent feelings about target stimuli possible, which participants regulate by engaging control processes. Thus, faces that can be perceived as expressing either surprise or fear activate the amygdala when judged to express fear and ventral MPFC when judged to express surprise (Kim, Somerville, Johnstone, Alexander, & Whalen, 2003). Similarly, attitude targets that elicit both positive and negative evaluations—such as abortion or a black face for a white participant who professes no explicit prejudice but exhibits prejudice one implicit tasks—tend to activate dorsal anterior cingulate and right lateral prefrontal regions implicated in cognitive control (Amodio, Harmon-Jones, & Devine, 2003; Cunningham, Raye, & Johnson, 2004). It remains for future research, however, to identify the specific patterns of social cognitive appraisal that lead stimuli to be perceived in neutral as compared to affectively arousing terms.

Studies examining the use of cognition to regulate the experience of emotion have identified at least two different types of *re*appraisal that engage prefrontal and cingulate control systems to downregulate emotional appraisal systems such as the amygdala. One strategy involves becoming psychologically distant and detached while observing an emotionally charged photo or film, which has been shown in behavioral experiments to be effective for downregulating negative emotion (Gross, 1998). Imaging studies have shown that this strategy engages prefrontal systems to regulate activation of appraisal systems related to sadness (Levesque et al., 2003, 2004), sexual arousal (Beauregard, Levesque, & Bourgouin, 2001), or negative emotion more generally (Ochsner, Ray, et al., 2004; Phan et al., 2005). A second strategy involves "looking on the bright side," or, "finding the silver lining," in an aversive event by reframing its meaning in terms that neutralize or even positivize its emotional punch (Lazarus & Alfert, 1964). For example, one could imagine that a sick man depicted in a photograph has a hearty constitution, feels little pain, and soon will be well. This strategy also activates prefrontal and cingulate control systems to downregulate amygdala responses to aversive images (Ochsner, Bunge, Gross, & Gabrieli, 2002; Ochsner, Ray, et al., 2004).

The specific locations of control and appraisal-related activations have varied considerably across studies, however, and it will be necessary for future research to provide within-study comparisons of strategy and stimulus type to determine how and why specific kinds of regulation are associated with specific neural dynamics. One study has directly compared the *self-focused* and *situation-focused* strategies described earlier, and found that they differentially depend on medial and lateral prefrontal cortex, respectively (Ochsner, Ray, et al., 2004).

It also will be essential that future work examine the numerous other forms of self-regulation typically studied by social psychologists, which range from the control of eating to the recovery of self-esteem after a failure or rejection (Ayduk, Mischel, & Downey, 2002; Heatherton, Polivy, Herman, & Baumeister, 1993; Steele, Spencer, & Lynch, 1993). Such work could help determine whether regulatory responses to social threat depend on psychological and neural processes similar to or different than those used to regulate responses to physical threats, such as pain (Eisenberger & Lieberman, 2004), or responses to any emotion-eliciting discrepancy between desired and actual outcomes (Higgins, 1996a, 1999). The goal here is to identify and understand the dynamics of core systems important for control on the one hand and emotional appraisal on the other. As discussed below, such models could have important implications for the development and dysfunction of social and emotional abilities (Ochsner & Gross, 2005).

Mapping the Road Toward the Future

All forms of research are journeys of discovery. With any luck, the journey follows a road that does not endlessly wind back on itself but moves forward toward its ultimate goal. In the case of SCN, that goal involves building multilevel models of socioemotional phenomena. Thus far, this chapter has been concerned with building normative models that describe the behavior of physically and psychologically healthy adults. But this is only a first step. Once a normative adult model has been established, it can and should be extended in numerous directions. As considered in this section, models can be applied toward understanding new domains of research and can be extended to additional levels of analysis.

When contemplating any journey it is often useful to have a map. To guide our discussion of new directions for future SCN research, Figure 3.6 maps the relationships between SCN and a number of allied disciplines, all of which are concerned with understanding the relationships between psychological processes, neural systems, and/or clinical outcomes. Construing the term "social neuroscience" broadly to refer to any research that links social-level variables to biological variables (Cacioppo, 2002; Cacioppo et al., 2002), this map charts the relationships among various disciplines within *social neuroscience*. The point of this map is to help visualize points of potential connection between SCN and other disciplines that could be the focus of collaborative work.

New Domains of Research

SCN is more of an interdisciplinary approach to asking and answering questions about (both intra- and interpersonal) social phenomena than it is a field with crisply defined topical boundaries. Understood this way, the SCN approach could be used to profitably address questions of interest to a number of closely related sister disciplines, some of which are depicted in the lower left of Figure 3.6.

For example, affective neuroscience, which as described earlier has been concerned with mapping the brain correlates of a basic set of emotional responses, potentially could benefit from SCN's emphasis on construal and context. For SCN, emotional experiences are many
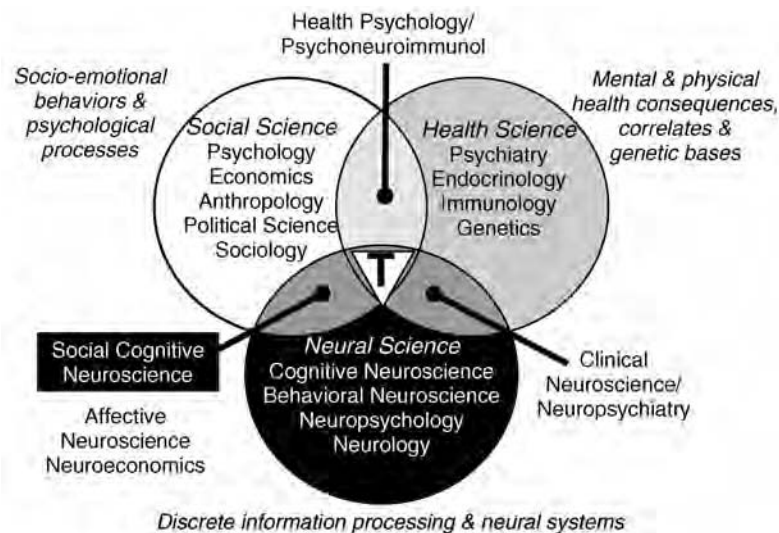
FIGURE 3.6. Diagrammatic map of the relationship between SCN and allied disciplines. The three large circles represent traditional disciplines in the social sciences, health sciences, and neural sciences. The concerns typically associated with each of these three major disciplinary categories are listed in the periphery. Interdisciplinary fields such as social cognitive neuroscience lie at the intersection of traditional disciplinary boundaries. Research that incorporates aspects of all three disciplinary categories would lie in the center of the figure, which is represented by the "translational triangle" designated with a capital *T*. Such *translational research* uses models of socioemotional behavior derived from basic science research with normative populations to address questions about the mechanisms underlying maladaptive mental and physical health outcomes. The map is intended to visualize the potential points of connection between SCN and other disciplines that could be the focus of collaborative work. For detail and explanation, see the section "Mapping the Road Toward the Future."

and varied and the product of cognitive construals or appraisals that may depend on a variety of neural systems. An important avenue for future research could be unpacking the ways in which different external situational and internal (states, goals, dispositions, etc.) contexts determine what type of emotional responses are generated (Feldman Barrett et al., in press; Ochsner & Feldman Barrett, 2001).

In like fashion, an SCN approach could benefit neuroscience research that recently has begun to emphasize the study of personality and individual differences (Hamann & Canli, 2004; Kosslyn et al., 2002). Although this research is explicitly focused on understanding the way in which internal contexts impact psychological and neural processes, it may benefit from SCN's additional emphasis on the way in which experience and learning impact the construal process as well. For example, apparent gender differences in neural responses to emotional stimuli could arise from ways in which cultural learning shapes the way men and women construe the meaning of emotional experiences (Wager & Ochsner, in press). These differences in construal may, of course, have neural correlates, but their origin may be attributable to the tuning of neural circuits via culture rather than innately specified genetic, endocrine, or other biological factors.

An SCN approach also could be used to help understand the development of and change in various social and emotional abilities across the lifespan. Cognitive neuroscience research has identified different developmental trajectories for systems related to cognition and emotion such that emotional appraisal systems such as

the amygdala reach adult size earlier in life than do prefrontal control systems, which undergo a rapid growth spurt between the ages of 8 and 12 and continue to structurally develop into one's late 20s (Diamond, 2002; Giedd, 2004; Luna & Sweeney, 2001; Luna et al., 2001). This differential sensitivity to aging continues into older adulthood, as age-related degeneration of the amygdala (and other structures related to emotion, like orbitofrontal cortex) is slow compared to degeneration observed in lateral prefrontal and cingulate systems related to cognitive control (DeCarli et al., 1994; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998; Salat, Kaye, & Janowsky, 2001). Behavioral data suggest that working memory and attentional capacities wax and wane with the growth and degeneration of prefrontal and cingulate control systems (Grady, 2002; Klingberg, Forssberg, & Westerberg, 2002; Milham et al., 2002; Reuter-Lorenz et al., 2000; Thomas et al., 1999), but it is not yet known if social and emotional functions wax and wane with structural changes in a similar fashion. The implications of developmental trends for person perception, self-perception, and emotional self-regulatory abilities could be an important topic for future SCN research—as would the potential impact of cultural learning history and life-stage transitions on the development of these abilities as well (Higgins & Eccles-Parsons, 1983).

The emerging field of neuroeconomics might similarly benefit from an SCN approach. Neuroeconomics is concerned with understanding the neural correlates of social exchanges, decisions, judgments, and predictions that have varying degrees of utility, or value, to a person. SCN

could be instrumental in unpacking the concept of utility in terms of both the neural systems associated with subjective utility and systems associated with computing the objective value of a commodity. SCN models of self-perception and self-regulation could also inform the way in which decision makers regulate their affective responses to choice options and decision outcomes. For example, depending on whether one anticipates a positive or negative choice outcome (e.g., winning a bet), cognitive control processes could generate either anticipatory eagerness or regret. Depending on which outcome is actually experienced, control processes could help regulate disappointment, sadness, regret, or other negative emotions (Larsen, McGraw, Mellers, & Cacioppo, 2004). As described earlier, current SCN research is examining the neural systems implicated in emotion regulation, and the findings of these studies could be relevant here. Various studies of judgment and decision making activate prefrontal systems such as those used for emotion regulation (e.g., Bechara, Tranel, & Damasio, 2000; Rogers et al., 2004; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), but the relationship between the two is not yet clear. The key is that SCN incorporates social psychology's emphasis on context and construal and provides a direct link to a large body of social psychological research examining the way in which these factors influence judgment and choice.

Finally, an SCN approach also could inform other hybrid disciplines whose emergence may be on the horizon. For example, a recent special issue of the journal *Political Psychology* was devoted to describing the possibility that *political neuroscience* research could use neuroscience methods to study phenomena typically of interest to political scientists. An SCN approach to this endeavor could help specify the ways in which content, context, and construal are related to the neural systems underlying political attitudes, political decision making, and related phenomena (Lieberman, Schreiber, & Ochsner, 2003).

Translational Research

As normative SCN models of person perception, self-perception, self-regulation, and other abilities are solidified, they can be extended to help explain how their underlying mechanisms contribute to the maintenance of mental and physical well-being. Such *translational research* seeks to apply the methods and findings of basic science research to understanding the causes and consequences of both psychological and physical ailments. Translational SCN research would connect research located in the lower left of Figure 3.6 with traditionally biomedical and psychiatric research located in the upper right of Figure 3.6, thus occupying the intersection of the three primary domains of research depicted in this figure. This intersection zone is represented by the "translational triangle" in the center of the figure, designated with a capital *T*.

A prime candidate for translational SCN work might be unpacking the functional consequences of structural and functional abnormalities in emotional appraisal and cognitive control systems that have been identified in vir-

tually every major mood, anxiety, and thought disorder, including depression, anxiety, posttraumatic stress disorder, obsessive–compulsive disorder, and schizophrenia (Bremner, Vythilingam, Vermetten, Vaccarino, & Charney, 2004; Davidson, Pizzagalli, Nitschke, & Putnam, 2002; Heckers et al., 2004; Mayberg, 2003; Quintana, Wong, Ortiz-Portillo, Marder, & Mazziotta, 2004; Rauch, Savage, Alpert, Fischman, & Jenike, 1997; Tillfors et al., 2001). Despite the fact that many of the distressing symptoms accompanying these disorders are often social or emotional in nature, the majority of extant functional studies have measured brain activity while participants are either "at rest," in the scanner or performing a cognitive task. A clear avenue for future SCN research will be to translate its basic models of normative functioning to clarifying why and how clinical populations are anhedonic, asocial, highly anxious, or depressed or experience other forms of affective and social dysregulation.

Equally important will be translating SCN models of self-perception and self-regulation into an understanding of how they relate to physical health outcomes. Health psychological research has identified relationships between social variables—such as the experience of shame or loneliness or the size of one's social support network—to the occurrence of common colds, cardiovascular disease, and level of immune functioning more generally (Bandura, 2004; Cacioppo et al., 1998; Dickerson & Kemeny, 2004; Kemeny, 2003; Kiecolt-Glaser, Cacioppo, Malarkey, & Glaser, 1992; Taylor et al., 2000; Uchino, Cacioppo, Malarkey, Glaser, & Kiecolt-Glaser, 1995). Current models of these effects either connect social level descriptions of interpersonal and regulatory behavior to other social-level descriptions of physical symptoms or bridge many levels of analysis between high-level descriptions of social variables and very low-level descriptions of molecular markers. These models do not yet make clear, however, how social, cognitive, and neural-level variables interrelate to produce adaptive or maladaptive health outcomes. SCN can play an invaluable role in filling in the missing levels of analysis, linking social variables to psychological processes, psychological processes to neural systems, and neural systems to transmitters, hormones, and endocrine systems.

Learning the Language

This chapter began with an everyday real-world scenario—waiting in the supermarket checkout line—that provides a glimpse into the metaphors by which we live. As the tabloid headlines reveal, human endeavors are easily described and understood in terms of personal or professional relationships. Much of this chapter used these two metaphors as the starting point for describing the historical development of SCN and the core principles that govern its experimental practice.

The supermarket checkout line reveals another important facet of human psychology, however, namely, the importance of communicating in a common language. Take a moment to think about the supermarket checkout line. In doing so, the reader may have spontaneously gen-

erated visual mental images of the scene. It is quite likely that the magazine headlines viewed in the mind's eye were printed in one's native tongue. Now imagine instead that the magazine covers are printed in an unknown foreign language. In this case, although one may be able to glean from the glossy cover photos that the stories inside are about people, the essential meaning of the stories is lost.

For an interdisciplinary field such as SCN, the importance of communicating in a common language cannot be underestimated. As described earlier in this chapter, there is always a danger that researchers in one domain will not be interested in the research products of their colleagues in another domain. This disinterest stems in no small part from an inability to understand cross-disciplinary jargon. Social psychologists may not understand the language of cortical and subcortical systems or fMRI scanners, and neuroscientists may not understand the language of attitudes, attribution, and person × situation interactions. Luckily for SCN, most newcomers to this field already have in common the cognitive, information-processing component of SCN's multilevel research language. Language pitched at this level provides descriptions of the psychological processes that link social phenomena and neural systems. This language, with some variation, includes terminology and concepts such as automatic and controlled processing, storage and retrieval, and selective attention that are part of social psychology and cognitive neuroscience as well as behavioral economics, psychiatry, and related disciplines.

A simple analogy can be used to illustrate the importance of the information-processing metaphor for understanding the mechanisms of socioemotional behaviors and unlocking the functions of neural systems. In the early 1800s, scholars in many nations were working to understand the meaning of Egyptian hieroglyphics. Their work focused on a stone slab discovered in the town of Rashid (*Rosetta*) in 1799 by Napoleon's invading army. Known as the *Rosetta Stone*, this slab contained a text written in three different languages. As depicted in the left panel of Figure 3.7, the top of the stone was written in hieroglyphics, a 3,000-year-old pictographic language. The middle and bottom portions were written in Demotic and Greek, which were the languages of literate Egyptians (of the time) and the government, respectively. In 1822, French scholar Jean-Francois Champollion deduced that repeating combinations of characters in each of the three texts referred to the royal name Ptolemy, and from there, he was able to link Greek and Demotic characters spelling the name to their hieroglyphics counterparts (Andrews, 1985).

The problem of translating hieroglyphics into Greek is not unlike the problem of translating the language of neuroscience into social psychological terminology—or vice versa. For the decoders of the Rosetta Stone, Demotic, the everyday language of literate Egyptians, was the link between modern Greek and ancient Hieroglyphics. For social cognitive neuroscientists, the everyday language of information processing is the link between neuroscience and social psychology, as illustrated by the right panel of Figure 3.7.
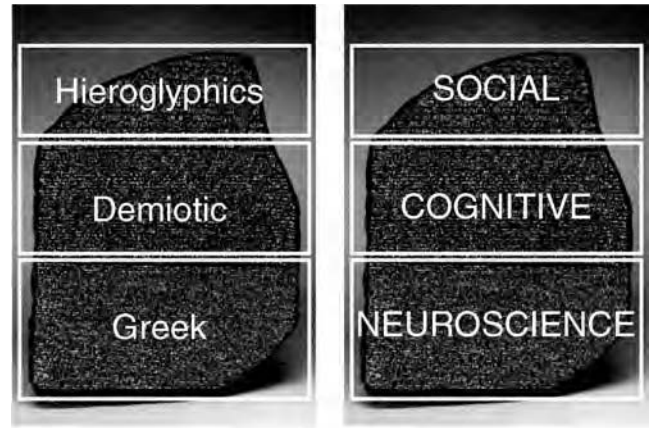
FIGURE 3.7. The ancient Egyptian *Rosetta Stone* provides an analogy for understanding how one can learn to translate between the differing languages used by social, cognitive, and neuroscience researchers. The original Stone contained the same text written in three different languages: hieroglyphics, Demotic, and Greek. The relationship between Greek and Egyptian hieroglyphics was decided when it was realized that Demotic characters spelling the name of a Pharaoh could be linked to their Greek and hieroglyphic counterparts. For social cognitive neuroscientists, the everyday language of information processing, originally derived from cognitive psychology, may provide a similar means for translating between neuroscience and social psychology.

Closing Comment: The Value of SCN

In the long run, SCN will succeed only if social psychologists and cognitive neuroscientists alike perceive the value of its approach. Toward that end, it will be important to recognize that SCN can be more than a simple addition of neuroscience data to social psychology or a new focus on social phenomena for cognitive neuroscience. SCN asks questions and aims to construct theories similar to, but importantly different than, those formulated by either of its parent disciplines. In comparison to social psychology, SCN offers the opportunity of constraining psychological theorizing through the use of neuroscience data that can use patterns of brain data to identify common and distinct processing systems underlying various behaviors. In comparison to cognitive neuroscience, SCN offers the opportunity to understand how contexts impact the way in which socioemotional contents are construed, thereby providing invaluable data about the functions associated with specific brain systems. For its practitioners, SCN can be an emergent discipline greater than the sum of its parental investment.

NOTES

1. During this time period paradigms for studying various forms of learning were developed, and theories to describe the nature of the learning process were advanced, that did not make reference to mental states. Ironically, these paradigms are used today to study fear, reward, and their underlying mental processes.
2. It is notable that, historically, there also has been debate concerning what is "social" about social cognition. Some have defined the term as we define it here in the context of SCN, but importantly distinguish it from the social psychology of cognition, which is quite different. For discussion, see Higgins (2000).
3. It also could be argued that cognitive psychology to a "step up," to become social cognition (Higgins, 2000).
4. As has been noted in many circles, this may be attributable to an academic system that places emphasis on distinguishing the contribution of independent individuals, rather than teams or groups, in order for individuals to be granted tenure.

REFERENCES

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278.

Allport, F. H. (1924). *Social psychology*. New York: Houghton Mifflin.

Amaral, D. G., Bauman, M. D., Capitanio, J. P., Lavenex, P., Mason, W. A., Mauldin-Jourdain, M. L., et al. (2003). The amygdala: Is it an essential component of the neural network for social cognition? *Neuropsychologia*, *41*(4), 517–522.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology*, *84*(4), 738–753.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*(13), 5627–5633.

Andrews, C. (1985). *The British Museum book of the Rosetta Stone*. New York: P. Bedrick Sons.

Ayduk, O., Mischel, W., & Downey, G. (2002). Attentional mechanisms linking rejection to hostile reactivity: The role of "hot" versus "cool" focus. *Psychological Science*, *13*(5), 443–448.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education and Behavior*, *31*(2), 143–164.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Bauman, M. D., Lavenex, P., Mason, W. A., Capitanio, J. P., & Amaral, D. G. (2004a). The development of mother–infant interactions after neonatal amygdala lesions in rhesus monkeys. *Journal of Neuroscience*, *24*(3), 711–721.

Bauman, M. D., Lavenex, P., Mason, W. A., Capitanio, J. P., & Amaral, D. G. (2004b). The development of social behavior following neonatal amygdala lesions in rhesus monkeys. *Journal of Cognitive Neuroscience*, *16*(8), 1388–1411.

Baumeister, R. F. (1998). The self. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 680–740). New York: McGraw-Hill.

Beauregard, M., Levesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *Journal of Neuroscience*, *21*(18), RC165.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*(Pt. 11), 2189–2202.

Beer, J. S., Heerey, E. A., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotion: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, *85*(4), 594–604.

Beer, J. S., Shimamura, A. P., & Knight, R. T. (2004). Frontal lobe contributions to executive control of cognitive and social behavior. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences: III* (pp. 1091–1104). Cambridge, MA: MIT Press.

Blakemore, S.-J., & Frith, U. (2004). How does the brain deal with the social world? *NeuroReport, 15*(1), 119–128.

Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: Where are we heading? *Trends in Cognitive Sciences*, *8*(5), 216–222.

Blascovich, J., Brennan, K., Tomaka, J., Kelsey, R. M., Hughes, P., Coad, M. L., et al. (1992). Affect intensity and cardiac arousal. *Journal of Personality and Social Psychology*, *63*(1), 164–174.

Blascovich, J., Mendes, W. B., Tomaka, J., Salomon, K., & Seery, M. (2003). The robust nature of the biopsychosocial model challenge and threat: A reply to Wright and Kirby. *Personality and Social Psychology Review*, *7*(3), 234–243.

Bremner, J. D., Vythilingam, M., Vermetten, E., Vaccarino, V., & Charney, D. S. (2004). Deficits in hippocampal and anterior cingulate functioning during verbal declarative memory encoding in midlife major depression. *American Journal of Psychiatry*, *161*(4), 637–645.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, *281*, 1185–1187.

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.

Brownell, H. H., Simpson, T. L., Bihrle, A. M., Potter, H. H., & Gardner, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia*, *28*(4), 375–383.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. *Journal of Cognitive Neuroscience*, *16*(1), 114–126.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., et al. (2004). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Neuron*, *42*(2), 323–334.

Buchel, C., Dolan, R. J., Armony, J. L., & Friston, K. J. (1999). Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 10869–10876.

Buchel, C., Morris, J., Dolan, R. J., & Friston, K. J. (1998). Brain systems mediating aversive conditioning: An event-related fMRI study. *Neuron*, *20*(5), 947–957.

Cacioppo, J. T. (1994). Social neuroscience: Autonomic, neuroendocrine, and immune responses to stress. *Psychophysiology*, *31*(2), 113–128.

Cacioppo, J. T. (2002). Social neuroscience: Understanding the pieces fosters understanding the whole and vice versa. *American Psychologist*, *57*(11), 819–831.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 173–191). New York: Guilford Press.

Cacioppo, J. T., Berntson, G. G., Malarkey, W. B., Kiecolt-Glaser, J. K., Sheridan, J. F., Poehlmann, K. M., et al. (1998). Autonomic, neuroendocrine, and immune responses to psychological stress: The reactivity hypothesis. *Annals of the New York Academy of Sciences*, *840*, 664–673.

Cacioppo, J. T., Berntson, G. G., Taylor, S. E., & Schacter, D. L. (2002). *Foundations in social neuroscience*. Cambridge, MA: MIT Press.

Cacioppo, J. T., Crites, S. L., Gardner, W. L., & Berntson, G. G.

(1994). Bioelectrical echoes from evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *Journal of Personality and Social Psychology*, *67*(1), 115–125.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (Eds.). (2000). *Handbook of psychophysiology* (2nd ed.). New York: Cambridge University Press.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*(11), 1077–1078.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., et al. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*(8), 1129–1138.

Canli, T., Desmond, J. E., Zhao, Z., Glover, G., & Gabrieli, J. D. (1998). Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport*, *9*(14), 3233–3239.

Cannon, W. B. (1987). The James-Lange theory of emotions: C critical examination and an alternative theory. By Walter B. Cannon, 1927. *American Journal of Psychology*, *100*(3–4), 567–586.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, USA*, *100*(9), 5497–5502.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., et al. (1999). In search of the self: A positron emission tomography study. *Psychological Science*, *10*, 26–34.

Crites, S. L., Jr., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorization: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology*, *68*(6), 997–1013.

Crosson, B., Radonovich, K., Sadek, J. R., Gokcay, D., Bauer, R. M., Fischler, I. S., et al. (1999). Left-hemisphere processing of emotional connotation during word generation. *NeuroReport*, *10*(12), 2449–2455.

Cunningham, W. A., & Johnson, M. K. (2007). Attitudes and evaluation: Toward a component process framework. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (pp. 227–245). New York: Guilford Press.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: FMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience*, *16*(10), 1717–1729.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), 1049–1056.

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences, USA*, *100*(4), 2157–2162.

Davey, L. M., Kaada, B. R., & Fulton, J. F. (1949). Effects on gastric secretion of frontal lobe stimulation. *Research Publications–Association for Research in Nervous and Mental Disease*, *29*, 617–627.

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience. *Psychological Bulletin*, *126*(6), 890–909.

Davidson, R. J., Pizzagalli, D., Nitschke, J. B., & Putnam, K. (2002). Depression: Perspectives from affective neuroscience. *Annual Review of Psychology*, *53*, 545–574.

DeCarli, C., Murphy, D. G., Gillette, J. A., Haxby, J. V., Teichberg, D., Schapiro, M. B., et al. (1994). Lack of age-related differences in temporal lobe volume of very healthy adults. *American Journal of Neuroradiology*, *15*, 689–696.

Decety, J., Chaminade, T., Grezes, J., & Meltzoff, A. N. (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation. *NeuroImage*, *15*(1), 265–272.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*(6), 3072–3077.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391.

Ebbinghaus, H. (1908). *Abriss der Psychologie* [Psychology: An elementary textbook]. Boston: D.C. Heath.

Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, *8*(7), 294–300.

Ellis, A. W. (1992). Cognitive mechanisms of face processing. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *335*, 113–119.

Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, *5*(4), 149–156.

Epley, N., Savitsky, K., & Gilovich, T. (2002). Empathy neglect: Reconciling the spotlight effect and the correspondence bias. *Journal of Personality and Social Psychology*, *83*(2), 300–312.

Erber, R. (1996). The self-regulation of moods. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self-regulation* (pp. 251–275). Mahwah, NJ: Erlbaum.

Eyal, T., Liberman, N., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology*, *86*(6), 781–795.

Fangel, C., & Kaada, B. R. (1960). Behavior "attention" and fear induced by cortical stimulation in the cat. *Electroencephalography and Clinical Neurophysiology*, *12*, 575–588.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge, MA: MIT Press.

Feldman Barrett, L., Ochsner, K. N., & Gross, J. J. (in press). Automaticity and emotion. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior*. New York: Psychology Press.

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. New York: Free Press.

Fiske, A. P., Kitayama, S., Markus, H. R., & Nisbett, R. E. (1998). *The cultural matrix of social psychology*. New York: McGraw-Hill.

Fossati, P., Hevenor, S. J., Graham, S. J., Grady, C., Keightley, M. L., Craik, F., et al. (2003). In search of the emotional self: An fMRI study using positive and negative emotional words. *American Journal of Psychiatry*, *160*(11), 1938–1945.

Fredrikson, M., Furmark, T., Olsson, M. T., Fischer, H., Andersson, J., & Langstrom, B. (1998). Functional neuroanatomical correlates of electrodermal activity: A positron emission tomographic study. *Psychophysiology*, *35*(2), 179–185.

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, *32*(6), 969–979.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *358*, 459–473.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind.' *Trends in Cognitive Sciences*, *7*(2), 77–83.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*(9), 396–403.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–432.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Ex-

ploring mechanisms for face recognition. *Vision Research*, *37*(12), 1673–1682.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504.

Gazzaniga, M. S. (Ed.-in-Chief). (1995). *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Gazzaniga, M. S. (Ed.-in-Chief). (2000). *The new cognitive neurosciences*. Cambridge, MA: MIT Press.

Gazzaniga, M. S. (Ed.-in-Chief). (2004). *The cognitive neurosciences* (3rd ed.). Cambridge, MA: MIT Press.

Giedd, J. N. (2004). Structural magnetic resonance imaging of the adolescent brain. *Annals of the New York Academy of Science*, *1021*, 77–85.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 89–150). New York: McGraw-Hill.

Gilbert, D. T., Fiske, S. T., & Lindzey, G. (1998). *The handbook of social psychology* (Vol. 2, 4th ed.). New York: McGraw-Hill.

Gilbert, D. T., & Wilson, T. D. (2000). Miswanting: Some problems in the forecasting of future affective states. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 178–197). New York: Cambridge University Press.

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, *131*(1), 76–97.

Goldenberg, G. (2004). An odd kind of fame—Stories of Phineas Gage. *Cortex*, *40*(3), 552–555.

Grady, C. L. (2002). Age-related differences in face processing: A meta-analysis of three functional neuroimaging experiments. *Canadian Journal of Experimental Psychology*, *56*(3), 208–220.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–562.

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224–237.

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology*, *14*(2), 233–238.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences, USA*, *95*(25), 15061–15065.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: Effects of a neocortical network on the limbic system. *NeuroReport*, *11*(1), 43–48.

Hariri, A. R., Mattay, V. S., Tessitore, A., Fera, F., & Weinberger, D. R. (2003). Neocortical modulation of the amygdala response to fearful stimuli. *Biological Psychiatry*, *53*(6), 494–501.

Harlow, H. F., Dodsworth, R. O., & Harlow, M. K. (1965). Total social isolation in monkeys. *Proceedings of the National Academy of Sciences, USA*, *54*(1), 90–97.

Harlow, H. F., & Harlow, M. (1962). Social deprivation in monkeys. *Sci Am*, *207*, 136–146.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *NeuroReport*, *11*(11), 2351–2355.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67.

Heatherton, T. F. (2004). *Introductory remarks to the Social Cognitive Neuroscience preconference*. Paper presented at the annual meeting of the Cognitive Neuroscience Society, San Francisco.

Heatherton, T. F., Macrae, C. N., & Kelley, W. M. (2004). What the social brain sciences can tell us about the self. *Current Directions in Psychological Science*, *13*(5), 190–193.

Heatherton, T. F., Polivy, J., Herman, C. P., & Baumeister, R. F. (1993). Self-awareness, task failure, and disinhibition: How attentional focus affects eating. *Journal of Personality*, *61*(1), 49–61.

Heckers, S., Weiss, A. P., Deckersbach, T., Goff, D. C., Morecraft, R. J., & Bush, G. (2004). Anterior cingulate cortex activation during cognitive interference in schizophrenia. *American Journal of Psychiatry*, *161*(4), 707–715.

Higgins, E. T. (1996a). *Emotional experiences: The pains and pleasures of distinct regulatory systems*. Hillsdale, NJ: Erlbaum.

Higgins, E. T. (1996b). The "self digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*(6), 1062–1083.

Higgins, E. T. (1997). *Biases in social cognition: "Aboutness" as a general principle*. Malden, MA: Blackwell.

Higgins, E. T. (1999). *Self-discrepancy: A theory relating self and affect*. New York: Psychology Press.

Higgins, E. T. (2000). Social cognition: Learning about what matters in the social world. *European Journal of Social Psychology*, *30*(1), 3–39.

Higgins, E. T., & Eccles-Parsons, J. E. (1983). Social cognition and the social life of the child: Stages as subcultures. In E. T. Higgins, D. N. Ruble, & W. W. Hartup (Eds.), *Social cognition and social development: A sociocultural perspective* (pp. 15–62). New York: Cambridge University Press.

Hodges, S., & Wegner, D. M. (1997). Automatic and controlled empathy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 311–339). New York: Guilford Press.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Ickes, W. (Ed.). (1997). *Empathic accuracy*. New York: Guilford Press.

Insel, T. R., & Fernald, R. D. (2004). How the brain processes social information: searching for the social brain. *Annual Review of Neuroscience*, *27*, 697–722.

Irwin, W., Davidson, R. J., Lowe, M. J., Mock, B. J., Sorenson, J. A., & Turski, P. A. (1996). Human amygdala activation detected with echo-planar functional magnetic resonance imaging. *NeuroReport*, *7*(11), 1765–1769.

Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, *24*(3), 771–779.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 22–36.

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain*, *125*(Pt. 8), 1808–1814.

Kaada, B. R., Andersen, P., & Jansen, J., Jr. (1954). Stimulation of the amygdaloid nuclear complex in unanesthetized cats. *Neurology*, *4*(1), 48–64.

Kagan, J. (1998). *Three seductive ideas*. Cambridge, MA: Harvard University Press.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Kanwisher, N., Stanley, D., & Harris, A. (1999). The fusiform face area is selective for faces not animals. *NeuroReport*, *10*(1), 183–187.

Kaplan, J. A., Brownell, H. H., Jacobs, J. R., & Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain and Language*, *38*(2), 315–333.

Keenan, J. P., Freund, S., Hamilton, R. H., Ganis, G., & Pascual-Leone, A. (2000). Hand response differences in a self-face identification task. *Neuropsychologia*, *38*(7), 1047–1053.

Keenan, J. P., McCutcheon, B., Freund, S., Gallup, G. G., Jr.,

Sanders, G., & Pascual-Leone, A. (1999). Left hand advantage in a self-face recognition task. *Neuropsychologia*, *37*(12), 1421–1425.

Keenan, J. P., Nelson, A., O'Connor, M., & Pascual-Leone, A. (2001). Self-recognition and the right hemisphere. *Nature*, *409*, 305.

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794.

Kemeny, M. E. (2003). The psychobiology of stress. *Current Directions in Psychological Science*, *12*(4), 124–129.

Kesler-West, M. L., Andersen, A. H., Smith, C. D., Avison, M. J., Davis, C. E., Kryscio, R. J., et al. (2001). Neural substrates of facial emotion processing using fMRI. *Brain Research: Cognitive Brain Research*, *11*(2), 213–226.

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences*, *8*(11), 501–507.

Kiecolt-Glaser, J. K., Cacioppo, J. T., Malarkey, W. B., & Glaser, R. (1992). Acute psychological stressors and short-term immune changes: what, why, for whom, and to what extent? *Psychosomatic Medicine*, *54*(6), 680–685.

Kim, H., Somerville, L. H., Johnstone, T., Alexander, A. L., & Whalen, P. J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *NeuroReport*, *14*(18), 2317–2322.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, *14*(1), 1–10.

Kluver, H., & Bucy, P. (1939). Preliminary analysis of functions of the temporal lobes in monkeys. *Archives of Neurology and Psychiatry*, *42*, 979–1000.

Kosslyn, S. M. (1999). If neuroimaging is the answer, what is the question? *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *354*, 1283–1294.

Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., et al. (2002). Bridging psychology and biology. The analysis of individuals in groups. *American Psychologist*, *57*(5), 341–351.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Review Neuroscience*, *2*(9), 635–642.

Kosslyn, S. M., & Maljkovic, V. M. (1990). Marr's metatheory revisited. *Concepts in Neuroscience*, *1*(2), 239–251.

Krueger, J. I. (2003). Return of the ego—Self-referent information as a filter for social prediction: Comment on Karniol (2003). *Psychological Review*, *110*(3), 585–590.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: A mixed-trial fMRI study. *Neuron*, *20*(5), 937–945.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 202–255). New York: Cambridge University Press.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., et al. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, *35*(11), 1437–1444.

Larsen, J. T., McGraw, A. P., Mellers, B. A., & Cacioppo, J. T. (2004). The agony of victory and thrill of defeat: Mixed emotional reactions to disappointing wins and relieving losses. *Psychological Science*, *15*(5), 325–330.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford, UK: Oxford University Press.

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *Journal of Abnormal and Social Psychology*, *69*, 195–205.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Levesque, J., Eugene, F., Joanette, Y., Paquette, V., Mensour, B., Beaudoin, G., et al. (2003). Neural circuitry underlying voluntary suppression of sadness. *Biological Psychiatry*, *53*(6), 502–510.

Levesque, J., Joanette, Y., Mensour, B., Beaudoin, G., Leroux, J. M., Bourgouin, P., et al. (2004). Neural basis of emotional self-regulation in childhood. *Neuroscience*, *129*(2), 361–369.

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin*, *126*(1), 109–137.

Lieberman, M. D. (2003). Reflective and reflexive judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. V. Hippel (Eds.), *Social judgments: Explicit and implicit processes* (pp. 44–67). New York: Cambridge University Press.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). San Diego, CA: Academic Press.

Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, *8*(6), 720–722.

Lieberman, M. D., Jarcho, J. M., & Satpute, A. B. (2004). Evidence-based and intuition-based self-knowledge: An fMRI study. *Journal of Personality and Social Psychology*, *87*(4), 421–435.

Lieberman, M. D., & Pfeifer, J. H. (2005). The self and social perception: Three kinds of questions in social cognitive neuroscience. In A. Easton & N. Emery (Eds.), *Cognitive neuroscience of social behavior* (pp. 195–235). New York: Psychology Press.

Lieberman, M. D., Schreiber, D., & Ochsner, K. N. (2003). Is political sophistication like riding a bicycle? How cognitive neuroscience can inform research on political attitudes and decision-making. *Political Psychology*, *24*, 681–704.

Luna, B., & Sweeney, J. A. (2001). Studies of brain and cognitive maturation through childhood and adolescence: A strategy for testing neurodevelopmental hypotheses. *Schizophrenia Bulletin*, *27*(3), 443–455.

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., et al. (2001). Maturation of widely distributed brain function subserves cognitive development. *NeuroImage*, *13*(5), 786–793.

MacLean, P. D. (1969). *A triune concept of the brain and behaviour*. Toronto: University of Toronto Press.

Macmillan, M. (2000). *An odd kind of fame: Stories of Phineas Gage*. Cambridge, MA: MIT Press.

Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, *14*, 647–654.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Mason, M. F., Banfield, J. F., & Macrae, C. N. (2004). Thinking about actions: The neural substrates of person knowledge. *Cerebral Cortex*, *14*(2), 209–214.

Mayberg, H. S. (2003). Positron emission tomography imaging in depression: A neural systems perspective. *Neuroimaging Clinics of North America*, *13*(4), 805–815.

Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., et al. (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *American Journal of Psychiatry*, *156*(5), 675–682.

Meaney, M. J. (2001). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual Review of Neuroscience*, *24*, 1161–1192.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *358*, 491–500.

Mendoza-Denton, R., Ayduk, O., Mischel, W., Shoda, Y., & Testa, A. (2001). Person X situation interactionism in self-encoding (I am . . . when . . .): Implications for affect regulation and social information processing. *Journal of Personality and Social Psychology*, *80*(4), 533–544.

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., et al. (2002). Attentional control in the aging brain: Insights from an fMRI study of the stroop task. *Brain and Cognition*, *49*(3), 277–296.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268.

Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-behavior profiles as a locus of consistency in personality. *Current Directions in Psychological Science*, *11*(2), 50–54.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences, USA*, *99*(23), 15238–15243.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, *24*, 4912–4917.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2005). The link between social cognition and the self referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(18), 1306–1315.

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences, USA*, *96*(4), 1680–1685.

Morrison, I., Lloyd, D., di Pellegrino, G., & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: Is empathy a multisensory issue? *Cognitive, Affective, and Behavioral Neuroscience*, *4*(2), 270–278.

Ochsner, K. N. (2005). Characterizing the functional architecture of affect regulation: Emerging answers and outstanding questions. In J. T. Cacioppo (Ed.), *Social neuroscience* (pp. 245–268). Cambridge, MA: MIT Press.

Ochsner, K. N., Beer, J. S., Robertson, E., Cooper, J., Gabrieli, J. D. E., Kihlstrom, J. F., et al. (2005). The neural correlates of direct and reflected self-knowledge. *NeuroImage*, *28*(4), 797–814.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*(8), 1215–1229.

Ochsner, K. N., & Feldman Barrett, L. (2001). A multiprocess perspective on the neuroscience of emotion. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions* (pp. 38–81). New York: Guilford Press.

Ochsner, K. N., & Gross, J. J. (2004). Thinking makes it so: A social-cognitive neuroscience approach to emotion regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 229–255). New York: Guilford Press.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences*, *9*(5), 242–249.

Ochsner, K. N., & Gross, J. J. (2005). Putting the "I" and the "me" in emotion regulation: A reply to Northoff. *Trends in Cognitive Sciences*, *9*(5), 242–249.

Ochsner, K. N., Knierim, K., Ludlow, D., Hanelin, J., Ramachandran, T., & Mackey, S. (2004). Reflecting upon feelings: An fMRI study of neural systems supporting the attribution of emotion to self and other. *Journal of Cognitive Neuroscience*, *16*(10), 1746–1772.

Ochsner, K. N., & Kosslyn, S. M. (1999). The cognitive neuroscience approach. In B. M. Bly & D. E. Rumelhart (Eds.), *Cognitive science* (pp. 319–365). San Diego, CA: Academic Press.

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, *56*(9), 717–734.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., et al. (2004). For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, *23*(2), 483–499.

Ochsner, K. N., & Schacter, D. L. (2000). A social cognitive neuroscience approach to emotion and memory. In J. C. Borod (Ed.), *The neuropsychology of emotion* (pp. 163–193). London: Oxford University Press.

Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science*, *11*(2), 62–66.

Öhman, A., Flykt, A., & Lundqvist, D. (2000). Unconscious emotion: Evolutionary perspectives, psychophysiological data and neuropsychological mechanisms. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 296–327). New York: Oxford University Press.

Osgood, C. E. (1976). *Focus on meaning: I. Explorations in semantic space*. Oxford, UK: Mouton.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Papez, J. W. (1958). Visceral brain, its component parts and their connections. *Journal of Nervous and Mental Disease*, *126*(1), 40–56.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neural and emotional stimuli. *Brain Research: Cognitive Brain Research*, *15*(1), 31–45.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, USA*, *99*(17), 11458–11463.

Phan, K. L., Fitzgerald, D. A., Nathan, P. J., Moore, G. J., Uhde, T. W., & Tancer, M. E. (2005). Neural substrates for voluntary suppression of negative affect: A functional magnetic resonance imaging study. *Biological Psychiatry*, *57*(3), 210–219.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*(2), 331–348.

Phillips, M. L., Bullmore, E. T., Howard, R., Woodruff, P. W., Wright, I. C., Williams, S. C., et al. (1998). Investigation of facial recognition memory and happy and sad facial expression perception: An fMRI study. *Psychiatry Research*, *83*(3), 127–138.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*, 495–498.

Posner, M. I., & DiGirolamo, G. J. (2000). Cognitive neuroscience: Origins and promise. *Psychological Bulletin*, *126*(6), 873–889.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Puce, A., & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *358*, 435–445.

Quintana, J., Wong, T., Ortiz-Portillo, E., Marder, S. R., & Mazziotta, J. C. (2004). Anterior cingulate dysfunction during choice anticipation in schizophrenia. *Psychiatry Research*, *132*(2), 117–130.

Ramachandran, V. S. (1995). Anosognosia in parietal lobe syndrome. *Conscious Cognitive*, *4*(1), 22–51.

Rauch, S. L., Savage, C. R., Alpert, N. M., Fischman, A. J., & Jenike, M. A. (1997). The functional neuroanatomy of anxiety: A study of three disorders using positron emission tomography and symptom provocation. *Biological Psychiatry*, *42*(6), 446–452.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology*, *12*(1), 95–114.

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., et al. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of Cognitive Neuroscience*, *12*(1), 174–187.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychological*, *80*(2), 340–352.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology*, *68*(3), 404–417.

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., et al. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, *55*(6), 594–602.

Rolls, E. T. (1999). *The brain and emotion*. Oxford, UK: Oxford University Press.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.

Ross, L., & Ward, A. (1996). *Naive realism in everyday life: Implications for social conflict and misunderstanding*. Hillsdale, NJ: Erlbaum.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, *4*(5), 546–550.

Salat, D. H., Kaye, J. A., & Janowsky, J. S. (2001). Selective preservation and degeneration within the prefrontal cortex in aging and Alzheimer's disease. *Archives of Neurology*, *58*, 1403–1408.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*, 1755–1758.

Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience. Toward strong inference in attributing function to structure. *American Psychologist*, *51*(1), 13–21.

Saxe, R. (2005). Against simulation: The argument from error. *Trends in Cognitive Sciences*, *9*(4), 174–179.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124.

Schacter, D. L. (1997). The cognitive neuroscience of memory: Perspectives from neuroimaging research. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences*, *352*, 1689–1695.

Schacter, D. L., Alpert, N. M., Savage, C. R., Rauch, S. L., & Albert, M. S. (1996). Conscious recollection and the human hippocampal formation: Evidence from positron emission tomography. *Proceedings of the National Academy of Sciences, USA*, *93*(1), 321–325.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379–399.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. New York: Oxford University Press.

Schmitz, T. W., Kawahara-Baccus, T. N., & Johnson, S., C. (2004). Metacognitive evaluation, self-reference, and the right prefrontal cortex. *NeuroImage*, *22*, 941–947.

Shin, L. M., Kosslyn, S. M., McNally, R. J., Alpert, N. M., Thompson, W. L., Rauch, S. L., et al. (1997). Visual imagery and perception in posttraumatic stress disorder. A positron emission tomographic investigation. *Archives of General Psychiatry*, *54*(3), 233–241.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*, 1157–1162.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, *64*(6), 885–896.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394.

Taylor, S. E. (1998). *The social being in social psychology*. New York: McGraw-Hill.

Taylor, S. E., Kemeny, M. E., Reed, G. M., Bower, J. E., &

Gruenewald, T. L. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, *55*(1), 99–109.

Taylor, S. F., Liberzon, I., Fig, L. M., Decker, L. R., Minoshima, S., & Koeppe, R. A. (1998). The effect of emotional content on visual recognition memory: A PET activation study. *NeuroImage*, *8*(2), 188–197.

Thomas, K. M., King, S. W., Franzen, P. L., Welsh, T. F., Berkowitz, A. L., Noll, D. C., et al. (1999). A developmental functional MRI study of spatial working memory. *NeuroImage*, *10*(3, Pt. 1), 327–338.

Tillfors, M., Furmark, T., Marteinsdottir, I., Fischer, H., Pissiota, A., Langstrom, B., et al. (2001). Cerebral blood flow in subjects with social phobia during stressful speaking tasks: A PET study. *American Journal of Psychiatry*, *158*(8), 1220–1226.

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*(1), 63–72.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*(3), 403–421.

Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences, USA*, *91*(6), 2016–2020.

Uchino, B. N., Cacioppo, J. T., Malarkey, W., Glaser, R., & Kiecolt-Glaser, J. K. (1995). Appraisal support predicts age-related differences in cardiovascular function in women. *Health Psychological*, *14*(6), 556–562.

van Baaren, R. B., Horgan, T. G., Chartrand, T. L., & Dijkmans, M. (2004). The forest, the trees, and the chameleon: Context dependence and mimicry. *Journal of Personality and Social Psychology*, *86*(3), 453–459.

van Baaren, R. B., Maddux, W. W., Chartrand, T. L., de Bouter, C., & van Knippenberg, A. (2003). It takes two to mimic: Behavioral consequences of self-construals. *Journal of Personality and Social Psychology*, *84*(5), 1093–1102.

Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., & Dolan, R. J. (2002). Neural response to emotional faces with and without awareness: Event-related fMRI in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia*, *40*(12), 2156–2166.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron*, *30*(3), 829–841.

Wager, T. D., & Feldman Barrett, L. (2004). *From affect to control: Functional specialization of the insula in motivation and regulation*. Available: www.apa.org/psycextra

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *NeuroImage*, *22*(4), 1679–1693.

Wager, T. D., & Ochsner, K. N. (in press). Sex differences in the emotional brain. *NeuroReport*.

Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: A meta-analysis of findings from neuroimaging. *NeuroImage*, *19*(3), 513–531.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience*, *3*(4), 255–274.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, *16*(1), 56–63.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655–664.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*(1), 117–142.

Winner, E., Brownell, H., Happe, F., Blum, A., & Pincus, D. (1998).

Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language*, *62*(1), 89–106.

Wright, R. A., & Kirby, L. D. (2003). Cardiovascular correlates of challenge and threat appraisals: A critical examination of the biopsychosocial analysis. *Personality and Social Psychology Review*, *7*(3), 216–233.

Young, A. W., & Ellis, H. D. (1989). Childhood prosopagnosia. *Brain and Cognition*, *9*(1), 16–47.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5), 889–898.

Zald, D. H., Lee, J. T., Fluegel, K. W., & Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain*, *121*(Pt. 6), 1143–1154.

Zysset, S., Huber, O., Ferstl, E., & von Cramon, D. Y. (2002). The anterior frontomedian cortex and evaluative judgment: An fMRI study. *NeuroImage*, *15*(4), 983–991.

Zysset, S., Huber, O., Samson, A., Ferstl, E. C., & von Cramon, D. Y. (2003). Functional specialization within the anterior medial prefrontal cortex: A functional magnetic resonance imaging study with human subjects. *Neuroscience Letters*, *335*(3), 183–186.

PART II
COGNITIVE SYSTEM

CHAPTER 4

Prediction

The Inside View

DAVID DUNNING

Friends may come and go, but there is one companion that never leaves a person's side. This companion is hardly the easiest of acquaintances, making a habit of mystery and suspense, cloaking moves in a constant uncertainty of what might happen next. In reaction, people spend a great deal of time, thought—and even money—to analyze this companion, to anticipate the twists and turns this friend might take. This work is never ending, and the companion's actions remain subject to unpredictability and surprise throughout a person's life.

This lifetime companion is the future, and a simple survey of everyday life reveals just how people spend a good deal of their hours thinking, talking, and calculating about what this friend—the future—might hold. This obsession arises in both personal and professional spheres. In their personal lives, people strive to anticipate what other people will do (e.g., *Will my boss give me a raise?*), what they themselves will do (i.e., *Will I make the high school basketball team?*), as well as what will happen in the world at large (e.g., *Will the rain hold off until the picnic?*). Trying to answer these questions, people may consult their friends, their elders, even a ouija board or astrological chart for some clue about how the future might behave.

This fixation on the future is equally apparent in people's professional lives. The prediction business is a $200 billion industry. There are at least 200,000 people licensed by the National Association of Security Dealers to estimate future stock prices and company earnings. Roughly $5 billion a year is devoted to forecasting the weather (Sherden, 1998). And these figures do not in-clude professions that have prediction as a core element in their activities, such as doctors providing prognoses of their patients' future health, personnel directors divining which job applicant will best fulfill job duties, and city planners calculating the environmental impact of new construction.

There is, however, a troubling truth about the task of prediction. In an observation that has been variously attributed to Nobel laureate Nils Bohr and to baseball great Yogi Berra, among others, prediction is hard, especially about the future. People strive effortlessly to anticipate the future, but their forecasts often tend to be far from perfect and sometimes prove dramatically wrong. People fail to hire the right person for job. They misestimate their chances of getting into the "right" law school. They find themselves working on projects far later than they anticipated working on them.

Even professionals whose careers depend on providing plausible projections meet with only modest success (Sherden, 1998). Despite the billions of dollars spent on weather prediction, meteorologists still accurately predict future weather only a few days into the future (Ahrens, 1991). Despite its obvious social consequences, forensic psychologists and psychiatrists predicting which criminals will prove dangerous again in the future achieve only meager achievement at the task, and they are often outperformed by simple mathematical formuli that can be quickly designed to do the same task (Dawes, 1979; Dawes, Faust, & Meehl, 1989; Meehl, 1954). And, despite a history of over 100 years, economists know for sure that the business cycle exists, but they cannot pre-

cisely project when the next boom or recession will be-
gin. As famous financial guru Warren Buffett put it, of-
ten the only apparent value of economic forecasting is to
make the fortunetellers look good.

This chapter focuses on the psychology of prediction
and outlines some of its principles. It describes how peo-
ple approach the important, ever-present, yet vexing task
of prediction, delineating the strategies that people
commonly pursue when they calculate what the future
holds for them as well as the usual outcome of those
predictions—that is, how accurate those predictions tend
to be.

In doing so, the chapter describes three different sets
of principles regarding the psychology of prediction.
First, the chapter enumerates *principles of outcome*, de-
scribing what patterns researchers typically find when
they explore the level of accuracy and bias in prediction.
Second, the chapter outlines *principles of process*, describ-
ing the psychological processes by which people reach
their predictions. Third, the chapter outlines *principles of
improvement*, describing alternative ways that people
could approach their predictions to ensure greater accu-
racy. Table 4.1 provides an outline the material con-
tained in this chapter, as well as a list of the types of prin-
ciples that are described and discussed in it.

In giving such an "inside" view to the psychology of
prediction, the chapter sets forth two assertions. The first
is that the real story of prediction is not how people go
about doing it but rather on what they *fail* to do but could
if they developed the habit; people could make more ac-
curate predictions if they paid more attention to crucial
information they potentially have at their disposal. The
second assertion is that prediction is an inherently diffi-

cult task. People very rarely have all the necessary infor-
mation at their disposal to make accurate forecasts. As a
consequence, people should not be expected to be over-
whelmingly accurate in the predictions they make—and
the meager success people achieve in personal and pro-
fessional realms should not be considered that much of a
surprise. After discussing these two assertions, I finish
the chapter by talking about additional steps, suggested
by recent research, that people can take to reach more ac-
curate forecasts of what future lays ahead of them.

PRINCIPLES OF OUTCOME: COMMON BIASES IN PREDICTION

Before discussing the strategies people follow in making
predictions, it is important to stipulate, with data, that
predictions tend to be far from perfect—and, thus, out-
line the principles of outcome that characterize typical
imperfections. History is replete with bad predictions. In
1895, Lord Kelvin pronounced that heavier-than-air fly-
ing machines were a physical impossibility. The Institute
for Boiler and Radiator Manufacturers in 1955 predicted
that U.S. homes would soon be heated and cooled by
their own household nuclear reactors; vacuum maker
Alex Lewyt that same year forecast that nuclear-powered
vacuum cleaners would be in the stores by 1965. Decca
Recording Company in 1962 rejected a contract with the
Beatles because it felt that guitar music was on the way
out (Kusterbeck, 2004).

Of course, not all predictions are bad ones. History is
also filled with accurate forecasts that could be just as eas-
ily recounted. Thus, beyond counting up the good pre-
dictions with the bad, is there any comprehensive way to
describe the worth or worthlessness of people's predic-
tions? Such a task is difficult, but the psychological litera-
ture provides clues about the systematic types of errors
that people make in their forecasts. In all, predictions
tend to be characterized by two general shortcomings,
which serve as two principles of outcome. The first prin-
ciple is that people tend to make predictions of the fu-
ture that are too optimistic. The second principle is that
people tend to be overly confident in their predictions,
overestimating the likelihood that those predictions will
prove right, whether those predictions point in an opti-
mistic or pessimistic direction.

Undue Optimism in Predictions

People tend to render predictions that are overly opti-
mistic in that the future they ultimately encounter of-
ten does not match the rosy outlook described in their
forecasts. Evidence for this first principle of outcome,
undue optimism, comes in two forms in the psycholog-
ical literature. First, when people make predictions,
they overpredict the occurrence of positive actions and
outcomes and underpredict negative ones. Second,
when people estimate how long it will take to com-
plete a task, they typically underestimate the amount
of time it will take.

TABLE 4.1. Outline of Principles of Prediction Discussed in This Chapter

Principles of outcome
  Undue optimism
    Overprediction of desirable events
    Planning fallacy
  Overconfidence

Principles of process
  Prediction as scenario-building process (adopting an "inside"
    approach)
  Incomplete scenario building
    Focusing on abstract and neglecting the concrete
    Focusing on central outcomes and neglecting alternatives
    Focusing on the optimistic and neglecting the pessimistic
    Focusing on distinctive features of events and neglecting
      shared ones (focalism)
    Focusing on the strength of evidence and neglecting its weight
  Limited utility of scenario building
    Unknown and unknowable situational details
    Inaccessibility of emotions and their impact

Principles of improvement
  Prediction as data-based process (adopting an "outside"
    approach)
  Cognitive repairs
  Aggregating predictions

Overprediction of Desirable Events

The best of all possible worlds tends to be the one that people believe will arrive in the future. For example, when people are asked to describe their past, they provide a portrait containing a mixture of positive and negative features. However, when they describe the future, the picture they paint tends to be uniformly positive (Ross & Newby-Clark, 1998).

One sees this rose-colored optimism in the predictions people make of their future behavior. People tend to overestimate the likelihood that they will buy a flower for charity, volunteer to work at an American Cancer Society fundraiser, vote in an upcoming election, volunteer to do a long and arduous experiment, and have a long-lasting romantic relationship (Epley & Dunning, 2000, 2004; Sherman, 1980). They overestimate how likely it is that they will challenge an aggressive opponent in a negotiation (Dieckmann, Tenbrusel, & Galinsky, 2003) and an employment interviewer who sexually harasses them (Woodzicka & LaFrance, 2001). College students overpredict their performance on their course examinations (Gilovich, Medvec, & Kerr, 1993). Women overpredict the effectiveness of the contraceptives they use, as well as whether they will use them in the month ahead (Hynie & Lydon, 1996).

One also sees this undue optimism revealed in people's actions. People spend significant amounts of money on home exercise equipment, yet a 1997 Sporting Goods Manufacturers Association survey found that the equipment was not regularly used in around a third of the relevant households. People buy memberships in health clubs—and then go so seldomly to the club that they would have been better off to have paid for each individual visit separately rather than for the rather expensive memberships (Della Vigna & Malmendier, 2002). People go on diets yet fail to keep the weight off. One typical reaction is to go on the diet again, causing a syndrome of yo-yo weight loss and gain that can be more injurious to health than just living with one's original weight (Polivy & Herman, 2002). Corporate CEOs announce mergers and then typically watch the stock price of their company and the one they want to acquire go *down*, not up, contrary to their optimistic assessment (Hayward & Hambrick, 1997).

The Planning Fallacy

People also display undue optimism when they specifically estimate how long it will take them to complete tasks or projects they are currently working on, a phenomenon known as the *planning fallacy* (Buehler, Griffin, & Ross, 1994). For example, taxpayers filing their returns typically predict that they will finish the forms 1 week sooner than they do in actuality—2 weeks sooner if they expect that they will receive a sizable refund (Buehler, Griffin, & MacDonald, 1997). College seniors completing a thesis "realistically" predict that they will complete their work 3 weeks sooner than they actually do—and even their "worst case" prediction tends to be 1 week too optimistic relative to the actual amount of time they take (Buehler et al., 1994).

The prevalence of the planning fallacy is not constrained to individuals making informal predictions in their personal life. Business predictions made by organizations tend to be characterized by the same bias. Research on "grand scale" projects, such as the Golden Gate Bridge, Erie Canal, and the Empire State Building, shows that these projects tend to take much longer to build and cost much more money than originally anticipated. Moreover, once built, the revenue stream these projects generate tends to come in much lower than initially estimated (Shapira & Berndt, 1997). In perhaps the most illustrative example, the Sydney Opera House in 1957 was predicted to open in 1963 at a cost of $7 million. It finally opened in 1973, in a scaled-down version, after $102 million had been spent (Buehler, Griffin, & Ross, 2002).

Overconfidence in Prediction

People also tend to imbue their predictions, whether optimistic or pessimistic, with too much confidence. As a long history of psychological research indicates, when people are asked to estimate how likely their predictions are to prove right, the likelihood they cite is too high relative to actual accuracy (for an informative review, see Lichtenstein, Fischhoff, & Phillips, 1982). For example, when an individual states that he or she is 80% certain to receive an "A" in a relevant course, the real likelihood that the student will receive that grade tends to be significantly lower. Indeed, even when people express 100% certainty in their predictions, they tend, depending on the circumstance, to be wrong roughly 20% of the time (Fischhoff, Slovic, & Lichtenstein, 1977; see also Dunning, Griffin, Milojkovic, & Ross, 1990; Vallone, Griffin, Lin, & Ross, 1990). This typical overconfidence serves as the second principle of outcome in the psychology of prediction.

Such overconfidence has been observed in a wide variety of contexts. Students predicting, for example, whether they will take part in the school play, change their career plans, question their decision to attend their college, take part in intramural sports, and change roommates tend to overestimate the accuracy of their predictions, on average, by 10 to 15 percentage points (Dunning & Story, 1991; Vallone et al., 1990). Professionals, too, make predictions with too much certainty. Several examples abound in the medical literature. Physicians diagnosing pneumonia dramatically overestimate the accuracy of their decisions (Christensen-Szalanski & Bushyhead, 1981). Surgical trainees do the same with their diagnoses of possible fractures after examining patient X-rays (Oksam, Kingma, & Klasen, 2000). Clinical psychologists overrate the accuracy of their diagnostic inferences after reading patient case materials (Oskamp, 1965). Similarly, professionals engaged in international security and politics (e.g., think-tank policy analysts and government intelligence analysts) overestimate how likely their predictions about world events will later prove accurate (Tetlock, 2002).

Why might prediction fall prey to these twin biases of undue optimism and unwarranted confidence? One answer to this question may lie in the strategy people adopt when they engage in the task of prediction. However, perhaps the better answer lies in what people *fail* to do when they follow that strategy.

PRINCIPLES OF PROCESS: I. THE SCENARIO-BASED NATURE OF PREDICTION

Prediction might be hard, but describing the basic strategy people use to reach predictions is somewhat easier. A good deal of evidence from cognitive and social psychology indicates that people estimate the likelihood of an event by constructing scenarios that would suggest it. To the extent that scenarios suggesting an event are simple, easy to construct, plausible, and numerous, people conclude that an event is more likely to happen (Atance & O'Neil, 2001; Dougherty, Gettys, & Ogden, 1999; Dougherty, Gettys, & Thomas, 1997; Kahneman & Lovallo, 1993; Kahneman & Tversky, 1982; Koehler, 1994; Koriat, Lichtenstein, & Fischhoff, 1980; Lagnado & Sloman, 2004; Mynatt, Doherty, & Dragan, 1993). In essence, the act of prediction is one of "mental simulation," that is, taking current conditions and building causal chains of events that would lead to the focal outcome in question. In the literature, this scenario-building strategy is referred to as the "inside approach," in that people set up a mental model of the relevant situation and explore its internal dynamics to see what outcomes are most plausibly suggested (Kahneman & Tversky, 1982; Lagnado & Sloman, 2004). In terms of a principle of process, this reliance on scenario building serves as an overarching principle in the psychology of prediction.

This inside, scenario-building view of prediction is evident in many corners of psychology. For example, in their discussion of the *representativeness heuristic*, Kahneman and Tversky (1973) suggested that people predict uncertain events by taking what they know of the world to see if it matches their model of some outcome in question. For example, if Donna loves books, speaks three languages, and shows little concern for money, one can easily spin a story that Donna will become a professor of comparative literature. Her talents and preferences squarely match the job profile of the typical comparative literature professor, and thus it is easy to simulate Donna favoring this career choice if she were exposed to it. The scenario needed to translate Donna's current characteristics to this future career choice involves neither too many steps nor ones that are implausible. In contrast, spinning scenarios in which Donna becomes an economist or a lawyer requires more numerous steps of uncertain plausibility, and thus these future career choices would seem less likely.

Asking people to describe their thoughts while they make predictions also reveals the storytelling approach people bring to the task. Dougherty and colleagues (1997) presented college students with vignettes and asked them to predict the most likely explanation for the final outcome. For example, students read about a firefighter who had died outside a building that was on fire and were asked to estimate the likelihood that he had died of smoke inhalation. Students were also asked explicitly to describe their thoughts as they reached their estimates. Students overwhelmingly thought in terms of scenarios, linking how the circumstances they read about would lead to smoke inhalation (e.g., the smoke was so thick he couldn't see) or would not (e.g., firefighters typically wear oxygen masks), and their likelihood estimates closely tracked how well their scenarios suggested smoke inhalation versus some alternative cause.

A scenario-building approach to prediction is also evident from Buehler and colleagues' (1994; see also Buehler & Griffin, 2003) work on the planning fallacy. In one study, Buehler and colleagues asked college students to think aloud that they predicted how long it would take them to complete a project they needed to get done for class. Students dwelled on mental simulations for getting the task done: A full 74% of thoughts focused on creating scenarios of how they would complete the task—either the plans they had for getting the task done or the obstacles that lay in their path. The bulk of the rest (15%) centered on specific deadlines they were laboring under, which also can be construed as a major ingredient in the scenarios they built to arrive at their estimates.

PRINCIPLES OF PROCESS: II. INCOMPLETE SCENARIO BUILDING

Running mental simulations to predict future events might be a more or less reasonable strategy, but it would be a more reasonable stratagem if people did not practice some unsound habits when constructing their simulations. People are not very thorough or exhaustive in the scenarios they build. Instead, when adopting an inside approach, they construct only partial scenarios of the future, systematically ignoring possibilities that they know, if pressed, are relevant. As such, an adequate discussion about the psychology of prediction must turn quickly from the tactics people follow in prediction to spend most of its time talking about what stratagems people eschew. Thus, paradoxically, although the first principle of process regarding prediction is that it is scenario based, the second overarching principle of prediction is that people hardly do a comprehensive job of building those scenarios. The real story of prediction must focus on what people ignore as they construct their forecasts.

In total, the psychological literature suggests five separate and systematic ways in which people's scenarios fail to be comprehensive representations of possible future events, serving as subprinciples under the overarching principle of incomplete scenario building. First, people often base their scenarios on only a few abstract, higher-order features of the events they strive to predict, particularly when those events are far off in the future. Second, they focus almost exclusively on the particular outcome their attention is drawn to, neglecting to simulate the likelihood that some other alternative outcome might

transpire. Third, they accentuate the positive too much in their scenarios while overlooking the negative. Fourth, when comparing the likelihood of one event to another, they focus too much on characteristics that differentiate the two events and too little on features the events have in common. Finally, when people collect information to build their scenarios, they often fail to take into account the reliability and validity of that information. All these habits lead to predictions that err systematically, as well as predictions that are held with too much confidence.

Focusing on the Abstract and Neglecting the Concrete

When simulating uncertain events, people do not construct an exhaustive set of concrete and detailed possibilities. Instead, they often stay at an abstract level, constructing only a partial simulation that dwells on a few central features of the relevant situation. This serves as the first subprinciple under the overarching principle of incomplete scenario building.

The fact that people fail to consider all possible concrete instantiations of an event has been shown by work on support theory (Rottenstreich & Tversky, 1997; Tversky & Koehler, 1994). For example, Tversky and Koehler asked some of their participants to predict the likelihood that someone had died of "natural causes" in the past year. The likelihood estimate participants gave was significantly smaller than the estimates given by another set of participants who were asked to estimate the likelihood of a number of concrete ways people can die of natural causes (e.g., cancer and heart disease). Similarly, participants asked to estimate the chance that someone died from homicide provided higher likelihood estimates when asked to consider separate sets of potential perpetrators (e.g., acquaintances and strangers) than when just considering the general category of homicide alone (Rottenstreich & Tversky, 1997).

Other work shows that overly simple scenarios that skate across details might underlie the planning fallacy. Kruger and Evans (2004) proposed that the planning fallacy arose because people fail to "unpack" the steps needed to complete complex tasks, such as holiday shopping, preparing food, getting ready for a date, or formatting a document. In their studies, they asked some of their participants to furnish more complete scenarios in which those participants detailed and considered all the concrete subtasks they would have to confront in order to finish a complex project. Relative to a control condition, participants who explicitly and deliberately enumerated all the relevant subtasks produced more pessimistic estimates of the time it would take to complete their overall tasks and in two experiments provided more accurate assessments.

Jorgensen (2004b) found a similar pattern when he examined software development teams estimating how much time it would take them to generate a new piece of software. Teams provided more accurate estimates when they pursued a bottom-up strategy, estimating how long it would take to complete each subtask and then simply summing the results to estimate the total amount of time to finish the development project. Teams provided more optimistic and less realistic estimates when they used a top-down strategy, focusing on the project as a whole.

This tendency to rely on simple, abstract, and stereotypical simulations is exacerbated to the extent that people consider events that dwell in the far off future. When focusing on distant events, people concentrate on a few general, higher-order, features of the situation, basing their simulation largely on schematic knowledge of the situation at hand. For example, if Archie asked Edith if she would like to take a vacation on Maui next summer, Edith might base her judgment on just a few general and schematic features that come to mind, such as warm weather, beaches, and comforting tropical breezes. What is left out of her simulation are lower-level, more concrete features of the situation, even though they obviously exist, such as the long plane flight to the island, the need to make hotel reservations, and whatever work might be left at the office while she is gone.

A good deal of recent research has shown that people tend to think of the distant future in terms of high-level features—building simple, abstract, and schematic scenarios—giving short shrift to lower-level features that may have a significant impact on their preferences, actions, and outcomes (Trope & Liberman, 2003; Vallacher & Wegner, 1987). People tend to think of more abstract features when they consider the distant future. For example, Liberman and Trope (1998) asked participants to describe what came to mind when they considered such actions as *locking a door* either in the distant future (e.g., next year) or some time more immediate (e.g., tomorrow). Participants tended to report more higher-ordered and abstract descriptions of the action when it took place in the distant future (e.g., *securing the house*) than when in took place more immediately (e.g., *putting the key in the lock*).

Other measures showed that people think more simply about events in the distant future than events that are almost upon them. When considering an event such as *moving out of an apartment* in the distant future, participants sorted objects associated with the event (e.g., computers, stereo, and clothes) into fewer groupings than they did when considering the event in a more immediate time frame (Liberman, Sagristano, & Trope, 2002). Still other measures show that people tend to think about future events stereotypically. Liberman et al. asked participants to list events that would constitute a "good" or "bad day," either 1 day or 1 year hence. The events listed for 1 year in the future tended to be less variable and more extreme, relative to those listed a day away. Similarly, in a different study, when participants were asked to describe how they would cope with daily events 1 year rather than 1 week into the future, participants described less variable coping strategies considering life 1 year ahead relative to 1 week ahead.

Basing predictions of the future on simple, abstract, and schematic simulations can produce errors, particularly mistakes about what one will prefer in the future. Liberman and Trope (1998), for example, asked college

students which of two class assignments they would like to complete in the future. One assignment had a desirable higher-order feature (its topic was romantic love) but an undesirable lower-order one (the paper to be read was not in the students' native language). The other assignment had an undesirable higher-order feature (its topic was the "attitude concept") but an attractive lower-order one (the paper to be read was in the student's native language). When the due date for the assignment lay 9 weeks in to the future, students preferred the assignment on the desirable topic. However, when the assignment was due within a week, students switched their preferences to the more mundane assignment that was easier to complete. Similar work has shown that when people consider gambling for money, their preferences in the distant future focus on the possible payoff of the gamble. However, as the time frame of those gambles approach, people give more weight to the concrete detail of whether the odds of winning are favorable (Sagristano, Trope, & Liberman, 2002).

People also show a tendency to lean on the simple and abstract when making predictions about other people in the distant future. When predicting another person's actions in the distant future rather than more immediately, people seek out more information about a person's global personality characteristics. People also predict that other people will act more consistently in the remote future than they will in the near future. They are also more likely to ignore important situational details when explaining another person's behavior in a distant time frame rather than explaining more immediate actions (Nussbaum, Trope, & Liberman, 2003).

Focusing on Central Outcomes and Neglecting Alternatives

In their simulations, people focus on the specific outcome in question and fail to pay equal attention to alternative ones. For example, if a student wonders whether she can get into Harvard Law School, she will likely focus on scenarios that would lead to her gaining admittance but fail to consider scenarios that would lead to her rejection. This focusing on central outcomes serves as the second subprinciple under the overarching principle of incomplete scenario building.

This neglect of alternatives has a long history in social psychology. For example, in their classic demonstration, Snyder and Swann (1978) showed that directing people to a focal outcome, such as asking participants whether another person was extroverted, prompted them to seek information that was consistent with that outcome over information that would be inconsistent, such as signs of whether the person was shy (for a review, see Snyder, 1992). As a consequence, the impressions people were left with were biased in the direction of the outcome considered.

In a similar vein, Dunning and Parpal (1989) asked students to consider whether purchasing prepared class notes would improve their course grades. Students stated that the notes would help quite a bit, unless their attention was focused on the alternative of not having the notes. When asked whether not having the notes would hurt their grade, students stated that the notes would not make much of a difference. Dunning and Parpal (see also Dunning & Madey, 1995) showed, with data, that students provided more muted predictions of the notes' impact because they were now simulating what they would do to compensate for their lack of notes. They could have simulated these compensatory efforts when asked the original question about whether the notes would help, but people tend to stay focused on simulating the event in question and neglect simulating the alternative.

Neglect of alternatives is also seen in research on support theory (Rottenstreich & Tversky, 1997; Tversky & Koehler, 1994). In one example, participants were asked to write down the last digit in their telephone number (from 0 to 9) and then to estimate the number of couples in the United States who had that number of children. Estimates for each number of children were then added together, and the resulting sum was 199%, suggesting that subjects had focused on scenarios that would produce the specific number of children they had been asked about but not about alternative numbers of children (Tverksy & Koehler, 1994, Study 2).

Other data more specifically showed that participants failed to consider concrete alternatives. Redelmeier, Koehler, Liberman, and Tversky (1995) gave doctors a detailed case study in which a woman complained of intense abdominal pain. They were then asked to judge the probability that the woman suffered from gastroenteritis, ectopic pregnancy, or some other malady. Doctors on average stated that there was a 50% chance that she suffered from some other malady. However, in a separate condition, when they were reminded of two other concrete diagnoses that lay in the "some other malady" category, their estimates rose to 69%.

This neglect of alternatives has been linked directly to overconfidence. Koriat, Lichtenstein, and Fischhoff (1980) asked college students to answer trivia questions and then to estimate the likelihood, for each answer they gave, that they had been right. They were also asked to provide any reasons suggesting their answer was right and any suggesting they were wrong. The confidence people expressed in their answers was more highly related to how many reasons participants cited supporting their answers than it was to the number of reasons contradicting their answers—suggesting that people were focusing on simulating why their answers were right rather than wrong. Of key importance, when students in a separate study were asked explicitly to write down reasons why an alternative answer might be correct, they were significantly less overconfident in the answers they gave. When students were explicitly asked to provide reasons supporting their answer, there was no change in their confidence levels, presumably because thinking of supportive reasons was a cognitive activity they were already spontaneously engaged in.

Neglect of alternatives may also explain why people tend to be too optimistic in their predictions. Positive events tend to be the focal event in people's predictions.

Predicting such events, they may concentrate on the forces that would produce these positive outcomes. If, instead, people were compelled to simulate the production of more negative alternatives, the predictions they make may become more pessimistic but also realistic. Evidence for this speculation comes from Hoch (1985), who asked students finishing business school to predict whether they would obtain high-salary jobs upon graduation. Students were unduly optimistic about their job prospects unless they were first asked explicitly to write down reasons why they might fail to achieve a lucrative job offer. In contrast, students asked to write down reasons why they would achieve a high-paying job were no different in their confidence level relative to a control condition. Again, this lack of an effect presumably was due to the fact that participants in the control condition had focused on these supportive reasons—but not contrary ones—without any prompting.

Alternative neglect may also explain why people express too much confidence in their judgments and predictions, whether those assessments are optimistic or pessimistic in nature. When college students were given case materials about hypothetical hospital patients who could be suffering from one of two diseases and symptom information about one of the diseases, their confidence increased or decreased depending on whether the symptoms were consistent with the disease. However, when given symptom information about the alternative disease the patient might be suffering, their confidence about the focal disease did not change (McKenzie, 1997; for similar results, see Arkes, Christensen, Lai, & Blumer, 1987; Koehler, 1994).

Focusing on the Optimistic and Neglecting the Pessimistic

In their scenarios, people also concentrate on the optimistic, whereas they underweight or simply disregard the pessimistic. This focusing on the optimistic serves as the third subprinciple of prediction under the overarching principle of incomplete scenario building. The planning fallacy arises, in part, because people generate upbeat scenarios about how they are going to complete future tasks. Newby-Clark, Ross, Buehler, Koehler, and Griffin (2000) asked college students to estimate how quickly they would complete a class assignment that was due within the next 3 weeks. The students were also asked to generate a realistic scenario for completion, a best-case scenario, and a worst-case one—estimating when they would complete the assignment under each scenario. The estimate based on the realistic scenario lay much closer to the one associated with the best case than it was the estimate related to the worst-case scenario. Estimates based on the realistic scenario also correlated much more highly with best-case scenario estimates than they were with worst-case ones. This relative neglect of the worst-case scenario arose even though participants were quite comfortable giving these pessimistic scenarios weight when predicting the completion times of others.

This predilection for accentuating the optimistic appears to be greater when predicting events falling in the distant future rather than the immediate. Eyal, Liberman, Trope, and Walther (2004) asked college students to discuss various changes in the ways final exams could be given in their class (e.g., changing to an honor system). If asked to discuss these changes 3 months before the final exam was to take place, students tended to dwell on positive reasons for the change rather than on reasons against the changes. However, this tendency reversed if asked 2 weeks before exams were scheduled; students focused less on the *pro* and more on the *con* of changing exam procedures.

People's goals also slide from the positive to the negative as events draw near. Pennington and Roese (2003) asked college students about their goals for an upcoming exam. When the exam lay 2 weeks ahead, participants were focused on optimistic goals, such as getting a high score, improving their grade-point average, and showing off their strengths. However, on the day of the exam, students were much more focused not on attaining the positive but on avoiding the negative, such as disappointing themselves and damaging their academic transcript.

A reliance on optimistic scenarios that dissipate over time might explain one commonly observed pattern of overconfidence that similarly evaporates as an event draws near. Gilovich and colleagues (1993) asked college students how well they were going to do on a college exam. On the first day of the semester, students on average thought they would perform at the 82nd percentile of the course—a result that obviously speaks of too much optimism. However, on the day of the exam itself, students more cautiously reported, on average, that they would perform at the 67th percentile. Again, this average estimate was too optimistic, but at least it was closer to an appropriate answer (which would be an average response centered on the 50th percentile).

Further data show that this temporal pattern in confidence estimates comes from the different types of simulations people perform for distant events relative to more immediate ones. Students in a follow-up study were asked to consider different experimental tasks they might be asked to complete. Some of the students were told that they were going to perform those tasks almost immediately, whereas others were told it would be a few weeks before they were brought back to the laboratory to confront the tasks. Students were then asked to rate how well they thought they would do in each task, as well as the factors that would promote or inhibit their performance. Students considering tasks far in the future concentrated on reasons why they expected superior rather than inferior performance. Students in the immediate condition were more mixed in the reasons they cited—causing them to be less certain about whether they would perform well (Gilovich et al., 1993, Study 4). Other data confirmed this pattern of mental simulation. People dwell on the positive when thinking about performances in the distant future but become more mixed when the time of that performance draws near. Indeed, forcing participants considering a distant event to think about it

negatively causes them to lower their confidence significantly (Sanna, 1999; Sanna & Meier, 2000).

Focusing on Distinctive Features of Events and Neglecting Shared Ones

In their scenarios, people also focus too narrowly on building a model of the central outcome in question, ignoring the impact that life in general will have. For example, suppose Larry was asked how quickly he would get his holiday shopping done this winter. To answer this question, Larry is likely to think about how he would go about shopping. He would think about whom he would have to buy gifts for, what types of gifts he would have to get, and what stores he would likely have to visit.

All these details are central to the focal outcome of *getting shopping done*, but they are not the exclusive set of details that matter. Many events that, on the face of it, have nothing to do with shopping might have everything to do with the length of time it takes Larry to get it done. Larry might catch the flu. His car might break down. He might have a big project to get done at the office before the end of the year. None of these events is central to holiday shopping. Rather, they are central to everyday life and would be shared across *any* type of outcome that Larry might be asked to predict, such as whether he would get his taxes done, be able to take that long vacation, or finish painting the house. These background circumstances are potentially important in determining when Larry would finish any major project he might consider, not just one coming around the holidays.

In short, people build simulations that are too narrowly focused on the events they are specifically asked about and fail to build a simulation factoring in how the background *sturm und drang* of everyday living, shared across numerous circumstances, might promote or hinder the outcome in question. That is, they simulate the event but forget that it may be crucial to also simulate "life." This habit has been referred to as *focalism* (Wilson, Wheatley, Meyers, & Gilbert, 2000) or the *focusing illusion* (Schkade & Kahneman, 1998). Focalism serves as the fourth subprinciple under the overarching principle of incomplete scenario building.

The impact of focalism has been most directly shown in predictions people make about their emotional reactions to future life events. Wilson and colleagues (2000) asked college students how much joy they would feel if their college football team won or how much despair they would feel if their college team lost. Their estimates tended to be too extreme. On average, students did not feel as happy as they predicted after a win, nor did they feel as distraught as they expected after a defeat.

Wilson and colleagues (2000) proposed that people mispredict these future feelings because of focalism. Students in their study concentrated too narrowly on the central event of the team's fate and failed to consider how the background of everyday life would tend to add its own little joys and frustrations into the mix, tempering the impact that the football team's fortunes might have on the student's mood. To test this notion, Wilson and colleagues asked roughly half of their respondents to write down the types of activities they would engage in during the few days after a football game. Reminded of all these other quotidian events, respondents provided much more muted predictions about how the football team's fortunes would affect their mood, relative to a control condition. Indeed, the predictions of this "defocused" group proved quite accurate, relative to the more extreme—and erroneous—predictions made by the control group.

Focalism in prediction has been shown in many variants. People simulate the impact of an event for them personally but often fail to take into account that the event might have an equivalent impact on other people. For example, when students are told that 10 points might be added to everyone's test grade in a course whose grades are "curved," they report being quite happy about the prospect. Of course, if everyone's grade is being raised in a course using curved grade guidelines, then the increase has no impact on anyone's grade. Students, however, tend not to recognize this fact, focusing too much on how the 10 points will increase their own grade but failing to "simulate" that this event will have an equal impact on everyone else (Windshitl, Kruger, & Simms, 2003; see also Moore & Kim, 2003).

Similarly, students playing poker bet more after they are told that there will now be a number of "wild cards" in the deck. In a sense, students now act as if their chance of winning has increased—but, of course, it has not. Everyone sitting at a table, on average, is helped to the same degree by the presence of wild cards. People, however, tend to miss this fact, simulating how the wild cards might help them personally gain better hands but neglecting to simulate the "background," that is, how the wild cards will help everyone else as well (Windshitl et al., 2003).

Focalism also arises when people explicitly compare two events against each other, leading to mispredictions. When comparing two events, people focus disproportionately on situational features that distinguish those events, thus giving those distinctive features great weight in any future prediction—often, too much weight. For example, U.S. respondents tend to believe that people living in California are happier and more satisfied with life than people living in the Midwest. Comparing the two locales, it is easy to spot the basis for this prediction: California, for example, has terrific year-round weather. The Midwest does not. If good weather is one component to happiness, then California easily beats the Midwest on this score.

However, basing a prediction of happiness predominantly on weather is focusing too much on what distinguishes California from the midsection of the United States. There are many other factors—common to both locations—that influence well-being much more than the weather, such as one's health, social relations, and traffic on the way to work, among many other variables (Diener & Seligman, 2004). Neglecting these shared variables in a prediction about California leads to a prediction that is too extreme (Schkade & Kahneman, 1998). In a similar vein, college students comparing different dormitories overweight the distinctive features of those dorms, such

as their location, in their predictions of satisfaction, whereas they underweight the impact of features that tend not to distinguish dorms, such as social life (Dunn, Wilson, & Gilbert, 2003).

Focusing on the Strength of Evidence and Neglecting Its Weight

In their scenario building, people also neglect one important aspect of the evidence they survey. They collate evidence in favor of a prediction (and, on occasion, evidence against that prediction) and then examine how well that evidence suggests that prediction. However, they often fail to ask how reliable or valid the evidence before them is. For example, if Harry is politically liberal, loves books, and is not concerned about money, it is easy to predict that he is more likely to be a humanities than an economic major, presuming for the moment that those characteristics and preferences do, in fact, predict preferences in college major. However, what if those pieces of evidence are unreliable, changeable, or outright wrong? Perhaps it is not true that Harry is politically liberal. Perhaps he loved one book this week but will rarely pick up another book again. In short, if the evidence on which people base their predictions is unreliable or mistaken, their predictions will be wrong no matter how valid their prediction model is.

A growing body of evidence shows that people place too much emphasis on the *strength* of evidence, that is, how much the evidence suggests one outcome over another, but not enough emphasis on the *weight* they should place on that evidence, that is, the extent to which the evidence is likely to be valid or reliable. This is akin to judging that a young scholar will be a successful one because of a strong letter of recommendation written by a mentor (strength), without calculating whether the letter writer tends to be a credible judge of another person's potential. This issue goes to the weight one should place on the evidence cited in the letter (Griffin & Tversky, 1992). This emphasis on strength and neglect of weight serves as the fifth subprinciple under the overarching principle of incomplete scenario building.

The following example, modeled after Griffin and Tversky (1992), illustrates one way in which people pay too much attention to the strength but not to the weight of evidence. Suppose one possessed a coin that, if flipped, would have a slight tendency (say, 60%) to come up either heads or tails. However, which way the bias goes is unknown. Now suppose the coin is flipped three times and comes up heads all three times. This is evidence that is strong on strength (i.e., all flips point to heads) but low on weight (i.e., there are only three pieces of evidence). In situations such as these, people are roughly 85% confident that that coin is biased toward heads, although the actual likelihood of bias, based on a Bayesian analysis, is that the coin is only 77% likely to be heads biased. In this case, overemphasizing strength and underemphasizing weight leads to a prediction that is overconfident.

Now, consider a second case in which the coin is flipped seventeen times and comes up heads 11 times.

This is a circumstance that possesses low strength, in that the coin turns up heads only around 67% of the time, but high weight, in that the result is based on a total of 17 flips. In circumstances such as these, people tend to be roughly 65% certain that the coin is heads biased, but the actual likelihood via a Bayesian analysis is 88%. The advantage for heads in this case may be slight, but it is based on a relatively large number of flips, and thus the advantage is more likely to represent the coin's true bias. However, if people focus primarily on the strength (e.g., the extremity) of the evidence but not on the weight they should give it, they may end up not confident enough in their prediction that the coin is biased toward heads.

Other work shows the same sort of prediction biases. When strength of evidence is high but the weight given to it should be low, people tend to be overconfident. When the strength of evidence is low but the weight it should be given is high, people tend to be underconfident. College students and business managers, for example, showed this pattern when trying to ascertain whether a piece of market prediction software is accurate or biased (Jiang, Muhanna, & Pick, 1996). In a similar vein, MBA students trading securities in an experimental market also relied too much on the strength of evidence of whether a security was a good one and too little about whether the evidence they were using was reliable. As a consequence, market prices for securities tended to be too high when those prices were propped up by unreliable information that was, nonetheless, extreme. Traders spent too much money for such securities—and too little for securities that had less strong evidence in their favor but evidence that should have been given great weight (Nelson, Bloomfield, Hales, & Libby, 2001). Traders selling high-strength/low-weight securities, or buying low-strength/high-weight securities, tended to profit much more handsomely than peers doing the opposite.

Summary

In sum, one major difficulty people have in their predictions is that their scenario-building efforts tend to be incomplete. They tend to stay at a higher-level, abstract level in their simulations and fail to simulate important concrete details, particularly when making predictions about distant events. They concentrate on focal outcomes and fail to build full simulations about alternative ones. They simulate optimistic scenarios but then dismiss pessimistic ones. They suffer from focalism in all its variants, tending to simulate the influence of central events without considering the impact of normal, everyday background events. Finally, people respond more to strength of evidence in their simulations while neglecting the weight they should assign to those events.

All these habits explain why prediction tends to be imperfect. If people do not simulate all the facets of the future that are relevant, they will be left with an incomplete and flawed understanding of what the future might hold. That said, incomplete scenario building may not be the primary reason why people's predictions tend to be biased or wrong.

PRINCIPLES OF PROCESS: III. THE LIMITED UTILITY OF SCENARIO BUILDING

The primary reason why people's predictions might be imperfect is that it is impossible to build comprehensive and accurate scenarios of what the future holds. Although the strategy of scenario building may often lead to predictions that are adequate and accurate, it suffers from severe limits to its utility. Scenario building, to put it bluntly, might be somewhat of a fool's errand. It may mislead because people often are just not in a position to simulate all the possible scenarios that may unfold in the future. They simply do not have all the information they need to assess the plausibility and likelihood of future events, and this observation serves as a second overarching principle in process regarding the task of prediction. Two strands of research demonstrate how people cannot be expected to anticipate the future because they do not possess, nor can be expected to possess, all the information they need to do so. One strand focuses on the power of situational details that are often unknowable. The second has to do with how well people predict the impact of emotion on their behavior.

Unknown and Unknowable Situational Details

A central insight developed from decades of social psychological research is that how an event unfolds depends heavily on the exact details and features of that situation. Social psychology is filled with examples of people's behaviors being significantly influenced by small changes in situational circumstances (Ross & Nisbett, 1991). For example, whether someone offers to help someone else in need depends on any number of seemingly small or transient factors that are just not knowable until the moment the situation arises, such as how much time people have on their hands to help (Darley & Batson, 1973), how attractive that other person is (Benson, Karabenick, & Lerner, 1976), how many other people are around (Latane & Darley, 1969), and what mood the person is in (Isen, Clark, & Schwartz, 1976). The problem for anyone making a prediction about whether someone will help in any particular situation is that which of these small details will obtain are just unknowable until the relevant moment arises, and so any scenario an individual might build to predict it will be imperfect, at best. This observation about the unknowability of situational details serves as a first subprinciple under the overarching principle that people often cannot expect to have complete knowledge about future circumstances.

People fail to recognize this fact as they make their predictions. They commonly build only a single scenario, assuming that this scenario contains all the complete information necessary to make a prediction (Dougherty et al., 1997; Klauer, Musch, & Naumer, 2000). Or, if they do consider multiple scenarios, they then dismiss all the scenarios they have built except for the one or the few they consider especially likely (Dougherty et al., 1997). In following this strategy, they fail to consider two important facts. The first is that whichever scenario or scenarios they are considering need not be the one that will actually arise in the future—and that they should consider as many alternative scenarios as they can plausibly generate. The second is that generating as many alternative scenarios as possible is, in the end, a doomed enterprise. If people try to conjure as many scenarios as possible, they will never know with certainty if they have generated all the ones they should. Given this, they should just concede the fact that they do not know how the situation should turn out, and so they should just go ahead and generate a prediction but be hesitant in the confidence they assign to it (Griffin & Ross, 1991).

Griffin, Dunning, and Ross (1990) demonstrated that people predicate their predictions on somewhat specific scenarios that they assume are complete and true. Griffin and colleagues asked participants to make predictions about future events, such as how much they would stay on the phone if called by a survey taker. Participants provided an estimate and then placed an upper and lower bound around that estimate so that they were 50% sure that the actual value would be inside those bounds (and feel 50% certain that the actual value would lie outside those bounds). Participants were then introduced to one of four different experimental conditions. In the control condition, participants were simply asked to reconsider their predictions and to provide once again a best guess, an upper bound, and a lower bound. Not surprisingly, predictions tended not to change. Participants expressed roughly the same level of confidence as in their first prediction, in that the upper and lower bounds they placed around their estimates were only 5% tighter.

In a second condition, participants were asked to describe the features of the situation as they had imagined them before reconsidering their predictions. Participants tended to describe only one scenario, provide roughly the same predictions again, with upper and lower bounds that expanded by only 3%. In a third condition, participants were also asked to describe the details of the scenario and then to assume that those details were *exactly* how the situation would develop. Participants again mostly reiterated their previous predictions and held roughly the same confidence in them, constricting the upper and lower bounds by only 3%.

Taken together, the responses of participants in these three conditions suggest that those in the control condition were already doing what participants in the second and third condition were asked to do explicitly. Participants in the control condition, before making a prediction, were constructing a scenario and then assuming its completeness and truthfulness. By doing so, they revealed little awareness that the future might unfold in very different ways. As a consequence, the upper and lower bounds they placed around their predictions—representing the degree of uncertainty they expressed over their predictions—stayed roughly the same. What participants seemed not to do spontaneously was to consider multiple scenarios, or to recognize that the specific way the future would develop was unknowable and so they should be relatively uncertain about any prediction they made.

Bolstering this analysis were participants' responses in a fourth condition in which they were explicitly asked to

consider alternative scenarios and reminded, in reality, that they did not know how the future would unfold. Put through this exercise, participants expanded their upper and lower bounds by 38%—expressing much less certainty about the predictions they provided. A follow-up study replicated this result: Participants expressed more uncertainty in their predictions after furnishing two scenarios of the future situation surrounding a prediction. To the extent that those two situations differed in their detail and the prediction they suggested, participants expressed more uncertainty about what would take place (Griffin et al., 1990, Study 3).

Inaccessibility of Emotions and Their Impact

Other features of future situations also appear to be unavailable to individuals as they build their prediction scenarios. People are obviously emotional beings. Each month, many individuals pass through episodes of happiness, sadness, boredom, fear, surprise, anxiety, and elation, just to name a few of the emotional states people may find themselves in. Despite lifetimes of experience with the impact of emotion on their thoughts and actions, people do not demonstrate an adequate ability to anticipate how similar emotional states will influence them in the future. People mispredict the intensity of their emotional reactions to future events, as well as the duration of those emotions. They also fail to anticipate how emotionally arousing situations will influence their behaviors and preferences. In essence, if not in the relevant emotional state at the time they make their prediction, people fail in their attempts to simulate how that emotion will influence their reactions to events. Without these accurate simulations, the scenarios on which they base their predictions tend to be mistaken. This observation serves as a second subprinciple under the overarching theme that people often cannot expect to have complete knowledge about future circumstances.

Overestimating Impact of Events on Emotion

The first type of error that people make is that they overestimate the impact of events on their emotional states, in terms of both intensity and duration (Wilson & Gilbert, 2003). Regarding intensity, people believe that just missing their subway train will produce ample levels of regret, but this is not actually the case (Gilbert, Morewedge, Risen, & Wilson, 2004). People believe that Christmas will give them much more joy than they actually experience. Students at the moment they receive an exam grade that is worse than they expected experience fewer negative emotions than they predicted they would a few weeks before; students getting a higher grade than they expected do not feel as elated as they thought they would (Buehler & McFarland, 2001).

People also overestimate how long the emotional impact of an event will last. Young assistant professors overestimate how long their happiness will be damaged by being denied tenure. Voters overestimate how long they will be happy if their gubernatorial or presidential candidate wins or how sorrowful they will be if their candidate loses. Young lovers overpredict how long their well-being will be injured by a romantic breakup (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998). Sports fans overestimate how long their mood will be influenced by wins or losses of a favorite team (Wilson et al., 2000). Health clinic patients overestimate how long they will be distressed if told they are HIV positive, as well as how long they will feel relieved if told they are HIV negative (Sieff, Dawes, & Loewenstein, 1999).

People overestimate the impact of emotions for many reasons (Wilson & Gilbert, 2003), but one particular aspect of emotions that people seem unable to simulate is how such emotions spur them to action. Emotions energize people, and negative emotions energize specific efforts to dispel those emotions. After failure or disappointment, people are quite skilled at finding ways to rationalize the experience away (Dunning, 2001; Kunda, 1990). However, in the scenarios that people build about the future, they seem not to recognize that a negative event will invigorate these rationalization efforts.

For example, Gilbert and colleagues (1998) interviewed college students putatively for a very lucrative short-term job. For some of the students, the decision about whether they were suitable for the job was made by a single, perhaps idiosyncratic, MBA student. For others, the decision was made collectively by a panel of MBA students. Students asked hypothetically about how they would react to being rejected for the job thought they would feel relatively bad regardless of whether the individual or the panel rejected them. However, those actually rejected by the panel felt significantly worse than those rejected by the lone individual. Presumably, students found it easier to rationalize away the judgment of the single individual than that of the panel, but students making predictions about how they would react did not demonstrate any insight into this possibility.

Further research shows that people fail to recognize how negative events spur efforts toward rationalization that ameliorate negative emotional states (see Gilbert & Ebert, 2002), leading them to reach decisions that potentially could cause them to be less happy in the long run. For example, students facing potential rejection for a date by a person of the opposite sex opt to take higher doses of a drug aimed at alleviating psychic pain. However, the pain they anticipate is often dispelled effectively by a level of rationalizing expertise that people appear not to know they have (Wilson, Wheatley, Kurtz, Dunn, & Gilbert, 2004).

Underestimating Impact of Emotions on Action

The inability to recognize how emotion activates rationalization stands as a special case of a more general misunderstanding about emotion. In their simulations of future events, people fail to grasp how much emotions will alter what they pay attention to in a situation and what behaviors they will display and preferences they will express. That is, people may overestimate the intensity and duration of emotions they will feel, but they underestimate just how much that emotion, even somewhat less-

ened, will still have a significant impact on thought and behavior.

Emotions have been shown to have a significant impact on what people pay attention to in the landscape of a situation in front of them. Emotional arousal causes people to pay more attention to stimuli in the environment that are evaluatively valenced (Halberstadt & Niedenthal, 1997; Niedenthal, Halberstadt, & Innes-Ker, 1999), or to any stimuli that might explain the emotion they are feeling (Dutton & Aron, 1974; Valins, 1966). Emotion also predisposes people to reach different assumptions about the meaning of events, such as whether they are in control of them (Lerner & Keltner, 2001).

Thus, in one emotional state, people may view a situation differently from the way they would in a different emotional state—and people appear capable of understanding only imperfectly how their thoughts and actions might change when going from one state to the next. Instead, when they make a prediction, they appear to predicate their predication on the way they view a situation currently, even when they know "intellectually" that their emotional state may be different in a future circumstance. They fail to correct for changes in emotional state, instead projecting their current view of a situation onto the future one in question. For example, Van Boven and Loewenstein (2003) approached people either just before or just after they had worked out in a gym, asking them how much they would be bothered by thirst if they took a hike in the mountains. Respondents reported that they would be more bothered by thirst if they had just finished their workout (and likely to be thirsty themselves) than if they had yet to start. In essence, people had projected their current evaluation of a thirst-quenching drink onto a very different situation.

Other researchers have shown that people, in one emotional state, fail to adequately simulate how their preferences would be different in a dissimilar emotional state. Read and van Leeuwen (1998) approached office workers just after lunch with a promise to return 1 week later with a free snack, either again just after lunch or instead late in the afternoon. Respondents were asked which snack they preferred, a candy bar or a piece of fruit. Respondents realized that they would be hungrier if the experimenters returned late in the afternoon and so were more likely to choose the candy bar over the fruit if the experimenter said he or she would return at 4:00 P.M.

However, respondents' current state of being full from a recent lunch introduced some bias into their decisions. Relative to a group approached with the same proposition at 4:00 P.M., respondents choosing just after lunch were significantly more likely to choose the fruit over the candy bar, regardless of when the experimenters promised to return. Current hunger states had been projected into future hunger states—even when respondents knew at some intellectual level that their state of hunger would probably be different. This tendency has been demonstrated elsewhere and is exacerbated if people are not given time to think through their preferences (Gilbert, Gill, & Wilson, 2002).

Other research has shown how people inadequately anticipate how emotions change the landscape of thought and action. In one such study, roughly half of the college students in a large lecture class were asked, hypothetically, whether they would volunteer to go up to the front of the class to dance to a classic funk tune for $5. Most of the students said they would not, but on average said they would do it if they were paid roughly $21. However, a hypothetical decision to dance in front of hundreds of people is not very anxiety inducing. Being asked for real is—but respondents could not simulate how the emotion of potential embarrassment would change their perception of the situation and influence their decisions to volunteer. When the other half of college students in the room were given an actual opportunity to dance in front of the class for $5, only 8% did so. On average, respondents in this group said they would need to be paid $53 to dance (Van Boven, Loewenstein, & Dunning, 2004; Van Boven, Loewenstein, Welch, & Dunning, 2004).

Summary

Recent research on the psychology of prediction suggests that prediction would be difficult—and predictions open to error—even if people built more complete and thoughtful scenarios about how the future might unfold. People simply do not have access to all the information they need to produce perfect predictions. They often do not know, nor cannot hope to know, all the important situational details they need to know to furnish accurate predictions. In addition, they do not have cognitive access to how emotional layers to situations will influence their thoughts, preferences, and actions. As a consequence, they tend to predict that they will behave in the future the way they prefer to behave now, not able to anticipate how emotions change the landscape of what people think about and how they react to emotion-laden situations.

AFTERTHOUGHTS: REVERSALS TOWARD PESSIMISM AND UNDERCONFIDENCE

Up to this point, this chapter has reviewed work on the two general tendencies researchers tend to see in prediction—undue optimism and overconfidence. However, the astute reader of behavioral and cognitive sciences knows that tendencies are just that—what people tend to do and not what they always do. These inclinations toward optimism and overconfidence may be common, but they are hardly universal. At times, people show the reverse, being too pessimistic or too underconfident in their predictions.

What can be said about these reversals of these general tendencies? To date in the literature, much less has been said about undue pessimism and underconfidence relative to their more common counterparts. However, some fragments of themes have emerged to anticipate when people will be too glum or too unsure about what the future holds.

Undue Pessimism

One central principle guiding this chapter's treatment of prediction is that people often rely on a single scenario when they assess what the future might hold. Often, that scenario is a positive one, for people are awash in a sea of positive information or beliefs that they can use to build a prediction on (Matlin & Stang, 1978). That positive information might influence their guiding scenario by suggesting a tentative frame or hypothesis for the individual to start with. It is on this initial positive frame that people elaborate their predictive scenarios.

But what if the most prominent or bulk of information people had regarding a prediction suggests a negative frame? For example, if people were asked whether they would ever own a horse that won the Kentucky Derby, most would quickly assume that they would not—for the bulk of information they have accessible to them suggests a quick hypothesis that runs in a negative direction (e.g., I do not know enough to buy the right horse or have enough money to buy that horse with!).

One can conjecture that if people quickly come up with tentative hypotheses that are pessimistic, it would be a circumstance in which their predictions would tend to be *too* pessimistic—for many of the same psychological processes enumerated previously that support undue optimism could also just as equally support undue pessimism. For example, recall that people tend to suffer from focalism when predicting how they will perform relative to their peers in a competition—focusing on how some advantage will benefit them greatly even though the advantage is shared by all (Moore & Kim, 2003; Windshitl et al., 2003). This focusing on the self also occurs when people consider disadvantages, with people thinking that the disadvantage will hurt them disproportionately. For example, negotiators tend to think that time pressure and deadlines disadvantage them more than they do the people with whom they are negotiating, even though the same deadline exposes both sides to the same burdens and complications (Moore, 2004).

One can also see how many of the other processes discussed in this chapter can support pessimistic predictions if people's prediction scenarios take a negative turn. If people start from a negative frame—neglect alternative scenarios, focus on the abstract and neglect concrete details of future situations, forget about the mitigating nature of everyday life, fail to consider the weight they should assign to the evidence they have, and ignore the fact that they often just do not have access to how future circumstances will specifically unfold—they are likely to make predictions that are overly glum. The destination of a prediction may be a reversal of the general case, but the cognitive route to that destination may be paved in much the same way as are overly optimistic forecasts.

Underconfidence

Circumstances that lead to underconfidence are somewhat better delineated in the psychological literature. People tend to be overconfident, but if the prediction or task in front of them is quite easy, they seem not to be sensitive to that fact—thus imbuing their predictions with too little confidence, not too much (Gigerenzer, Hoffrage, & Kleinbolting, 1991; Lichtenstein & Fischhoff, 1977; Suantak, Bolger, & Ferrell, 1996). This underconfidence in the face of easy tasks has elicited much empirical and theoretical attention, and several explanations exist for it. Some explanations focus on statistical or artifactual accounts—for example, one can explain underconfidence for easy tasks by simply noting that it is difficult to be overconfident on a task one is getting right nearly all the time (see Juslin, Winman, & Olsson, 2000, for a review of statistical explanations for the hard–easy task effect on over- and underconfidence).

But, again, one can also imagine that the same psychological mechanisms that produce overconfidence in general also work to produce underconfidence when tasks are easy. For example, if people ignore the weight they should give to the evidence in front of them and, instead, concentrate solely on its strength, it would produce underconfidence when evidence is reliable and the weight that should be given to it is great (Griffin & Tversky, 1992).

The same goes for other indicators of accuracy. In adopting an inside view and focusing on scenario building, people tend to ignore past experience and the "outside" data that this experience provides. Sometimes that past experience suggests a decisive prediction; if an event has happened frequently in the past it should be confidently predicted to occur in the future. Thus, if one has bought at least one soda at some point in each week in the past, one should be confident that one will buy at least one in the week coming up. If one can never get one's coworker on the phone during lunchtime, one should confidently predict that calling that coworker at 12:15 P.M. will not work today. However, if people ignore such valuable data when making predictions, they will be underconfident in their predictions when those data clearly suggest a compelling conclusion (Dunning et al., 1990; Vallone et al., 1990). If they are overconfident when they go against decisive data, they may prove underconfident when their predictions are consistent with the data.

Individual Differences—And the Lessons They Suggest

There are also individual differences in the tendency to be optimistic. One particularly telling individual difference centers on *defensive pessimism* (Norem & Cantor, 1986). Defensive pessimists tend to be relatively anxious individuals who deal with the anxiety of upcoming challenges, such as a test in a college course, by setting low expectations about themselves and then running through all the various negative outcomes they might achieve. This strategy can be compared to that of *strategic optimists*, who deal with the same anxiety by making rosy forecasts of the future and then avoiding any contemplation of what might come next. Although defensive pessimists predict more doom and despair than do their more

optimistic counterparts, their level of actual achievement tends to be equivalent once the challenge arrives (for a review, see Norem & Chang, 2002).

However, beyond demonstrating that there are stable individual differences in optimism and pessimism, the literature on defensive pessimism suggests one potentially important facet of the psychology of prediction that, as of yet, has not perhaps been given the attention it richly deserves. So far in this chapter, I have discussed how people reach the predictions they make but have not discussed *why* people decide to make a prediction and choose the particular forecast they decide to endorse. The guiding assumption underlying this chapter is that people have one simple goal in making predictions—they wish to accurately anticipate the future and so their goal is to reach a prediction that will prove correct. People harbor just one motive—the one to be accurate—even if their cognitive habits often circumvent them from achieving that goal.

Tellingly, the habits of defensive pessimists, however, seem to be aimed at achieving a goal other than accuracy—that of reducing the level of anxiety they typically experience (Norem & Cantor, 1986). They predict doom and gloom not necessarily because they believe that is how the future will unfold; they predict doom and gloom because it makes them feel better. Similarly, strategic optimists predict rosy futures to achieve the same aim of reducing any apprehension and disquiet they may feel (Norem & Chang, 2002).

Given this reason, one may wonder what other motivations attach to the act of prediction that having little or nothing to do with the motivation to be accurate. As of yet, no exhaustive catalog of motivations exist, but other research findings affirm the notion that people make predictions for reasons that lie outside the goal of accuracy. For example, people tend to be less optimistic (and more realistic) in their predictions before they reach a final decision about how to act in a given situation. However, once they have committed to a course of action, they become more optimistic and sanguine in their predictions, especially concerning how their actions will lead to success (Gollwitzer & Kinney, 1989; Puca, 2004; Taylor & Gollwitzer, 1995).

This difference stands to reason if one presumes that the act of prediction is motivated by different goals before and after people have decided on a course of action. Before deciding, people are in a *deliberative mindset* in which they wish to accurately weigh which decision is the best to make. This goal pulls for realism and impartiality. However, after deciding, people are no longer interested in accuracy but are, instead, in an *implementational mindset* that aims at energizing themselves to succeed. In this situation, optimistic predictions are more motivating than realistic ones, and so people craft predictions that serve this goal toward fostering a happy outcome (Gollwitzer, 1990).

In sum, work on individual differences suggests that variations in goals might be the key factor that lead people to be overly optimistic versus pessimistic in the predictions they make. Such differences in goals may also influence whether people chronically appear to be over- or underconfident in their predictions—and recent research does suggest that people do differ reliably in the level of confidence they express about their predictions (West & Stanovich, 1997). However, to date, the relations among individual differences, goals, optimism, and confidence have yet to be exhaustively explored, and so remains a topic of tremendous potential for research.

PRINCIPLES OF IMPROVEMENT: ENHANCING PREDICTIVE ACCURACY

The preceding sections enumerate habits that typically lead people to make erroneous predictions, or at least predictions that are too optimistic and held with too much confidence. Given these problems, what is a person to do to improve his or her prediction-making ability? What habits should people adopt to form more perfect forecasts? What principles of improvement should they adopt?

The research literature contains many suggestions for improving predictive accuracy. First, one can construct a more comprehensive scenario-building analysis. Instead of focusing on optimistic scenarios, one should consciously set out to build pessimistic scenarios and then give them weight. Instead of focusing on circumstances that might produce a focal event (e.g., a company to succeed with its new product), one should also focus on conditions that might promote alternative outcomes (e.g., failing with the new product). Instead of succumbing to focalism, and just concentrating on features of situations that are relevant to an event, one could take deliberate pains to simulate how the normal hurly-burly of everyday life might affect one's feelings, preferences, and actions.

These steps may alleviate some predictive error and overconfidence. However, as some previously discussed material suggests, the details of the future are always going to be unknown and unknowable, leading any scenario-building strategy to be incomplete and, thus, imperfect. Given this, what is one to do?

Adopting the Outside View

The literature suggests that one primary way to improve predictive accuracy—that is, one principle of improvement—is to set aside scenario building and adopt a completely different strategy. Instead of taking the "inside view," focusing on the individual case to be predicted, analyzing its internal dynamics and assessing how certain outcomes are potentially caused, one should instead take an "outside view." Taking an outside view means recognizing that a particular situation is one example of a class of similar events and then surveying previous situations in this class to tally which outcomes have commonly followed (Kahneman & Lovallo, 1993; Kahneman & Tversky, 1982; Lagnado & Sloman, 2004; Lovallo & Kahneman, 2003). For example, an inside approach to predicting when one will finish one's tax returns is to enumerate the steps one has to take to com-

plete the task and to list the obstacles that might fall in one's way. This approach focuses on the internal dynamics of completing one's tax return and fails to reference similar situations and tasks in the past. Adopting an outside view pushes aside these efforts at mental simulation to instead collate data. One predicts when the tax return will be done by counting up when one has completed one's taxes in the past—or even by tallying when other people typically get their taxes done.

The extant literature suggests that people reach more accurate predictions when they pursue an outside approach to prediction rather than an inside one. This is akin to an experience related by Lovallo and Kahneman (2003), who described a group of education professionals working to revise a school system's curriculum. When each member of the group was asked to predict when the group would finish its work, the most pessimistic prediction was between 2 and 3 years. However, when group members were asked explicitly to adopt an outside view and to consider their experiences with similar groups in the past, one group member with extensive experience with curriculum revision conceded that similar groups took at least 7 years to complete their mandate, if they finished their task at all. Not surprisingly, this particular group finished its specific task 8 years later.

More formal research shows that people provide more accurate predictions when they adopt an outside view. For example, in a study on the planning fallacy, Buehler and colleagues (1994) asked college students to complete an hour-long computer tutorial program aimed at teaching them basic principles in psychology. When participants in a control condition were asked to predict when they would complete the tutorial, they tended to underestimate how much time it would take by over a day, and only 29% completed the tutorial by the date and time they predicted. However, when participants were asked to list when they had completed similar assignments in the past (and then reminded that past experiences are relevant to predicting current ones), they showed no tendency to underestimate the amount of time it would take them to finish the tutorial. Indeed, 60% of participants in this condition completed the assignment by the date and time they said they would.

Similar findings have been found in more real-world settings. Software developers provide more accurate estimates of when new software products will be completed if they can recall similar software projects from the past and use those past experiences in their estimates for calculating the time to complete current or future projects (Jorgensen, 2004a, 2004b). Ostrom and Shrauger (1986) found that college students who used "personal base rates," that is, they referred to the frequency with which they performed behaviors in the past, were more accurate in their predictions of future behavior (such as whether would fall in love or change their hairstyle), especially when they considered behaviors with extremely high or extremely low base rates.

Other data suggest that individuals achieve higher accuracy and avoid overconfidence when they use the behavior of other people as a guide to their own. College students make more accurate predictions about future behavior when they predict that they perform behaviors with high base rates than when they make predictions "going against" the base rate (Dunning & Story, 1991; Shrauger, Mariano, & Walter, 1998; Vallone et al., 1990). Curiously, although people appear to have accurate impressions of the base rates of behavior among their peers (Epley & Dunning, 2000, 2004; Nisbett & Kunda, 1985), they do not seem to give this knowledge much weight when they make predictions about individual people, whether it be themselves or a particular other person (Dunning et al., 1990; Epley & Dunning, 2000). As a consequence, they make erroneous predictions and suffer from overconfidence (Dunning et al., 1990; Vallone et al., 1990).

Cognitive Repairs

A second corrective habit is potentially not as radical. Instead of throwing away scenario-based predictions *in toto*, an alternative strategy would be to recognize that scenario-derived predictions carry some validity—they just have to be adjusted for any systematic error these predictions contain. For example, research on the planning fallacy shows that people systematically underestimate the amount of time it will take them to complete the projects they undertake, but people's predictions do contain information, and those predictions correlate significantly with their actual prediction times—with such correlations sometimes reaching as high as .75 (e.g., Buehler et al., 1994). That is, although people show a systematic bias in their estimates, their estimates do accurately discriminate between those who will take less time to complete a task and who will take more.

A similar pattern of error and accuracy arises for the prediction of socially desirable behavior. Overall, people significantly overpredict the likelihood that they will perform such desirable acts as voting in an election and maintaining their current romantic relationship. Despite this overall error, people's predictions carry some information: People who predict they will vote actually do so 75% of the time; those who predict that they will not, vote only 20% of the time. Once again, although people show *bias* overall in their predictions, in that they overpredict the likelihood of performing a desirable behavior, their predictions are not useless—in that their predictions exhibit *discrimination*. Their predictions can be used to discriminate who is the most likely to perform a behavior and who is not (Epley & Dunning, 2004).

The potentially useful principle of improvement to use in prediction would be to ask people for their predictions, which are likely to carry some discriminative information, and then to apply a *cognitive repair* to those predictions to remove the systematic bias that usually contaminates such predictions. One such cognitive repair would be to "dial down" the amount of optimism displayed in a prediction. If a person thinks she is 80% likely to vote, a cognitive repair would be to lower that likelihood to 70% or 60%. Such cognitive repairs are commonly practiced in the business world. For example,

Microsoft Corporation commonly asks its software developers to estimate how long it will take them to finish a new product and then takes that prediction and inflates it automatically by 30–50% if the product is a new operating system. This "repaired" or corrected prediction is the one the company uses in its planning. Similarly, structural engineers are typically asked to estimate how much concrete a structure will need to remain stable. That prediction is then taken and multiplied by a factor between 3 and 8 to guard against overconfidence (Heath, Larrick, & Klayman, 1998).

Aggregating Predictions

A good deal of psychological, economic, and sociological research also suggests that there is wisdom in numbers, in that an aggregated prediction from a group of individuals often outperforms the prediction of any individual. In 1924, Gordon asked 200 college students to rank-order the weights of several everyday objects. She then averaged the ranks given by the students and found that these average ranks predicted far better the weight of the objects than did the ranks given by the typical individual student. Indeed, the average rank given by the group outperformed the estimates of all but 5 of the 200 students in the study.

Since this classic study, research has discovered that aggregating people's judgments works as well for predictions of the future as it does for estimates of weight. Aggregated forecasts outperform individual ones, and often are almost as accurate if not as accurate as the best-performing individual group member (Clemens, 1989; Einhorn, Hogarth, & Klempner, 1977; Hill, 1982; Yaniv, 2004). This finding holds true in domains as diverse as mental health professionals predicting the outset of mental disease (Goldberg, 1965), economists forecasting turns in the economy (Zarnowitz, 1984), and business executives and sales managers predicting future advertising revenue (Ashton, 1986). Not many forecasts have to be aggregated to achieve the full benefit of averaging people's predictions together (Yaniv, 2004). For example, Libby and Blashfield (1978) discovered that the maximum benefit of aggregation was obtained by combining the forecasts of only 3 to 6 experts, although Hogarth (1978) suggested that up to 20 individuals should be consulted, especially if people differed in their estimates.

That, however, is not all. When individual group members are asked to indicate their confidence in their predictions, and these confidence estimates are aggregated, the level of confidence expressed at the aggregate level tends to match the level of accuracy achieved by the group quite well (Plous, 1995). That is, predictions at the aggregate level tend not to be overconfident—or at least not as overconfident as the predictions made by the individuals whose judgments are used to construct the aggregate. This serves as a third principle of improvement.

Aggregate predictions provide superior forecasts to individual predictions because of simple statistical verities. As discussed throughout this chapter, the predictions that individuals provide may be incompletely thought out and imperfectly rendered, but they do tend to bear some relationship, sometimes tenuous, to the truth. The rest of the prediction is made up of error caused by mistakes in analysis, incomplete or biased information, and important data that are simply unknowable. If a group of individuals tends to make predictions that cluster somewhat around the truth, those predictions will tend to diverge from each other and the truth because of forces that produce error. Aggregating those predictions will retain the part of the prediction that is related to the truth while getting rid of some of the individual errors, leading to a more accurate prediction (Einhorn et al., 1977). This analysis, however, depends on two assumptions. First, individual predictions must bear some relation to the truth. Second, the errors that people make must diverge from one another. If all individuals in a group make an error in the same direction (e.g., such as government officials overestimating the robustness of the economy), then the aggregate will retain this bias and fail to prove more accurate (Einhorn et al., 1977).

In sum, a long history of research suggests that the best way to improve one's predictions is to seek out the predictions of other people and aggregate them with one's own (Yaniv, 2004). This research, in addition, demonstrates the imprudence of two closely related strategies that would seem to be as good—if not better than—aggregating individual predictions.

First, if aggregate predictions merely match (or sometimes prove to be slightly worse than) the predictions of the best-performing individuals in a group, why not just identify those best performers and adopt whatever prediction they make? At first blush, this strategy would seem to be sound, but it is not for two reasons. First, people have a very difficult time discerning who among them are the best performers (Libby, Trotman, & Zimmer, 1987; Miner, 1984; Trotman, Yetton, & Zimmer, 1983; Yetton & Bottger, 1982). Second, the very nature of an aggregate prediction is that it gets rid of the error associated with individual predictions. Any individual predictor, even best performers, will be vulnerable to having a greater degree of error associated with his or her prediction. The very presence of that error puts best performers at risk for making predictions that, at times, are wrong. Because aggregate predictions remove, at least in part, that error, this risk is reduced.

The second imprudent strategy would be to bring group members together to hash out their predictions before aggregating them. Groups that interact with each other typically produce products that are not as good as those produced by "groups" formed by merely combining the efforts of individuals working separately (Diehl & Stroebe, 1987; Guzzo, 1986; Hill, 1982; Mullen, Johnson, & Salas, 1991). This is also true in the realm of forecasts. Plous (1995) asked groups of college students to estimate the answers to such questions as the year that Wolfgang Mozart died. Groups talking together produced worse estimates than those produced by averaging the individual estimates of people working independently.

Group interaction detracts from the benefits of aggregation because that interaction causes people to become

redundant with one another. Working separately and independently, each individual is likely to base his or her prediction on at least a slightly different set of information, background knowledge, beliefs, and theories. In this way, each individual brings something unique to the prediction process that, when aggregated with the unique insights of other individuals, may bring the group aggregate closer to reality. However, when people talk, they tend to dwell on the information and knowledge that they share and, thus, neglect perspectives that each uniquely brings to the task (Stasser & Titus, 1985, 1987), and that might pull the group's estimate closer to the truth. As a consequence, group interaction prompts group members to think of the prediction problem in the same way, and thus to become biased in their prediction in the same direction. Aggregation under this circumstance fails to cancel the bias that group members have acquired. As a consequence, group predictions fail to prove superior to individual predictions. This holds true even when groups adopt "best practices" that are often associated with good decision making, such as assigning one group member to play a devil's advocate to the rest of the group's thinking (Plous, 1995).

Group interaction, however, carries another deleterious consequence. Although interacting groups do not necessarily outperform individual judgments, groups that interact tend to reach more extreme and polarized judgments (Myers & Lamm, 1976), as well as judgments that are held with more confidence (Ono & Davis, 1988; Paulhus, Dzindolet, Poletes, & Camacho, 1993; Sniezek & Henry, 1989; Stephenson, Clark, & Wade, 1986), particularly when they reach wrong decisions (Punchohar & Fox, 2004). Interacting groups display more overconfidence than groups formed by aggregating estimates generated individually and then averaged together (Plous, 1995).

Economists have increasingly recognized the value of aggregating predictions to forecast future events better, but they have supplied their own twist. Across several different domains, economists have set up *prediction markets* (also known as *information markets* or *events futures*) in which participants buy and sell financial securities based on their expectations of future events. From the prices of securities traded, researchers derive the probability that the event will occur (Pennock, Lawrence, Giles, & Nielsen, 2000; Wolfers & Zitzewitz, 2004).

The most famous example of such a prediction market is the Iowa Electronic Market (IEM) in which participants trade, for example, securities based on their predictions of who will win presidential and congressional elections. For example, in 2000, the IEM offered a security that paid $1 if Al Gore won the popular vote in the presidential election. People could pay from $.00 to $1 for the security, with the price offered roughly translating to the probability that Al Gore would win the popular vote, if they were acting rationally. For example, if a person bid $.52 for the Al Gore security, that meant that he or she would "break even" in expected value if there were a 52% chance that Al Gore would win (Wolfers & Zitzewitz, 2004).

Many such prediction markets have proliferated in recent years. The Hollywood Stock Exchange offers securities used to predict the opening weekend box office of individual movies, as well as Oscar winners. The Foresight Exchange trades securities tied to possible scientific discoveries. Tradesports.com issues securities tied to the outcomes of sporting events, as well as world events (such as whether Saddam Hussein would still be in power as of September 30, 2003) (Pennock et al., 2000; Servan-Schreiber, Wolfers, Pennock, & Galebach, 2004; Wolfers & Zitzewitz, 2004). Companies have started to run internal prediction markets, forecasting, for example, which new drugs are likely to be successful and which software packages will be developed before the deadline (Wolfers & Zitzewitz, 2004).

According to the logic of prediction markets, the market price serves as an accurate and unbiased estimate of the likelihood of future events. This accuracy arises because the market price contains the sum total of individual estimates driven by all the information and background knowledge that traders bring to the market. In addition, markets should draw only those people who have some information and expertise. People who have little information on which to base a prediction should avoid the market. But more important, markets provide individuals with incentives for trying to get the prediction right. Those incentives might be financial, but at times they can simply be the incentives inspired by competitive spirit and a need to maintain reputation. Such incentives should prompt individuals to set aside personal biases and may even motivate people to research their predictions (Wolfers & Zitzewitz, 2004).

Prediction markets, when tested, do seem to have an enviable record of accuracy. For example, prediction markets tied to elections appear to predict the actual outcome more accurately than do opinion polls. For example, the IEM missed the final popular vote total of U.S. presidential elections from 1988 through 2000 by less than 1.5%, whereas the Gallup poll missed on average by 2.1%. Across 15 cases in which the IEM was pitted against the prediction derived from opinion polls, the market proved more accurate in nine of the cases (Berg, Forsythe, Nelson, & Rietz, in press). In a study of sporting events, a market involving nearly 3,000 National Football League fans, the price set by the market outperformed all but roughly a dozen of the market's individual participants, regardless of whether market participants traded securities involving real or play money (Servan-Schreiber et al., in press). The Hollywood Stock Exchange provides estimates of opening weekend box office receipts that tend to diverge from actual receipts by only 4% on average, a rate that equals the most well-known prediction models used by the film industry. When it comes to predicting who will win the Oscar, the Hollywood Exchange predicts better than panels of experts (Pennock et al., 2000).

To be sure, prediction markets do not rid predictions of all biases that people display in their forecasts. Such markets do not completely stop people from acting on optimistic premises. Political partisans, for example, still

favor their party in election markets (Forsythe, Reitz, & Ross, 1999). In addition, markets, like people, still tend to overestimate the likelihood of rare events (Wolfers & Zitzewitz, 2004). Markets also are not necessarily accurate when dealing with events for which there is little public information available (such as who will be elected as the next Pope) (Wolfers & Zitzewitz, 2004). But, as a general mechanism for aggregating individual knowledge from diverse perspectives into an aggregate estimate, prediction markets, as demonstrated by an impressive first-generation of research, have shown much promise. They may not provide perfect predictions, but the level of perfection found in their predictions far exceeds that of any alternative. Future research will have to be conducted to see if this initial promise continues to be revealed or whether this promise is bounded by circumstances that have yet to be identified.

CONCLUDING REMARKS

At the beginning of this chapter, I noted all the time, energy, and expense that people spent focusing on the future and trying to anticipate it. I then went on to chronicle how, despite the effort that people expend on prediction, the forecasts they provide tend to possess some systematic biases. People tend to be too optimistic, and they tend to express too much confidence in the predictions they provide.

The extant research suggests that these biases share a common source. People build mental scenarios of the future, simulating what events might arise in the future and then noting what outcomes those events might imply. The scenarios that people build, however, tend to be narrow and incomplete. They neglect pessimistic turns that the future might take. They give short shrift to scenarios that might imply alternative outcomes other than the one the individual focuses on.

More problematic, however, is the impossibility of anticipating all the situational features and forces that might influence behavior once that situation arrives. Human behavior depends importantly on situational details, but often those details are unknowable. Human behavior also depends on emotion, and an emerging body of evidence suggests that people do not adequately anticipate how emotion influences the preferences people form and the actions they take once emotion is aroused.

However, such a portrait of the human ability to render predictions would be unnecessarily bleak. People may provide imperfect predictions when they build scenarios and adopt an "inside view" to prediction, but they do have information available that would help to form more accurate predictions if they used it.

In essence, if people formed new habits when they approached predictions or decided to use a different strategy, they might have the wherewithal to enhance the accuracy of their predictions. One new strategy would be to adopt an "outside," more data-driven, approach to decision making—looking at similar situations that have happened in the past or to other people and using those ex-

periences as a basis for their predictions. People would also profit if they collated their predictions with those of other people. People may base their individual predictions on incomplete information and imperfect expertise, but pooling their flawed suspicions with those of others appears to aggregate valid suspicions while getting rid of erroneous ones. Psychologists have discovered this principle by pooling the estimates of their respondents. Economists have discovered this principle as well in their explorations of market behavior.

Given all this, it appears that the time is ripe for further research on the psychology of prediction. Past research has shown, in some detail, what people fail to do when they render predictions, and how those failures damage the accuracy of people's forecasts. The future could, instead, turn away from these failures to make proposals about what people could or should do—what habits they could adopt—to render more accurate predictions. Are there ways to improve on the outside approach to prediction? Are there times in which the inside view is superior to the outside approach? How should outside tactics be blended with inside information to take advantages of the merits of both? And what about aggregating predictions? Are there other more valuable ways to combine the predictions that different individuals make? Putting this all together, is there a recipe of "best practices" that would aid people to provide the most accurate forecasts possible?

There are many questions to answer, and it is difficult to predict what the answers of those questions would look like.

REFERENCES

Ahrens, C. (1991). *Meteorology today: An introduction to weather* (4th ed.). St. Paul, MN: West.

Arkes, H., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods for reducing overconfidence. *Organizational and Human Decision Processes*, *39*, 133–144.

Ashton, R. H. (1986). Combining the judgments of experts: How many and which ones? *Organizational Behavior and Human Decision Processes*, *38*, 405–414.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Science*, *5*, 533–539.

Benson, P. L., Karabenick, S. A., & Lerner, R. M. (1976). Pretty pleases: The effects of physical attractiveness, race, and sex on receiving help. *Journal of Experimental Social Psychology*, *12*, 409–415.

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (in press). Results from a dozen years of election futures research. In C. Plott & V. Smith (Eds.), *Handbook of experimental economic results*.

Buehler, R., & Griffin, D. (2003). Planning, personality, and prediction: The role of future focus in optimistic time predictions. *Organizational Behavior and Human Decision Processes*, *92*, 80–90.

Buehler, R., Griffin, D., & MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. *Personality and Social Psychology Bulletin*, *23*, 238–247.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*, 366–381.

Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: The causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 250–270). New York: Cambridge University Press.

Buehler, R., & McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin*, *27*, 1480–1493.

Christiansen-Szalanski, J. J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 928–935.

Clemens, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*, 559–583.

Darley, J., & Batson, C. D. (1973). From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, *27*, 100–108.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674.

Della Vigna, S., & Malmendier, U. (2002). *Overestimating self-control: Evidence from the health club industry*. Unpublished manuscript, University of California–Berkeley and Stanford University.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, *53*, 497–509.

Diener, E., & Seligman, M. E. P. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest*, *5*, 1–31.

Diekmann, K. A., Tenbrunsel, A. E., & Galinsky, A. D. (2003). From self-prediction to self-defeat: Behavioral forecasting, self-fulfilling prophecies, and the effect of competitive expectations. *Journal of Personality and Social Psychology*, *85*, 672–683.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model of judgments of likelihood. *Psychological Review*, *106*, 180–209.

Dougherty, M. R. P., Gettys, C. F., & Thomas, R. P. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes*, *70*, 135–148.

Dunn, E. W., Wilson, T. D., & Gilbert, D. T. (2003). Location, location, location: The misprediction of satisfaction in housing and lotteries. *Personality and Social Psychology Bulletin*, *29*, 1421–1432.

Dunning, D. (2001). On the motives underlying social cognition. In N. Schwarz & A. Tesser (Eds.), *Blackwell handbook of social psychology: Volume 1. Intraindividual processes* (pp. 348–374). New York: Blackwell.

Dunning, D., Griffin, D. W., Milojkovic, J. H., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, *58*, 568–592.

Dunning, D., & Madey, S. F. (1995). Comparison processes in counterfactual reasoning. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 103–132). Hillsdale, NJ: Erlbaum

Dunning, D., & Parpal, M. (1989). Mental addition versus subtraction in counterfactual reasoning: On assessing the impact of personal actions and life events. *Journal of Personality and Social Psychology*, *57*, 5–15.

Dunning, D., & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology*, *61*, 521–532.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, *30*, 510–517.

Einhorn, H. J., Hogarth, R. M., & Klempner, E. (1977). Quality of group judgment. *Psychological Bulletin*, 158–172.

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology*, *79*, 861–875.

Epley, N., & Dunning, D. (2004). *The mixed blessing of self-knowledge in behavioral prediction*. Unpublished manuscript, Cornell University.

Eyal, T., Liberman, N., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology*, *86*, 781–795.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 552–564.

Forsythe, R., Rietz, T., & Ross, T. (1999). Wishes, expectations and actions: Price formation in election stock markets. *Journal of Economic Behavior and Organization*, *39*, 83–110.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.

Gilbert, D. T., & Ebert, J. E. J. (2002). Decisions and revisions: The affective forecasting of changeable outcomes. *Journal of Personality and Social Psychology*, *82*, 503–512.

Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, *88*, 430–444.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science*, *15*, 346–350.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*, 617–638.

Gilovich, T., Kerr, M., & Medvec, V.H. (1993). The effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, *64*, 552–560.

Goldberg, L. R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis vs. neurosis from MMPI. *Psychological Monographs*, *9*(Whole no. 79), 1–28.

Gollwitzer, P. M. (1990). Action phases and mind-sets. In R. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 2. Foundations of social behavior* (pp. 53–92). New York: Guilford Press.

Gollwitzer, P. M., & Kinney, R. F. (1989). Effects of deliberative and implemental mind-sets on illusion of control. *Journal of Personality and Social Psychology*, *56*, 531–542.

Gordon, K. (1924). Group judgments in the field of lifted weights. *Journal of Experimental Psychology*, *7*, 398–400.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology*, *59*, 1128–1139.

Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319–359). San Diego, CA: Academic Press.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*, 411–435.

Guzzo, R. A. (1986). Group decision making and group effectiveness in organizations. In P. S. Goodman (Ed.), *Designing effective work groups* (pp. 34–71). San Francisco: Jossey-Bass.

Halberstadt, J. B., & Niedenthal, P. M. (1997). Emotional state and the use of stimulus dimensions in judgment. *Journal of Personality and Social Psychology*, *72*, 1017–1033.

Hayward, M. L. A., & Hambrick, D. C. (1997). Explaining the pre-

miums paid for large acquisitions: Evidence of CEO hubris. *Administrative Science Quarterly*, *42*, 103–127.

Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizations compensate for the shortcoming of individual learners. *Research in Organizational Behavior*, *20*, 1–37.

Hill, G. W. (1982). Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin*, *91*, 517–539.

Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 719–731.

Hogarth, R. M. (1978). A note on aggregating opinions. *Organizational Behavior and Human Performance*, *21*, 40–46.

Hsee, C. K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology*, *86*, 680–695.

Hynie, M., & Lydon, J. E. (1996). Sexual attitudes and contraceptive behavior revisited: Can there be too much of a good thing? *Journal of Sex Research*, *33*, 127–134.

Isen, A., Clark, M., & Schwartz, M. (1976). Duration of the effect of good mood on helping: "Footprints in the sand of time." *Journal of Personality and Social Psychology*, *34*, 385–393.

Jiang, J. J., Muhanna, W. A., & Pick, R. A. (1996). The impact of model performance history information on users' confidence in decision models: An experimental examination. *Computers in Human Behavior*, *12*, 193–207.

Jorgensen, M. (2004a). A review of studies on expert estimation of software development effort. *Journal of Systems and Software*, *70*, 37–60.

Jorgensen, M. (2004b). Top-down and bottom-up expert estimation of software development effort. *Information and Software Technology*, *46*, 3–16.

Juslin, P., Winman, A., & Olsson, H. (2000). Naïve empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*, 384–396.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*, 17–31.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovik, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, *107*, 852–884.

Koehler, D. J. (1994). Hypothesis generation and confidence in judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 461–469.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118.

Kruger, J., & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Personality and Social Psychology*, *40*, 586–594.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kusterbeck, A. (2004). 1955. In P. Smallman (Ed.), *The old farmer's almanac 2005*. Dublin, NH: Yankee Publishing.

Lagnado, D., & Sloman, S. (2004). Inside and outside probability judgment. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 157–176). Boston: Blackwell.

Latané, B., & Darley, J. (1969). Bystander "apathy." *American Scientist*, *57*, 244–268.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*, 146–159.

Libby, R., & Blashfield, R. K. (1978). Performance of a composite as a function of the number of judges. *Organizational Behavior and Human Performance*, *21*, 121–129.

Libby, R., Trotman, K. T., & Zimmer, I. (1987). Member variation, recognition of expertise, and group performance. *Journal of Applied Psychology*, *72*, 81–87.

Liberman, N., Eyal, T., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology*, *85*, 781–795.

Liberman, N., Sagristano, M., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology*, *38*, 523–535.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*, 5–18.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know also know more about how much they know? *Organizational Behavior and Human Performance*, *20*, 159–183.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

Lovallo, D., & Kahneman, D. (2003). Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review*, *8*, 56–63.

Matlin, M., & Stang, D. (1978). *The Pollyanna principle: Selectivity in language, memory, and thought.* Cambridge, MA: Schenkman.

McKenzie, C. R. M. (1997). Underweighting alternatives and overconfidence. *Organizational Behavior and Human Decision Processes*, *71*, 141–160.

Meehl, P. (1954). *Clinical versus statistical predictions: A theoretical analysis and review of the evidence.* Minneapolis: University of Minnesota Press.

Miner, C. F. (1984). Group versus individual decision making: An investigation of performance measures, decision strategies, and process losses/gains. *Organizational Behavior and Human Performance*, *33*, 112–124.

Moore, D. A. (2004). The unexpected benefits of final deadlines in negotiation. *Journal of Experimental Social Psychology*, *40*, 121–127.

Moore, D. A., & Kim, T. G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology*, *85*, 1121–1135.

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology*, *12*, 3–23.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 6072–627.

Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information relevance, working memory, and the consideration of alternatives. *Quarterly Journal of Experimental Psychology*, *46*, 759–778.

Nelson, M. W., Bloomfield, R., Hales, J. W., & Libby, R. (2001). The effect of information strength and weight on behavior in financial markets. *Organizational Behavior and Human Decision Processes*, *86*, 168–196.

Newby-Clark, I. R., Ross, M., Buehler, R., Koehler, D. J., & Griffin, D. (2000). People focus on optimistic and ignore pessimistic scenarios while predicting their task completion times. *Journal of Experimental Psychology: Applied*, *6*, 171–182.

Niedenthal, P. M., Halberstandt, J. B., & Innes-Ker, A. H. (1999). Emotional response categorization. *Psychological Review*, *106*, 337–361.

Nisbett, R. E., & Kunda, Z. (l985). Perception of social distributions. *Journal of Personality and Social Psychology*, *48*, 297–311.

Norem, J. K., & Cantor, N. (1986). Anticipatory and post hoc cushioning strategies: Optimism and defensive pessimism in "risky" situations. *Cognitive Theory and Research*, *10*, 347–362.

Norem, J. K., & Chang, E. E. (2002). The positive psychology of negative thinking. *Journal of Clinical Psychology*, *58*, 993–1001.

Nussbaum, S., Trope, Y., & Liberman, N. (2003). Creeping dispositionism: The temporal dynamics of behavior prediction. *Journal of Personality and Social Psychology*, *84*, 485–497.

Oksam, J., Kingma, J., & Klasen, H. J. (2000). Clinicians' recognition of 10 different types of distal radial fractures. *Perceptual and Motor Skills*, *91*, 917–924.

Ono, K., & Davis, J. H. (1988). Individual judgment and group interaction: A variable perspectives approach. *Organizational Behavior and Human Decision Processes*, *41*, 211–232.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Clinical Psychology*, *29*, 261–265.

Ostrom, T. M., & Shrauger, J. S. (1986). Self-prediction: Exploring the parameters of accuracy, *Journal of Personality and Social Psychology*, *51*, 1044–1057.

Paulus, P. B., Dzindolet, M. T., Poletes, G., & Camacho, L. M. (1993). Perception of performance in group brainstorming: The illusion of group productivity. *Personality and Social Psychology Bulletin*, *19*, 78–89.

Pennington, G. L., & Roese, N. J. (2003). Regulatory focus and temporal distance. *Journal of Experimental Social Psychology*, *39*, 563–576.

Pennock, D. M., Lawrence, S., Giles, C. L., & Nielsen, F. A. (2000). *The power of play: Efficiency and forecast accuracy in web market games* (NEC Research Institute Technical Report 2000–168). Princeton, NJ: NEC Laboratories.

Plous, S. (1995). A comparison of strategies for reducing interval overconfidence in group judgments. *Journal of Applied Psychology*, *80*, 443–454.

Polivy, J., & Herman, C. P. (2002). If at first you don't succeed: False hopes of self-change. *American Psychologist*, *57*, 677–689.

Puca, R. M. (2004). Action phases and goal setting: Being optimistic after decision making. *Motivation and Emotion*, *28*, 121–145.

Punchochar, J. M., & Fox, P. W. (2004). Confidence in individual and group decision making: When "two heads" are worse than one. *Journal of Educational Psychology*, *96*, 583–591.

Read, D., & van Leeuwen, B. (1998). Time and desire: The effects of anticipated and experienced hunger and delay to consumption on the choice between healthy and unhealthy snack food. *Organizational Behavior and Human Decision Processes*, *76*, 189–205.

Redelmeier, D., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified alternatives. *Medical Decision Making*, *15*, 227–230.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation*. New York: McGraw-Hill.

Ross, M., & Newby-Clark, I. R. (1998). Construing the past and future. *Social Cognition*, *16*, 133–150.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, *104*, 406–415.

Sagristano, M., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: Money now, odds later. *Journal of Experimental Psychology: General*, *131*, 364–376.

Sanna, L. J. (1999). Mental simulations, affect, and subjective confidence: Timing is everything. *Psychological Science*, *10*, 339–345.

Sanna, L. J., & Meier, S. (2000). Looking for clouds in a silver lining: Self-esteem, mental simulations, and temporal confidence changes. *Journal of Research in Personality*, *34*, 236–251.

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy?: A focusing illusion in judgments of life satisfaction. *Psychological Science*, *9*, 340–6.

Servan-Schreiber, E., Waolfers, J., Pennock, D., & Galebach, B. (2004). Prediction markets: Does money matter? *Electronic Markets*, *14*, 243–251.

Shapira, Z., & Berndt, D. J. (1997). Managing grand-scale construction projects: A risk-taking perspective. *Research in Organizational Behavior*, *19*, 303–360.

Sherden, W. A. (1998). *The fortune sellers: The big business of buying and selling predictions*. New York: Wiley.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*, 211–221.

Shrauger, J. S., Mariano, E., & Walter, T. (1998). Depressive symptoms and accuracy in the prediction of future events. *Personality and Social Psychology*, *24*, 880–892.

Sieff, E. M., Dawes, R. M., & Loewenstein, G. (1999). Anticipated versus actual reaction to HIV test results. *American Journal of Psychology*, *112*, 297–311.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, *43*, 1–28.

Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, *36*, 1202–1212.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*, 1467–1478.

Stasser, G., & Titus, W. (1987). Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology*, *53*, 81–93.

Stephenson, G. M., Clark, N. K., & Wade, G. S. (1986). Meetings make evidence? An experimental study of collaborative and individual recall of a simulated police interrogation. *Journal of Personality and Social Psychology*, *50*, 1113–1122.

Suantak, L., Bolger, F., & Ferrell, W. R. (1996). The hard-easy effect in subjective probability calibration. *Organizational Behavior and Human Decision Processes*, *67*, 201–221.

Taylor, S. E., & Gollwitzer, P. M. (1995). Effects of mindset on positive illusions. *Journal of Personality and Social Psychology*, *69*, 213–226.

Tetlock, P. E. (2002). Theory-driven reasoning about plausible pasts and probable futures in world politics. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 749–762). New York: Cambridge University Press.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–422.

Trotman, K. T., Yetton, P. W., & Zimmer, I. R. (1983). Individual and group judgments of internal control systems. *Journal of Accounting Research*, *21*, 286–292.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology*, *4*, 400–408.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3–15.

Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology*, *58*, 582–592.

Van Boven, L., & Loewenstein, G. (2003). Projection of transient visceral states. *Personality and Social Psychology Bulletin*, *29*, 1159–1168.

Van Boven, L., Loewenstein, G., & Dunning, D. (2004). *The illusion of courage in social predictions: Underestimating the impact of fear of embarrassment on other people*. Unpublished manuscript, University of Colorado.

Van Boven, L., Loewenstein, G., Welch, E., & Dunning, D. (2004). *The illusion of courage: Underestimating fear of embarrassment in public performance decisions*. Unpublished manuscript, University of Colorado.

West, R. F., & Stanovich, K. E. (1997). The domain specificity and generality of overconfidence: Individual differences in performance estimation bias. *Psychonomic Bulletin and Review*, *4*, 387–392.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 345–411). New York: Elsevier.

Wilson, T. D., Wheatley, T., Kurtz, J., Dunn, E. W., & Gilbert, D. T. (2004). When to fire: Anticipatory versus post-event reconstrual of uncontrollable events. *Personality and Social Psychology Bulletin*, *30*, 1–12.

Wilson, T. D., Wheatley, T., Meyers, J. M., & Gilbert, D. T. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*, 821–836.

Windshitl, P. D., Kruger, J., & Simms, E. (2003). The influence of egocentism and focalism on people's optimism in competitions. *Journal of Personality and Social Psychology*, *85*, 389–408.

Wolfers, J., & Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, *18*, 107–126.

Woodzicka, J. A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues*, *57*, 15–30.

Yaniv, I. (2004). The benefit of additional opinions. *Current Directions in Psychological Science*, *13*, 75–78.

Yetton, P. W., & Bottger, P. C. (1982). Individual versus group problem solving: An empirical test of a best member strategy. *Organizational Behavior and Human Performance*, *29*, 307–321.

Zarnowitz, V. (1984). The accuracy of individual and group forecasts from business and outlook surveys. *Journal of Forecasting*, *3*, 11–26.

CHAPTER 5

Expectancy

Neal J. Roese
Jeffrey W. Sherman

Contemplating the future consequences of present actions has a proud lineage among us primates, and is one of the secrets of what is still, by and large, the stunningly successful story of humans on Earth.

—Carl Sagan (1997, p. 91)

Nearly every human society has reserved a special place for those denizens claiming the ability to predict the future. Prophets, astrologers, shamans, witch doctors, and scientists are but a few examples of the numerous occupations predicated on foretelling what has yet to occur. Wielding runes, charts, crystal balls, hallucinogens, or mathematics, practitioners of prophecy have achieved widely varying degrees of predictive success. Yet even a little success goes a long way, for enormous is the value of prophecy to any society or any individual. Prophecy brings power in its purest sense: power to acquire that which is most desirable and avoid that which is most repugnant, power to achieve victory and conquest, power to bypass famine and flood. Regardless of the scope of prediction, the essential value remains invariant: to predict the future is to navigate it more effectively.

"Expectancy" is a generic term referring to beliefs about the future. From short-term dinner plans for the weekend to long-range forecasts of financial investments, people think about the future and use such thoughts in ongoing judgment, reasoning, decision making, and behavior. Expectancy is a core construct of psychology, a signature building block of cognition that is at once common among animals yet also uniquely human. It is common in the sense that the brains of all ambulatory organisms have evolved to abstract and record survival-oriented patterns of information (food here, predators there) that may then guide subsequent behavior. Yet expectancy is also uniquely human in the sense that we alone seem to have the capacity to create detailed imaginings of future possibilities, to erect vivid simulations of

environments and situations that have never before existed, and to coordinate ongoing behavior, often involving many people, to actualize those possibilities through effort and invention. Expectancy is perhaps a special case of the more general mental time travel capability inherent in episodic memory, which Tulving (e.g., 2002) also has argued to be unique to human cognition.

Expectancies are beliefs about a future state of affairs, subjective estimates of the likelihood of future events ranging from merely possible to virtually certain. This definition is taken from the expectancy chapter that appeared in the first volume of this handbook (Olson, Roese, & Zanna, 1996), and the review that follows is intended to build directly on that earlier chapter. The tone and argument of this chapter are similar and where conclusions remain the same, we restate them only briefly and refer back to the earlier chapter for elaboration. This chapter advances beyond the earlier one in three main ways. First, new insights into the tension between rigidity versus revision of mental representation, particularly in the face of disconfirming information, permit a more rigorous portrait of the expectancy construct. Second, the mushrooming literature on stereotyping (which may be defined partly in terms of expectancies for particular social groups) has yielded a variety of insights that must be integrated into a general overview of expectancy. Third and most centrally, this review is organized around a functional perspective rooted to principles of effective behavioral control, coordination, and automaticity. Expectancy is first and foremost an instantiation of those core cognitive mechanisms geared to action and survival.

The chapter is organized around the following sections. We first establish our theoretical framework by explaining the functional basis of expectancies. We then discuss determinants of expectancies, framed in terms of content-neutral parameters. The consequences of expectancies are then reviewed, structured by the central principle that behavior regulation is the primary function of expectancies, with cognitive and affective consequences operating in support of that primary function.

FUNCTION

The most general and basic function of expectancies is to guide effective behavior. Expectancies constitute information gleaned from past experience, but to be effective, they must be sufficiently lean to be deployed rapidly, particularly when processing resources are taxed. This section focuses first on how expectancies are used to regulate behavior, then on the efficiency of expectancy use.

Behavior Regulation

Expectancies are tools for survival. By anticipating future fortune or misfortune (i.e., by constructing a cognitive map that suggests and directs means of acquisition and avoidance), an organism is in a vastly better position subsequently to acquire and avoid successfully (Bandura, 1986; Higgins, 2000; Irwin, 1944; Rotter, 1954; Tolman, 1932). In short, to act effectively in the world is to draw on information gleaned from previous experience. The expectancy is where past and future meet to drive present behavior. This is the essence of the expectancy construct, a psychological mechanism that is first and foremost a tool for guiding behavior and, hence, ultimately a tool for survival.

Several core functions (or motives) have occupied center stage in recent social psychological theory, including needs for accuracy, control, improvement, affiliation, and affect regulation (e.g., Fiske, 2003; Sanna, 2000; Sedikides & Strube, 1997). Perhaps the most basic function underlying the majority of these conceptions is simply that of behavior control, the regulation of single and sequential actions to ensure survival. Effective management of ongoing behavior subsumes accuracy and control functions and may be put in the service of improvement and affiliation. Affect regulation (mood maintenance, mood improvement, self-enhancement, etc.) is a distinct function best explored separately, but we argue from the start that it is a secondary function of expectancies, largely subservient to the primary function of behavior control. But human beings do vastly more than merely survive; they plan marriages and careers, they move in and out of communities, and they spend an enormous amount of time attending to details both large and small to turn ideas into reality. Expectancy is not just a tool for survival but a necessity for modern living. For this reason, the functional basis of expectancy is best appreciated not at the level of discrete thoughts but at the level of complex thought systems.

Take the example of a roadtrip, say, driving from Chicago to Memphis. Such a journey over hundreds of miles requires not just an expectancy ("I'll be there on Sunday") but a system of expectancies, consisting of at least four different kinds of anticipatory cognitions: (1) a *superordinate goal* (i.e., the destination, Memphis, the ultimate endpoint that gives coherence to the other expectancy components), (2) *plans* (i.e., a set of subordinate goals specifying particular actions that must be implemented to reach the superordinate goal, such as routine car maintenance, buying gas, and making arrangements for an overnight stay midway), (3) *semantic expectancies* (i.e., derived from semantic memory, these form an interlocking web of implicit background assumptions, ranging, for example, from traffic laws, the map layout of interstate freeways, and the location of fast-food outlets), and (4) *episodic expectancies* (derived from similar past experiences stored in episodic memory, such as the last roadtrip down the I-57 highway or the degree of congestion during the last holiday long weekend). Successful goal completion rests on this interlocking set of multiple expectancies.

Regulatory feedback loops are a defining feature of goals (Austin & Vancouver, 1996), and indeed they pervade an expectancy system simultaneously and at multiple levels. A negative feedback loop is one in which the current state is compared to an ideal or expected state, with discrepancies between the two directly eliciting control changes designed to reduce the discrepancy. Returning to the roadtrip example, at the very lowest level, that of the briefest time duration and most subordinate of goals, sits a feedback loop for the mere act of driving. Yet driving is a behavior of tremendous complexity partly obscured by its automaticity. Driving requires continuous online monitoring of visual signals indicating position on the road, proprioceptive feedback revealing motion and acceleration, and symbolic information conveyed by signaling devices such as speedometers and signposts. This information is compared to expectancies, in the form of desired position or speed, and deviations between the desired state and the actual state require rapid correction using pedals and steering wheel. Absence of such correction (e.g., when sleeping at the wheel) is disastrous. This short-term, low-level feedback loop is but one example; yet the same conceptual operation, widely known as a TOTE unit (see Miller, Galanter, & Pribram, 1960), occurs over longer time periods and for larger goals. People compare their current marriages, careers, and other long-running experiences against expectancies taking the form of dreams, ideals, or obligations to others; they note discrepancies; and they engage subsequently in behavior aimed at correcting or reducing those discrepancies (Carver & Scheier, 1998; Higgins, 1987; Markus & Nurius, 1986).

The concept of a negative feedback regulatory loop emphasizes the twin aspects of stability and change with regard to both the incoming current state information and its behavioral response. Current-state information that is similar to the expected state embodies a situation that is normal and requires no change in behavior; state

information that is suddenly dissimilar creates a situation that is abnormal and does require change in behavior. In normal situations, then, expectancies serve to furnish background assumptions, and here their influence is silently implicit. In abnormal situations, however, expectancies become a jarring reminder of how things "ought to have been," defining in precise terms the ways in which the current situation has deviated, and thereby suggesting (or activating) information relevant to behavioral correction (Roese, 2001). Activation of cognitive efforts to better understand what amounts to a failure of prediction is the immediate result, but so too is activation from memory of information semantically related to the newly changed situation, information that creates new expectancies but also suggests new compensatory or corrective behaviors. In a subsequent section, we return to the enormous body of evidence that details what happens when expectancies are confirmed (normal situation) or disconfirmed (abnormal situation).

Behavior control therefore requires continuous online processing in the form of continuous comparison, or pattern matching, between the current state and the expected state. Very likely in parallel to this comparative process is the conceptually similar online comparison between the current state and recent past state. *Processing fluency* is a construct that captures the moment-by-moment degree of deviation detected in such comparisons, with high fluency characterizing the smooth flow of incoming state information that matches closely either the expected state or past state templates (e.g., Benjamin & Bjork, 1996; Jacoby & Dallas, 1981; Johnston & Hawley, 1994; Whittlesea & Williams, 2001). Processing dysfluency characterizes the detection of a mismatch and is the functional equivalent of an alarm system (Lieberman, Gaunt, Gilbert, & Trope, 2002). The perception of fluency and dysfluency may be explicit ("Why did she turn right with her left turn signal blinking?") or subtle and implicit (a momentary stutter in the ongoing processing of visual information that may not be consciously experienced, but may trigger resolution processes). Here there is an exciting new connection of cognitive function to structure with the identification of the anterior cingulate as the probable brain site at which this online monitoring of processing fluency occurs (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Lieberman et al., 2002). Activity in the anterior cingulate indeed corresponds to rapid shifts in emotion, including pain and anxiety, and also attendant behavioral correction. Crucially, anterior cingulate activity is heightened when expectancies are disconfirmed (Carter et al., 1998). Also implicated in detection of expectation disconfirmation is the orbitofrontal cortex, a region with direct neuronal connections to the anterior cingulate (Berns, McClure, Pagnoni, & Montague, 2001; Camille et al., 2004; Thorpe, Rolls, & Maddison, 1983). From this vantage point, the key mechanistic underpinning of expectancy confirmation and disconfirmation is, respectively, activation of similar versus dissimilar information from memory and fluent versus dysfluent processing.

We argue that most expectancies are accurate. Those expectancies based on semantic memory (i.e., based on slowly learned, general world knowledge) are largely accurate, mostly implicit, and utterly essential to effective behavior. Past research on expectancies in particular, but also knowledge activation in general, however, has tended to overlook semantic knowledge in favor of those expectancies rooted to episodic memory (i.e., based on rapidly learned specific instances), which are more often explicit and perhaps more prone to bias. To be sure, expectancies can produce judgmental error, but they do so only rarely against a backdrop of magnificent, silent, and often unappreciated success at overall behavior control.

Efficiency

For expectancies to function well, however, they must deliver information not only accurately but also efficiently. Expectancies that cannot be applied quickly and easily will be of little use in situations requiring swift decisions and nimble action. Indeed, such situations heighten the value of applying expectancies to guide behavior. With insufficient time to comprehend a situation in terms of its unique array of specifics, the ability to bring past experience to bear can be a lifesaver. Empirical evidence from multiple domains supports the idea that expectancies are especially likely to be relied on under trying circumstances. For example, people rely to a greater extent on expectancies relevant to persuasion (e.g., expert opinions are valid) and intergroup behavior (e.g., skinheads are unfriendly) when processing capacity is constrained (for reviews, see Petty & Wegener, 1998; Sherman, Macrae, & Bodenhausen, 2000). More broadly, there is considerable evidence that human judgment and behavior proceeds via *bounded rationality* and that relatively simple rules (a kind of expectancy) are relied on to provide relatively accurate information in a very efficient way (e.g., Gigerenzer, Todd, & ABC Research Group, 1999; Kahneman & Tversky, 1973; Simon, 1956). The efficiency with which expectancies may be applied is affected by a number of important properties and is reflected in the ways that expectancies influence the ongoing encoding of expectancy-relevsant information.

Summary

Expectancies are mental constructions used to guide and regulate behavior. As such, they are best conceptualized as tools for survival. Expectancies guide behavior with great efficiency, meaning that they provide useful information rapidly and with little demand on processing resources. These main ideas are refined further in the next section.

DETERMINANTS AND PARAMETERS

Expectancies vary along several dimensions, or parameters, and the particular envelope of variation along all these parameters for any one expectancy goes a long way toward characterizing its determinants and conse-

quences. Using a parametric approach to categorizing expectancies was a central feature of the earlier chapter (Olson et al., 1996), and it may be contrasted to a typological approach that defines discrete subtypes of expectancies. Miceli and Castelfranchi (2002) used this latter approach, distinguishing between, for example, forecasts (predictions of events believed to be likely), hopes (future events that are thought to be desirable though not necessarily likely), and hope-casts (a predicted future event that is desirable, likely, and which therefore ought to occur). These authors also used mirror-image terms for negatively valenced events (fears and fear-casts). But because of its content-neutral usefulness in encapsulating numerous judgment domains, we retain the parametric approach employed by Olson et al., modifying it slightly to reflect new theoretical insights to have appeared in the intervening years. Five parameters are discussed: likelihood, confidence, abstractness, accessibility, and explicitness.

Likelihood

Perhaps the most basic way of describing an expectancy is in terms of likelihood of occurrence. An expectancy describes an event that may or may not occur with some degree of probability, often conveniently expressed using a scale ranging from 0 to 100% (or 0 to 1). Thus, some expectancies refer to events believed to be low in likelihood ("Dave's attempt to quit smoking has about a 10% chance of success"; "This policeman is unlikely to be friendly"), moderate in likelihood ("Boeing has a 50% chance of meeting its quarterly earnings target"; "Bill is neither particularly introverted or extraverted"), or high ("I am 90% certain that I will eat grilled fish on Saturday"; "All professors are absent-minded"). Olson and colleagues (1996) used the term "subjective expectancy" to denote such variable likelihood beliefs. The extremity of subjective expectancies critically influences the manner and extent to which those expectancies guide behavior and cognition and may be confirmed or disconfirmed by experience.

The determinants of expectancy likelihood reflect the more general input sources for all beliefs, namely, information derived from past experience, social learning, the popular media, and the like. Mood has been shown to influence likelihood estimates, such that positive and negative moods increase perceived likelihoods of positive and negative events (Johnson & Tversky, 1983), but this effect appears to derive largely from the informational cues inherent in particular affective states (DeSteno, Petty, Wegener, & Rucker, 2000). Importantly, when existing expectancies are confirmed, their subjective likelihood may increase. For example, previous success on a particular task increases expected likelihood of future success (Feather, 1966; Feather & Saville, 1967).

People make inferences and attributions that mediate between the input of information and the output of expectancies. For example, consensus information increases entity attributions (Kelley, 1967), which may be taken to reflect belief in an external, objective reality. In this way, consensus information makes expectancies

seem more factual, a process that likely contributes to the conversion of initially subjective expectancies into factual ones ("Everyone agrees so it must be true"). Also, when attributions reflect a belief in stability rather than instability (Weiner, 1985), the resulting expectancy is higher in likelihood ("Chuck failed the exam because of weak ability, which is a stable disposition; I expect Chuck will fail the next exam too"). Importantly, this attribution mechanism underlies the *theory of hopelessness depression*, which suggests that when self-attributions for negative events are stable and global, expectations for the future become chronically bleak, resulting in depressive symptoms (cf. Abela & Seligman, 2000; Abramson, Metalsky, & Alloy, 1989). These are just a few examples of the many ways that interpretive cognitive mechanisms shape the perceived likelihood of future events.

Confidence

Any belief may be held with varying degrees of certainty or confidence. It is important to emphasize that confidence is orthogonal to likelihood. That is, high confidence is not the same thing as a belief in the high likelihood of occurrence. Take the flip of a fair coin. The likelihood of the coin landing with heads up is 50%, and knowing that the coin is fair, an observer would expect this likelihood with extremely high certainty. Both low- and high-probability events may be expected with both high or low confidence ("Jack is very confident that his chances of winning the spelling bee are about 20%"; "Analysts are only somewhat confident of the projected 95% success rate of the new missile system"). Confidence and likelihood are sometimes conflated in studies of expectancy, but we hope future researchers more clearly distinguish between them.

The determinants of expectancy confidence overlap partly with those that influence likelihood. For example, information derived from experience and communication with others can dictate to a large extent the confidence with which the individual expects a particular outcome. Direct personal experience tends to have a greater impact on confidence than indirect experience conveyed by others, as indicated by research on attitudes (Fazio & Zanna, 1981). In either case, increases in experience in a domain increase both the confidence with which expectancies are held and the likelihood that they are abstract. Thus, because abstract expectancies are typically based on a greater sample of experience than episodic expectancies, abstractions are held with greater confidence. Attributions and other interpretive cognitions influence confidence as well as perceived likelihood. Finally, past confirmation of expectancies also increases confidence (see Olson et al., 1996, for further discussion of these points).

Another determinant of expectancy confidence is the interconnection between the expectancy and other beliefs. The greater the degree of interconnection among semantically related beliefs, the greater the confidence with which such beliefs are held. In research on the *hindsight bias* (Hawkins & Hastie, 1990; Roese, 2004), it has been shown that people have difficulty disregarding the

information contained in an outcome in trying to recall their earlier expectancies for that outcome (e.g., after observers learned that the outcome of O. J. Simpson's notorious 1995 murder trial was acquittal, they misrecalled their earlier predictions for acquittal as being higher than they actually were; Bryant & Brockway, 1997). In other words, integration of new information into existing knowledge structures enhances confidence with which particular elements of the knowledge structure are held to be true, and this effect is very nearly the same in both retrospective and prospective judgments (Fischhoff, 1976; Gilovich, Kerr, & Medvec, 1993; Koriat, Lichtenstein, & Fischhoff, 1980). Thus, when people make predictions, they tend to bring to mind attendant information that is consistent with the prediction, which directly fuels overconfidence; manipulations that encourage consideration of alternative future outcomes mitigate that overconfidence (Dougherty, Gettys, & Thomas, 1997; Dunning, Griffin, Milojkovic, & Ross, 1990; Griffin, Dunning, & Ross, 1990; Hirt & Markman, 1995). Like skyscrapers, beliefs structured on interconnected scaffolding stand confidently taller.

This principle of belief interconnection underlies several more specific determinants of both expectancy likelihood and confidence. When people engage in vivid imagination, or *mental simulation*, of particular future event sequences, they subsequently believe the event to be more likely and are more confident of this belief (e.g., Anderson, 1983; Kahneman & Tversky, 1982; Koehler, 1991). Similar effects occur when people are asked to provide explanations (Ross, Lepper, Strack, & Steinmetz, 1977; Sherman, Skov, Hervitz, & Stock, 1981; Wilson & LaFleur, 1995) as to why a particular event might occur in the future. Furthermore, individual differences have also been explored; for example, people higher in need for closure tend to be more confident about future prospects (Hirt, Kardes, & Markman, 2004; Kruglanski & Webster, 1996).

Abstractness

The parameter of abstractness contains enormous implications for the functional basis of expectancies, particularly in terms of efficiency, depth, and temporal dependence. Abstractness refers to the variation between concrete and specific representations (as instantiated by episodic memory) and abstract generalizations that summarize experience across multiple events, people, and contexts over time (as exemplified by semantic memory).

Semantic Expectancies Are Efficient

Semantic expectancies may be used more efficiently than episodic ones. Semantic expectancies are preexisting knowledge structures that are extracted from ongoing experience, stored in memory, and retrieved when needed. By contrast, episodic expectancies must be formulated on the spot before they can be applied. Although both processes (retrieval vs. formulation) may occur automatically (e.g., Hintzman, 1986; Smith & Zarate, 1992), there nevertheless remains a clear difference

in that semantic expectancies deliver accurate, generalizable knowledge relatively more rapidly and with smaller demands on available resources.

The advantage in efficiency of semantic over episodic expectancies is especially evident when individuals encounter novel circumstances (McClelland, McNaughton, & O'Reilly, 1995; Nosofsky, Palmeri, & McKinley, 1994; Sherry & Schacter, 1987). In such cases, ad hoc expectancy generation runs into several problems. First, the levels of temporal, spatial, and contextual details preserved in episodic memories may inhibit the application of such knowledge to novel situations that do not possess those same features. To quickly extract generalities from multiple concrete episodes in an ad hoc fashion is a challenge (e.g., DeLosh, Busemeyer, & McDaniel, 1997). Second, the predictive validity of such expectancies depends on the number and breadth of experiences on which they are based: The larger the sample size, the greater the validity (e.g., McClelland et al., 1995). Third, it is simply more efficient to extract and store generalities in an ongoing fashion than it is to recalculate them every time they are needed by retrieving and summarizing a subset of episodes (Bruner, Goodnow, & Austin, 1956; Hamilton & Mackie, 1990; Klein, Cosmides, Tooby, & Chance, 2002; Nosofsky et al., 1994). Indeed, research demonstrates that the retrieval and application of specific episodes is more easily disrupted than is application of abstract knowledge (e.g., Johnson, Hashtroudi, & Lindsay, 1993; Rothbart, Fulero, Jensen, Howard, & Birrell, 1978; Sherman & Bessenoff, 1999; Tulving, 1983). Thus, extracting and storing abstract features of experience for subsequent use are more efficient ways to turn the past into functionally accurate expectancies than are retrieving and summarizing a large number of relevant episodes in an ad hoc fashion.

Episodic Expectancies Provide Depth

It is also important for effective behavior control to retain and develop domain- and context-specific information that may provide more situationally accurate expectancies than is possible via abstract knowledge. For example, though from past experience we might expect that dogs in our neighborhood are friendly and harmless, it would be important to know that the Rottweiler named Spike who lives around the corner is aggressive and dangerous. Recent evidence points further to the functional interplay between semantic and episodic knowledge. First, specific episodes may be retained and associated with a relevant semantic expectancy, so that when the latter is activated, so too are specific exceptions to the general rule embodied in that semantic expectancy (Bartlett, 1932; Graesser, Gordon, & Sawyer, 1979; Klein et al., 2002; Nosofsky et al., 1994). Some have argued that the main evolutionary purpose of episodic memory is to store instances that violate general expectancies about the world (e.g., McClelland et al., 1995; Schank, 1982; Sherry & Schacter, 1987).

Very likely is a developmental sequence through which people construct a full complement of expectancies that maximize their ability to go beyond the specifics of past

experiences to predict novel situations, at the same time maximizing the specificity, applicability, and accuracy of those expectancies. Initially, expectancies are based on particular episodes, but as experience accumulates, individuals extract semantic expectancies that summarize events across stimuli, time, and situations. Sometimes, however, expectancies become overly general and yield inaccuracies. When confronted with these inaccuracies, expectancies are refined, and more subordinate, narrowly defined expectancies are developed. We learn that not all dogs, roads, or skinheads are the same, and we develop expectancy subcategories. As discussed in detail in a subsequent section, subcategorization is often a direct result of expectancy disconfirmation. In short, expectancies are initially narrow and specific, then become broader and more general, and finally settle at midlevels that balance breadth with depth.

Abstractness Is Temporally Dependent

Expectancies become more abstract the further into the future the individual looks. Events that are imminent, by contrast, are conceptualized more in terms of concrete details. Under the rubric of *construal-level theory* (Trope & Liberman, 2003), recent research shows that in expectancies focusing on the distant future, people use fewer categorical distinctions and are more likely to rely on cognitive simplifications such as the correspondence bias, as compared to expectancies focusing on temporally nearer events. Abstractness as described within construal-level theory can help to explain other kinds of temporal shifts (e.g., those involving value and motivation).

People place greater value on objects or events that are temporally close than far in the future (Loewenstein, 1987; Loewenstein & Prelec, 1993). Given a choice, people prefer rewards that are available in the short term over those available in the long, even if getting something sooner means trading off a bit in value, a piece of human nature that any lender, creditor, or loan shark can readily confirm. Several reasons for this temporal discounting have been suggested, including factors centering on the affective versus cognitive basis of expectancies (i.e., affective expectancies show greater temporal dependence than cognitively rooted expectancies; e.g., Loewenstein, Weber, Hsee, & Welch, 2001; Metcalfe & Mischel, 1999) and on the magnitude of expected value (i.e., smaller values show greater temporal shifts than larger values; e.g., Chapman, 1996; Thaler, 1981).

Motivation also varies with temporal distance. Using the framework of *regulatory focus theory* (Higgins, 1997), Pennington and Roese (2003) showed that goals emphasizing promotion focus (i.e., those aimed acquisition, accomplishment, and improvement to the status quo) tend to aim further into the future, whereas prevention goals (i.e., those emphasizing caution, security, and preservation of the status quo) target the more immediate future. Moreover, people tend to "brace for the worst" by pondering greater potentials for mishap when events loom temporally close than far (Sanna, 1999; Shepperd, Findley-Klein, Kwavnick, Walker, & Perez, 2000). Curi-

ously, it may seem that the temporal dependence of value and motivation are contradictory: A person wants something better in the short term but is motivated to achieve something better in the long term. This contradiction, however, is illusory and disappears when firm definitions of value and motivation are enforced. People may find a thing more desirable now than later, yet they may nevertheless expect to attain even more desirable things in the more distant future. This basic principle is well exploited by creditors who offer "pay no interest for a year" deals. Consumers are unwilling to wait to save money for desired products and so pay interest to get the products today yet are curiously overconfident of their ability to pay off high-interest loans in the more distant future, apparently failing to recognize that neither their income nor their spending habits are likely to change that much.

As suggested already, the increasing abstractness with which increasingly distant future events are construed may be the underlying determinant of both shifts in value and motivation. The concrete tangibility, or "bird in the hand" aspect of temporally near events, may in and of itself confer value, while the abstract generality of temporally distant events may in and of itself invite broader, more maximal promotion goals aimed at comprehensive improvement of life circumstance. Temporal-construal theory points to this integration, and new research in the near (rather than distant) future will likely tackle it more directly.

Accessibility

Accessibility reflects the ease with which the expectancy is brought to conscious attention. As with other forms of knowledge, accessibility reflects the likelihood with which knowledge will be applied to subsequent judgment. Expectancies may be highly accessible because of frequency or recency of prior activation from memory (Higgins, 1996). Moreover, the experience of accessibility itself may feed into judgments: Expectancies may be optimistic, for example, not only as a function of the accessibility of positively valenced information but also as a function of the inference that positive information feels easy to bring to mind (Sanna & Schwarz, 2004).

Whether an expectancy is confirmed or disconfirmed is an important determinant of expectancy accessibility. When an outcome confirms an expectancy, only the merest gist of information regarding the outcome is abstracted and stored in memory, and the expectancy itself remains relatively inaccessible. By contrast, the very occurrence of expectancy disconfirmation makes the expectancy more accessible and instigates sense-making activity aimed at explaining the discrepancy between what is and what was expected. The cascade of cognitive processes stemming from expectancy disconfirmation are considered in detail in a subsequent section, but here we pause to mention two different processes that may result in the *apparent* increase in expectancy accessibility as a function of the experience of surprise. The first process involves, as we have mentioned, the heightened accessibility of a preexisting expectancy. For example, if an of-

fice coworker is seen wearing no clothing, the surprise serves to make accessible an implicit semantic expectancy regarding normal business attire. A second process, suggested by *norm theory* (Kahneman & Miller, 1986), is the absence of a preexisting expectancy. A feeling of surprise may nevertheless result if the outcome evokes the online construction of a new standard of comparison (or norm), assembled via the rapid integration of relevant exemplars, which is then contrasted to the outcome in question. Although this sort of *postcomputed* judgment process produces effects that mimic the effects of *a priori* expectancies, relatively little research has explored either deeper processing differences or the different circumstances in which these two processes arise (see McGill, 1993; Medvec, Madey, & Gilovich, 1995, for exceptions).

Explicitness

Expectancies also differ in the extent to which they are explicit (able to be consciously reported) versus implicit (unconsciously held). At the most basic level, an expectancy may be little more than an association between a concept and an attribute. Such associations are often held without any conscious awareness of their existence (e.g., Greenwald & Banaji, 1995). However, many complex expectancies, such as those that regulate driving a car, dining at a restaurant, or conversing with a friend, are also held at an implicit level. Indeed, the vast majority of the expectancies that guide our behavior are likely doing so without our awareness (Bargh & Ferguson, 2000; Nisbett & Wilson, 1977). The expectancy that can be reported explicitly and accurately is the exception, and expectancies likely only become explicit either when we are asked directly by others to articulate them or when we explicitly formulate them in our own minds when striving toward goals.

This is not meant to suggest that people cannot verbalize many expectancies when directly questioned about them. However, these verbalizations may be linked tenuously, at best, to the nature of the underlying patterns of covariation among attributes and concepts represented in memory systems. Though explicitly considered and reported expectancies are surely influenced by underlying patterns of association, they also are influenced by many other factors, including a host of self-promoting and self-presentational concerns that may distort or conceal the accurate expression of expectancy. Sometimes, as in the case of racial stereotypes, people are unwilling to express their true beliefs, even if they are aware of their nature. In many other circumstances, people are simply unable to introspect and identify the expectations that guide their perceptions, judgments, and behavior. The so-called unwilling and unable problem is longstanding in research on attitudes (Eagly & Chaiken, 1993), and has, over the years, led to a proliferation of measurement techniques designed to circumvent people's reluctance and inability to respond accurately, from the so-called bogus pipeline technique of the 1970s (reviewed in Roese & Jamieson, 1993) to the more recent explosion of indirect

measures, of which one was termed (with tongue in cheek) a "bona fide pipeline" (Fazio, Jackson, Dunton, & Williams, 1995; see Fazio & Olson, 2003, for review of indirect measures).[1]

One important consequence of the development of indirect measures is that it is now possible to measure and observe the correspondence between implicitly held and explicitly reported expectancies. In particular, interesting problems arise when the content of these expectancies are at odds with one another. What does it mean when such a conflict exists? What are the implications for judgment and behavior?

Most frequently, such dissociations have been treated as evidence that people may retain and store both implicit and explicit expectations of the same entity or outcome (e.g., Wilson, Lindsay, & Schooler, 2000). In the stereotyping literature, such dissociations are often described as reflecting the existence of separate implicit versus explicit stereotypes. The idea is that multiple representations of the same object or event exist in memory simultaneously in different memory systems.

An alternative view, and the one we endorse here, is that these dissociations reflect differences in the processes required and permitted by the direct and indirect measurement tasks, rather than differences in underlying representations (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Fazio & Olson, 2003; Roediger, 1990). That is, responses on different measurement tasks are just that: responses to tasks, rather than direct reflections of underlying representations. There is no isomorphism between task response and underlying representation. In this view, an important distinction between direct and indirect measures is the extent to which they permit conscious intentions and motivations to influence responses. Whereas indirect measures minimize these processes (though they do not eliminate them; e.g., Conrey et al., in press), direct measures do not. Direct, explicit measures demand subjective judgments that are influenced not only by the underlying systems of associations in memory but also by factors such as people's metatheories about what they believe, beliefs about what they would like to believe or should believe, and beliefs about what other people would like them to believe. Thus, in this view, dissociations between direct and indirect measures often reflect "downstream" processes that occur as an underlying, association-based "signal" is transformed into an explicit response, rather than reflecting true differences in the underlying expectancy "signal."

By no means do we intend to suggest that explicitly reported expectancies are meaningless or unimportant. Many important expectancies may be ill formed or may not exist at all until they are constructed and/or made explicit ("Where do you think this relationship is going?"; "What kind of wedding should we have?"; "How many children should we have?"). Moreover, regardless of their relation to implicit expectancies, the act of making an expectancy explicit undoubtedly has important consequences (e.g., making explicit the expectation that a relationship will lead to marriage). For example, explicitly

stating an expectancy increases the extent to which that expectancy guides subsequent thought and behavior (e.g., Kiesler, 1971; Rusbult, Martz, & Agnew, 1998). Finally, the overwhelming focus on implicit and explicit beliefs that are socially sensitive, such as those surrounding intergroup attitudes, has likely exaggerated the extent to which implicitly held and explicitly reported expectancies are dissociated. Rather little research of this sort has been conducted on other kinds of expectancies, a shortcoming we hope future research is able to address.

Ultimately, the most important question when considering implicit and explicit expectancies is to what extent does each actually predict behavior? So far the tentative answer is: It depends. Obviously, when there is no discrepancy between implicit and explicit expectancies, the two should influence behavior in similar ways. However, when the two are at odds, the key moderator appears to be the manner in which behavior is measured. Just as direct measures of belief permit greater influence of intent and motivation, so too do direct measures of behavior. Thus, it is not surprising that explicit measures (e.g., of racial attitudes and stereotypes) do a better job predicting explicit behaviors (e.g., verbal responses to white and black confederates, explicit evaluations of the confederates, racially relevant jury decisions, and judgments of the legitimacy of the Rodney King verdict) than do implicit measures of those beliefs (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1995). When conscious intent is permitted influence, it similarly influences expressions of belief and related expressions of behavior. Obversely, measures of behavior that minimize the role of conscious intent (e.g., eye contact and other nonverbal behaviors) are better predicted by indirect measures of belief that also minimize such factors than by explicit measures (e.g., Dovidio et al., 1997; Fazio et al., 1995). Thus, the most valuable measure of expectancy depends on exactly what it is that a researcher wishes to predict from it.

Summary

This section reviewed five parameters that may be used to characterize expectancies: likelihood, confidence, abstractness, accessibility, explicitness. Also discussed were the various influences on these parameters. Together, these parameters capture deeper commonalities in expectancies that differ widely in overt content. The principles by which expectancies guide effective behavior are reviewed in the next section.

BEHAVIORAL CONSEQUENCES

The primary function of expectancies is to guide behavior successfully and effectively. Expectancies do so by informing cognition. In this section we outline how expectancies facilitate successful behavior, then turn in the subsequent section to the more detailed topic of cognitive consequences.

Semantic Expectancies Provide General Guides for Behavior

The most basic function of expectancies is that they establish a set of broadly generic roadmaps for ongoing behavior. Semantic expectancies are those that are abstract and typically implicit. Buying food, navigating about town, getting work done, and enjoying leisure activities on the weekend all rest on vast networks of generic knowledge that comprises semantic memory. Novel situations are dealt with effectively by reliance on assumptions drawn from generic schematic knowledge.

Expectancies for Success Facilitate Success

Although there are numerous qualifications and caveats, it may generally be concluded that belief in future success facilitates future success (Lewin, Dembo, Festinger, & Sears, 1944; Oettingen & Mayer, 2002; Vroom, 1964). Moreover, commitment to a goal and course of action also facilitates performance (Locke, Shaw, Saari, & Latham, 1981; Mitchell, 1974). Expectancies for success produce effects on behavior by way of increases in confidence (Feather, 1966) and task persistence (e.g., Battle, 1965; Carver, Blaney, & Scheier, 1979). In this way, minor setbacks do little to interfere with the broader optimism that keeps the individual on track toward goal completion. There are two main mechanisms by which expectancy effects on performance occur. First, optimism in the sense of expectations of both personal efficacy (ability to succeed) as well as outcome success produces positive affect, which has a general motivating effect that energizes ongoing action (e.g., Bandura & Locke, 2003; Erez & Isen, 2002). Second, elaboration of specific plans fosters implementation intentions that guide ongoing action (e.g., Brandstätter, Lengfelder, & Gollwitzer, 2001; Gollwitzer, 1999; Pham & Taylor, 1999).

These optimism effects perhaps represent the default, but certainly not the only way expectancies influence goal-related behaviors. Drawing again on the regulatory focus theory (Higgins, 1997), the aforementioned effects center on promotion motivation, which involves focus on the attainment of desired future outcomes. Keeping one's "eye on the ball," so to speak, facilitates the eagerness that sustains pursuit of desired outcomes. But people may also strive toward prevention goals, which involve preserving the status quo by keeping at bay those outcomes that are not desired. Under this different motivational orientation, vigilance and defensive pessimism (i.e., focusing on undesirable future possibilities) facilitate success at keeping current circumstances from getting worse (McMullen & Markman, 2000; Norem & Cantor, 1986; Norem & Illingsworth, 1993; Showers, 1992).

On average, however, expectancies for personally relevant outcomes tend to be optimistically biased. True, people can be optimistic about their ability to engineer future success in both promotion and prevention, yet the former is more clearly aligned to the construct of optimism than the latter (Grant & Higgins, 2003). In a subse-

quent section, we note how optimism produces positive affect and hence may be used strategically for affect regulation. In the present context, it seems that a biased construal of reality that "spins" the future more positively than is objectively warranted might be behaviorally advantageous, as it may facilitate performance through either of the two mechanisms named above. But does such bias have drawbacks?

Optimistic Bias Shifts as a Function of Commitment to a Course of Action

As numerous authors have debated, there is a tension between the costs and benefits of bias versus accuracy, for inaccurate forecasts may impede successful action (e.g., Baumeister, 1998; McNulty & Karney, 2004). More generally, blind optimism in the face of obvious setbacks would seem to preclude efforts at remediation. However, it appears that optimistic bias is contingent upon the stage of progress toward goal completion. That is, before a course of action has been committed to, individuals are relatively unbiased, which is useful in facilitating accuracy-motivated assessment of available options. But once commitment to a course of action has occurred, optimism increases, which may then work to facilitate ongoing performance in the manner described above (Gollwitzer & Kinney, 1989; Taylor & Gollwitzer, 1995).

Anticipation of Setbacks Facilitates Corrective Action

In the pioneering early years of manned space flight, NASA engineers anticipated and mapped out in detail numerous failure scenarios (engine failure, navigation failure, explosive decompression, etc.), planned responses for each such scenario, then practiced these responses in dress rehearsals called mission simulations. People do pretty much the same thing. They anticipate possible problems and proactively plan solutions (Aspinwall & Taylor, 1997; Sanna, 2000). Research on anticipatory regret has revealed that people routinely take into consideration the consequences of potential future decisions and actions, then decide and act to avoid future regret (Zeelenberg, 1999). Much research in this tradition has emphasized the potential for bias: For example, individuals sometimes select objectively worse outcomes to bypass the potential for future regret (Zeelenberg, Beattie, van der Pligt, & de Vries, 1996) and they tend to mispredict the amount of regret that is actually felt (Crawford, McConnell, Lewis, & Sherman, 2002; Gilbert, Morewedge, Risen, & Wilson, 2004). As noted earlier, a prevention focus (whether defined as a momentary state or chronic individual difference) that evokes vigilant action can be effective particularly under those circumstances in which obstacles are many (Norem & Cantor, 1986; Showers, 1992). At a more basic level, this research collectively underscores the ability of individuals to create detailed simulations of potential future problems for the specific purpose of guiding proactive avoidance behavior. In short, there is functional value to worrying.

Expectancies Can Be Self-Fulfilling

As the previously noted principles show, expectancies can create their own reality. Such effects may enhance performance and improve the individual's life circumstances, but they may also constitute counterproductive bias. A widely studied example has been variously termed "self-fulfilling prophecy" and "behavioral confirmation." One of the simplest examples is the self-erasing effect of predictive error (Sherman, 1980). When individuals make explicit predictions for their own future performance, the act of explicit expectancy formation renders consistent information more accessible from memory, making this information more likely to guide subsequent behavior and thus creating a push toward behavioral consistency. With behaviors shifting to confirm prior expectancies, predictive "errors" may become less apparent over time.

The self-fulfilling prophecy centers on the effect of expectancies on interpersonal behavior, particularly when the perceiver's impressions or stereotypes suggest expectancies about a target person with whom the perceiver is interacting. Such expectancies guide the perceiver's behavior during interactions with target individuals, with the resulting expectancy-consistent behavior serving to elicit further expectancy-consistent behavior on the part of the target individual. Early research focused in particular on how the expectancies of teachers might influence the achievement behavior of students (Jussim, 1986; Jussim & Harber, 2005; Rosenthal & Jacobson, 1968), but subsequent research has revealed the generality of such effects, for example, in expectancies centering on relationship partners (Downey, Freitas, Michaelis, & Khouri, 1998; McNulty & Karney, 2002), gender stereotypes (Eagly, Wood, & Diekman, 2000; Snyder, Tanke, & Berscheid, 1977), and racial stereotypes (Word, Zanna, & Cooper, 1974). Such effects occur automatically, typically without awareness on the part of either the perceiver (Chen & Bargh, 1997) or the target (Vorauer & Miller, 1997). Yet even so, such effects tend to be relatively weak, when viewed against the backdrop of the general accuracy of expectancies (Jussim & Harber, 2005).

Self-fulfilling expectancy effects tend to be reduced by heightened accuracy motivation and by explicit awareness of the expectancy (Miller & Turnbull, 1986). A general principle extending to any assimilative priming effect is that such effects diminish when the perceiver becomes aware of the prime (Lombardi, Higgins, & Bargh, 1987; Strack, Schwarz, Bless, Kübler, & Wänke, 1993). Moreover, awareness on the part of the target individual, especially when he or she does not like the implications of the expectancy (as in the case of recognition of being unfairly stereotyped), can create what amounts to contrast effects, in that the target may deliberately behave in a manner that contradicts the expectancy to emphasize individuality or autonomy (Brehm, 1966; Neuberg, 1989). Given that these latter effects are motivated, they tend to occur when the expectancy is negative but not when it is positive.

Placebo effects are another example of the self-fulfilling nature of expectancies. A placebo effect occurs when a medical treatment (e.g., a pill) produces a physical impact not by way of any physical effect (e.g., the pill contains no active drug) but rather by way of the perceiver's belief in the efficacy of the treatment (i.e., an expectation of health improvement). In short, people may sometimes be fooled into wellness. Although debate about underlying mechanism continues (e.g., Stewart-Williams & Podd, 2004), the generality of the placebo effect as medical fact and its basis in expectancy is well established (Ross & Olson, 1981) and was reviewed in detail in the previous version of this chapter (Olson et al., 1996).

Summary

This section formed the conceptual heart of this chapter, in that our main argument is that expectancies work principally to guide effective behavior. Expectancies for success facilitate success, but they are especially likely to fuel behavioral progress once the individual has committed to a course of action. Individuals routinely anticipate future difficulty and proactively avoid it. These examples of expectancies working to create their own reality may also be problematic, as when self-fulfilling prophecies serve to perpetuate inaccurate and unfair prejudices. Through all of the aforementioned effects, expectancies exert effects on behavior as mediated by a variety of further cognitive processes, which are utterly critical for continued effectiveness of behavior in light of changing circumstances and learning.

COGNITIVE CONSEQUENCES

The cognitive consequences of expectancies have dominated past research, particularly with regard to their role in attention, encoding, representation, and memory. At a basic level, expectancies guide processing in a manner that is self-perpetuating. Once useful expectancies have developed, our cognitive system is rather conservative about altering or replacing them. Clearly, it would be dysfunctional to abandon effective knowledge too easily. Yet it would also be dysfunctional if expectancies were so stubborn to the facts as to be unrevisable in the face of contradiction (Piaget, 1952). As several theorists have noted, a cognitive system that is either too flexible or too stable would be at an evolutionary disadvantage (e.g., Johnston & Hawley, 1994; Sherry & Schacter, 1987; Tulving, Markowitsch, Kapur, Habib, & Houle, 1994). Thus, despite the generally conservative nature of expectancies, specialized processes exist to maintain vigilance for inaccuracy, to facilitate the encoding of unexpected events, to enhance the integration of those events into expectancies, and to improve memory for the events.

Expectancies are functional because they help to maximize the ratio of useful information gained for effort expended (e.g., Sherman, 2001; Sherman, Lee, Bessenoff, & Frost, 1998). This principle accounts for many of the specific ways in which expectancies influence attention, encoding, representation, and memory, particularly when processing resources are constrained. As resource-conserving devices, it is in these difficult conditions that the functional advantages of expectancies are most acute, and that expectancies are most likely to guide information processing. At the highest level of analysis, the cognitive consequences of expectancies are secondary and subservient to the primary function of behavior regulation. Expectancies are retained, discarded, or tweaked specifically to furnish more useful information to guide ongoing behavior. The cognitive processes surrounding expectancies, and in particular reactions to their confirmation or disconfirmation, represent tools for survival.

Expectancy Confirmation and Disconfirmation

Cognitive consequences hinge critically on whether incoming information confirms or disconfirms an expectancy. The majority of research on the cognitive consequences of expectancies has focused on differences in how expectancy-consistent versus expectancy-inconsistent information is processed. The experience of expectancy confirmation may be seen as the cognitive equivalent of the seafaring condition of "situation normal," meaning that all is well, no new crew action is required, and the currently operative behavior (speed, direction, degree of readiness) is satisfactory. For goal expectancies, situation normal corresponds to a regulatory loop in which current versus ideal conditions approximate, hence corrective action is presently unnecessary.

By contrast, a disconfirmed expectancy is the equivalent of "all hands on deck," meaning that the current situation represents the potential for danger and thus demands at the very least (1) heightened vigilance, but perhaps also (2) corrective action. Expectancies that are disconfirmed may also represent (3) inaccuracy, thus demanding some sort of conceptual repair work to restore or improve accuracy. These three aspects, vigilance, problem-solving, and belief repair, constitute the three primary imperatives of disconfirmed expectancies.

Thus, the functional significance of expectancy-consistent versus expectancy-inconsistent information is quite different. These functional considerations as well as the need for efficient processing largely determine the manner in which expected and unexpected information is attended to, encoded, represented, and remembered. We describe these processes in order of psychological events from initial seeking of and exposure to information, through encoding, representation, and memory.

Information Seeking

Expectancies influence the types of information perceivers seek in the environment. Often expectancies are essentially hypotheses about the world, and individuals seek information to test their validity. This occurs primarily for subjective rather than factual expectancies (which tend to be taken for granted). Far from being evenhanded, however, diverse research indicates a ten-

dency to seek out information that confirms rather than disconfirms the expectancy (Klayman & Ha, 1987; Lord, Ross, & Lepper, 1979; Skov & Sherman, 1986). Indeed, this bias extends perhaps also to the manner in which scientists test theories (MacCoun, 1998). A more detailed discussion of this pattern appeared in the Olson and colleagues (1996) chapter.

Processing Fluency

Exposure to expected or unexpected information influences implicit psychological responses within milliseconds. Several theories have converged on the idea that online processing—attention to and awareness of current experience—involves continuous pattern matching between the incoming sensory stream and semantically related information in memory (e.g., Lieberman et al., 2002; Srull & Wyer, 1989; Whittlesea, 1997). Processing fluency describes the extent to which this pattern matching flows smoothly or is interrupted by mismatches (i.e., expectancy disconfirmations) (e.g., Benjamin & Bjork, 1996; Johnston & Hawley, 1994; Whittlesea & Williams, 2001). The subjective experience of dysfluency (i.e., surprise) may embody mismatches involving perceptual (What should it look like?) or conceptual (What does it mean?) features (Jacoby & Dallas, 1981; Roediger, 1990; Whittlesea & LeBoe, 2003). In sum, expected stimuli produce subjective feelings of fluency and comfort, whereas unexpected stimuli produce feelings of dysfluency and surprise.

Processing fluency represents the first stage at which the cognitive system registers a confirmed or disconfirmed expectancy. The implicit perception of dysfluency is the starting point for the cascade of cognitive consequences that come next.

Attention

Though people may not actively seek out unexpected information, when exposed to it, they will generally attend carefully—more carefully than to expected information. The processing dysfluency that results from expectancy disconfirmation may act as a functional imperative to attend. Research indeed shows that attention is rapidly and automatically directed toward stimuli that are in some way surprising (e.g., Barthalow, Fabiani, Gratton, & Bettencourt, 2001). Interestingly, this principle forms the basis of the enormously successful *violation-of-expectation paradigm*, which has been used to reveal age of onset of particular representational categories in infants (Baillargeon, 2004). Indexed by visual gaze duration, infants who attend longer to an object or event are presumed to have been surprised and hence interested, thus revealing that their brain had in some way represented an expectancy regarding that object or event (Baillargeon, Spelke, & Wasserman, 1985; Wang, Baillargeon, & Brueckner, 2004; Wilcox, Nadel, & Rosser, 1996). This method has revealed that infants as young as 2½ months of age understand that their physical world consists of objects having continuity (i.e., they exist continuously in space and time) and solidity (i.e., two objects cannot occupy the same space at the same time) (Baillargeon, 2004; Spelke, Breinlinger, Macomber, & Jacobson, 1992).

An important moderator of these effects is the availability of processing resources. Under cognitive load, the tendency to attend more carefully to unexpected than expected information is enhanced. Indeed, in these conditions, attention actually shifts from expected and toward unexpected information in the visual field (Sherman et al., 1998). Differences in the conceptual fluency of expected and unexpected information may contribute to this effect. Because expected information is easily assimilated to existing knowledge and is therefore easily comprehended, little attention is required during encoding. The fit between the information and existing knowledge may be briefly noted, with attention then redirected to more novel and potentially important information in the environment. This attention shift is more likely to occur when capacity is depleted because it is under those conditions that the cognitive system is most pressed for efficiency. Such a process embodies both the stability (via conceptual fluency and pattern matching) and plasticity (by directing attention toward unexpected information) of expectancies.

Individual motivational differences also moderate these effects. People who are motivated to see group stereotypes as malleable (i.e., "incremental theorists") are especially likely to shift attention from stereotype-consistent and toward stereotype-inconsistent information when they were under cognitive load (Plaks, Stroessner, Dweck, & Sherman, 2001). In addition, prejudiced people who also are prevention oriented are more likely to attend to stereotype-violating information (Förster, Higgins, & Strack, 2000; Förster, Higgins, & Werth, 2004).

Interpretation

The most readily visible cognitive consequence of expectancies is their influence on how individuals see and understand the world around them. This influence comes in two forms. First, expectancies may act as heuristics in providing direct input into judgments. Generally speaking, a heuristic is a quick and resource-frugal judgment that is accurate often but not always. The use of expectancies as heuristics has been demonstrated in many domains of psychology. For example, in persuasion research, expectancies about source expertise (experts are to be trusted) often influence persuasion to a greater extent than systematic reasoning about the strengths and weaknesses of given arguments (e.g., Petty & Wegener, 1998). In stereotyping research, stereotypes are often relied on to make judgments about others to the exclusion of individual behaviors (e.g., Sherman et al., 2000). In both cases, the expectancies are particularly likely to drive judgments when people either are unmotivated to process carefully or are unable to do so (e.g., if they are under cognitive load).

The second way that expectancies influence an individual's understanding of the world is through their influence on interpretation. The idea that people see what

they expect to see, or interpret events and objects in a manner that assimilates experience to the expectancy, pervades numerous theoretical constructs in psychology, including notions of coding system, frame, schema, script, and stereotype. Much early research showed how concepts, inferences, or category labels influence subsequent interpretation (Bruner, 1957; Bruner et al., 1956; Darley & Gross, 1983; Hastorf & Cantril, 1954; Higgins, 1996; Kelly, 1955; Vallone, Ross, & Lepper, 1985; Wilson, Lisle, Kraft, & Wetzel, 1989); more recent research has shown similar effects with complex goal-oriented expectancies (Kawada, Oettingen, Gollwitzer, & Bargh, 2004) and in active seeking of self-verifying information (Swann, Stein-Seroussi, & Giesler, 1992). In general, this expectancy-assimilation effect is stronger to the extent that the expectancy is stronger and the stimuli more ambiguous (e.g., Alba & Hasher, 1983; Budescu, Kuhn, Kramer, & Johnson, 2002; Swann & Ely, 1984; Trope, 1986; Tuckey & Brewer, 2003).

Conceptual versus Perceptual Encoding

Information that fits expectancies is more easily understood. One consequence of this comprehension advantage, however, is that people do not attend carefully to expected information (e.g., von Hippel, Jonides, Hilton, & Narayan, 1993). A related consequence is that people do not encode the perceptual details (e.g., physical features) of expected information carefully. Rather, the basic conceptual gist of such information is extracted, but little else. In contrast, though unexpected information is often poorly comprehended, it is attended to carefully, and the physical details are encoded well (Sherman, Lee, et al., 1998). As with attention, the perceptual encoding advantage for incongruent information is greater when resources are low, again attesting to the flexible efficiency of expectancy use (e.g., Sherman, Conrey, & Groom, 2004). The careful encoding of the details of incongruent information is another means by which the cognitive system preserves plasticity in the face of expectancy confirmation. Retaining these details helps individuals reconstruct the facts surrounding unexpected events at a later time, when new information and greater resources may be available to help make sense of them. Indeed, differences in the conceptual and perceptual encoding of expected and unexpected events have significant influences on the manner in which these stimuli are represented in memory and subsequently remembered. These issues are addressed in detail below.

Coping with Disconfirmation

Disconfirmed expectancies are at root discrepancies between cognition and reality. They embody failures of prediction and thus constitute inaccuracy within the individual's cognitive model of reality. Such a failure may or may not demand a conceptual fix, but to ascertain which, effort is directed at investigating the nature and source of the failure and correcting it. Summarized next are those higher-order consequences of disconfirmed expectancies that are aimed at cognitive repair.

Disconfirmed Expectancies Evoke Processing That Is Resource Demanding

Whereas confirmed expectancies result in relatively automatic processing, disconfirmed expectancies recruit processing that is more effortful (Bargh & Thein, 1985; Stern, Marrs, Millar, & Cole, 1984; Wilson et al., 1989). Initial perceptual processing takes longer for unexpected than expected stimuli (Jentzsch & Sommer, 2002; Matt, Lethold, & Sommer, 1992). Overall, surprise demands deeper and more careful analysis of relevant information, aimed at explaining and understanding the predictive failure. This core principle linking expectancy disconfirmation to systematic processing echoes through several theories, for example, those aimed at attitudes (Chaiken, Liberman, & Eagly, 1989) and linguistic interpretation (Burgoon, 1993).

Disconfirmed Expectancies Activate Sense Making

Explanation and attribution correspond to an attempt to make sense of an outcome. Because disconfirmed expectancies may constitute danger (as in the case of avalanches, attacking muggers, or outgroup mobs), cognitive effort designed to make sense of them is an essential ingredient for guiding subsequent behavior (e.g., fleeing, fighting, and negotiating).

Sense making involves retrieving information from memory that forms the basis for new explanations (Ahn, Novick, & Kim, 1995; Kelley, 1967), which aim to bridge the gap between prior understanding and current experience (Ahn et al., 2003). Three classes of sense-making activity all reveal evidence of activation by disconfirmed expectancy. *Causal attribution* involves the most basic process of identifying the cause of a particular outcome (John stepped on Susan's feet because he lacks coordination); *counterfactual thinking* involves the more elaboratively narrative articulation of how the outcome might have come about had the key causal condition(s) been different (John might have been a better dancer with more practice in high school); and *hindsight bias* involves the metaperception of confidence that the outcome in question was sensible and predictable (I just knew John would step on Susan's feet). Evidence is most voluminous in support of the activation of causal reasoning by expectancy disconfirmation (e.g., Hastie, 1984; Kanazawa, 1992; Wong & Weiner, 1981); similar evidence is nevertheless available for counterfactual thinking (Roese & Olson, 1997; Sanna & Turley, 1996) and hindsight bias (Roese, 2004; Schkade & Kilbourne, 1991).

Sense Making Results in One of Four Inferential Products

IGNORING

In this first case, the sense-making activity uncovers few insights and little new information in memory or is in some way truncated. The result is that no inferential product becomes available, and the discrepancy is essentially ignored. Examples include the cognitive dissonance theory concept of trivialization (e.g., Simon, Greenberg,

& Brehm, 1995) and the motivated shallow processing of self-threatening hypothetical behaviors (Sedikides & Green, 2000).

TAGGING

Though sometimes sense-making failures will be ignored, in other cases, they will be tagged for future examination and use. As noted already, even if unexpected information cannot be clearly explained, it may receive considerable attention, and the details of the event may be encoded carefully. This allows people to reaccess this information at a later time when comprehension may be more successful. Moreover, unexpected events may be "tagged" onto existing expectancies, so that when expectancies are activated, so too are individual exceptions to those expectancies (e.g., Klein et al., 2002; McClelland et al., 1995). This helps to constrain the reach of expectancies and maintain their plasticity.

BRIDGING

In this third case, the sense-making activity focuses on erecting a conceptual bridge between the expectancy and the disconfirming event, in effect explaining away the discrepancy. Importantly, bridging adds new inferential information while preserving the integrity of the underlying schematic understanding. According to *cognitive dissonance theory* (Festinger, 1957), for example, discrepancies between cognitions may be resolved by adding new cognitions. A particularly common way of bridging discrepancies is by elaborating an exception to a general rule in terms of a subtype category. For example, attributions for disconfirmed interpersonal expectancies tend to focus on external and unstable rather than internal and stable causes (Crocker, Hannah, & Weber, 1983; Feather, 1969; Kulik, 1983). In the context of stereotyping, an explanation for an unexpected group-member behavior may be accomplished by noting a subtype, or new subclass of the stereotype, in which the surprising act is explained in terms of an exception to the general rule embodied in the stereotype (Hewstone, Johnston, & Aird, 1992; Kunda & Oleson, 1995, 1997). These various sense-making efforts conspire to preserve the integrity of the extant expectancy in light of disconfirming evidence, but importantly, they also point to the progressive creation of successively more detailed, multifaceted, and flexible representations of the domain in question. Thus, original expectancies are maintained, but their generality of application is reduced. In short, this bridging principle is an indicator of the deeper, functional process by which schematic knowledge is elaborated in light of ongoing experience, thereby providing successively more effective guides for subsequent behavior. People learn from mistakes.

REVISING

In this fourth case, the discrepancy prompts a reassessment and revision to the original conceptual underpinnings of the expectancy. Unlike bridging, which preserves the integrity of the underlying schema, revising involves changes to the underlying schema at a foundational level. For example, if one goes to McDonald's with the expectation of dining on filet mignon, the magnitude of the error demands a reworking of the essential informational components of the expectancy, along with, perhaps, expansion of the knowledge base supporting it (elaboration of the categorical distinction between fast food and fine dining; gathering new insights from restaurant reviews, etc.). This process has been termed "conversion" in some writings (Hewstone et al., 1992; Piaget, 1952; Rothbart, 1981; Weber & Crocker, 1983), but our emphasis includes not only the dramatic shift in schematic valence denoted by conversion but also the valence-neutral expansion of schematic detail exemplified by the development of expertise (Tanaka & Taylor, 1991), sometimes referred to as *bookkeeping* (Rothbart, 1981).

Discrepancy Magnitude and Schema Complexity Determine the Inferential Products of Sense Making

Theorists have pointed to two main determinants of which type of inferential product tends to emerge from the sense-making activity evoked by a disconfirmed expectancy: the magnitude of the discrepancy between expectancy and outcome and the degree of complexity or sophistication of the underlying schematic basis of the expectancy. For the first determinant, discrepancy magnitude, small discrepancies will be more likely to be ignored than larger discrepancies. Large discrepancies typically result in subtyping processes, whereby separate, specialized subcategories are created to account for the discrepant stimuli or events. In contrast, moderate discrepancies tend to result in slow and steady expectancy revision (e.g., Rothbart & Lewis, 1988; Weber & Crocker, 1983).

Schema complexity, or the degree to which knowledge about the relevant domain is developed, also plays a role. With new schemas, as exemplified by low expertise (e.g., a novice squash player trying to grasp the intricacies of the game), initial attempts at reaching a coherent understanding result in an emphasis on confirmatory search and openness to divergent new information. In this case, the emphasis is on developing useful inferential tools rather than with testing the boundaries of those tools. Moreover, it is more difficult to note discrepancies from weak expectancies in the first place. With greater development of the underlying schema, discrepancies between expectancy and experience are more likely to be noticed and tend to be processed more deeply, resulting in bridging effects (cf. Srull, Lichtenstein, & Rothbart, 1985; Tanaka & Taylor, 1991). It is probably also the case that weak expectancies are more likely to involve revision, but as expertise, accuracy, and certainty grow, both disconfirmations as well as revisions become far less likely (Karniol, 2003). Those disconfirmations that do occur involve mainly bridging (subtyping, in particular).

Representation

Differences in the ways that expected and unexpected events are encoded affect the manner in which the information is subsequently represented in memory. Expectancy-congruent information is not attended to carefully and the details are not thoroughly encoded (e.g., Sherman, Lee, et al., 1998). Rather, the basic conceptual gist meaning is extracted via assimilation to prior expectancies, and little else is retained. Thus, expected events are likely to be retained primarily in abstract, semantic form. In contrast, unexpected events are attended to carefully and the details are more thoroughly encoded. In part, this is simple necessity in that such events cannot be well understood in light of existing knowledge. Accordingly, unexpected events are more likely stored as detailed, context-specific episodes (Sherman, Klein, Laskey, & Wyer, 1998).

Memory

Memory for expected versus unexpected stimuli differs in a number of respects. These differences result from the ways in which congruent and incongruent information is attended to, encoded, represented, and retrieved. The differences are moderated by several important variables.

Recall and Recognition

On measures of both free recall and recognition, memory is superior for unexpected than expected events (e.g., Alba & Hasher, 1983; Stangor & McMillan, 1992). Even when a congruent event can be remembered, its source may not be. For example, stereotypical behaviors are often falsely attributed to people who did not commit them (e.g., Mather, Johnson, & De Leonardis, 1999; Sherman & Bessenoff, 1999).

Such effects center mainly on encoding processes. As noted, incongruent events draw our attention, and because they challenge extant beliefs, individuals expend effort toward explaining them (e.g., *bridging*). This deeper, more elaborate encoding of incongruencies increases their memorability in a number of ways (e.g., Craik & Lockhart, 1972). First, the relatively greater amount of time perceivers spend considering these events in working memory increases their general accessibility (e.g., Higgins, 1996). Second, the attention and detail given to encoding the item-specific features of these events increases the likelihood that the features can be used subsequently as retrieval cues (e.g., Einstein & Hunt, 1980). Third, in trying to make sense of incongruent events, they become associated with other information in memory, creating a wider network of pathways through which they may be retrieved (e.g., Srull & Wyer, 1989). Fourth, the attention to detail and sense making associated with encoding unexpected events increases the likelihood that they will be stored episodically, with details intact (e.g., Klein et al., 2002; McClelland et al., 1995; Sherman, 2001). In contrast, congruent events receive relatively little attention and elaboration and are

likely to be represented abstractly, making it difficult to remember them accurately.

Response Bias, Search Strategies, and Familiarity Increase True and False Memory of Expected Information

Estimates of memory for expected events are often inflated by response biases to report congruent events. Thus, on a recognition test, for example, correct recognition of congruent items may be quite high. However, the high incidence of false alarms on these items indicates that performance is driven largely by a bias to respond positively, rather than by accurate memory. Because congruent items (including foils) fit well with general expectancies, perceivers set low thresholds for claiming their verity (e.g., Stangor & McMillan, 1992). These same biases also appear to account, in part, for stereotypical biases in source memory (Spaniol & Bayen, 2002). The implications for eyewitness testimony are significant.

False memories of expectancy-congruent events may also be based on a feeling of processing fluency associated with those events. Because they fit with expectancies, these events may feel familiar even if they did not actually occur. This familiarity may be misinterpreted as being due to prior exposure, rather than to expectancy fit, leading people to falsely judge typical events as likely to have occurred (e.g., Greenwald & Banaji, 1995; Sherman et al., 2004).

Both true and false memory for expected information is also facilitated by retrieval strategies that take advantage of our expectancies. Thus, even though congruent events may be poorly encoded, they may be remembered when people use their expectancies to generate events that may have occurred. Often, when people are asked to remember what happened, they may reframe the question to themselves as "What is likely to have happened?," leading to a hypothesis-confirming search for expectancy-congruent information (e.g., Hirt, 1990; Hirt, Erickson, & McDonald, 1993). This same process may lead to the construction of typical events that did not occur.

Moderators of Memory Effects

One important moderator of all the effects described earlier is expectancy strength. As expectancies become clearer, more coherent, or more focused, incongruent events are increasingly surprising, but by the same token congruent events are increasingly taken for granted (e.g., Srull et al., 1985). The associated increased disparity in the extent to which expected and unexpected events are attended to, encoded, and stored episodically affects memory for the events. Thus, the advantage in accurate memory for incongruent events increases with expectancy strength. So, too, do response biases, feelings of familiarity, and expectancy-driven search strategies increase for congruent events (e.g., Sherman & Frost, 2000; Sherman, Klein, et al., 1998).

A second moderator of memory effects is the level of processing capacity available during encoding. When under cognitive load, people tend to attend more care-

fully to and encode the details of incongruent than congruent information. Accordingly, the memory advantage for unexpected events is greater when capacity has been restricted during encoding (e.g., Sherman & Frost, 2000). At the same time, diminished capacity increases the extent to which people rely on guessing strategies, feelings of familiarity, and biased search strategies in remembering typical or expected events. The increased reliance on these factors may increase both true and false memories for congruent events.

Summary

This section summarized the numerous cognitive consequences of expectancies. We attempted to integrate the general offshoots of expectancies, such as biased hypothesis testing, the expectancy heuristic, and expectancy-assimilative interpretation effects, with the more specific offshoots of expectancy confirmation versus disconfirmation, including effects in processing fluency, attention, interpretation, encoding, sense making, representation, and memory. All these cognitive consequences are aimed at extracting useful information from experience for the specific purpose of guiding subsequent effective action.

AFFECTIVE CONSEQUENCES

Affective responses to expectancies may be viewed as regulatory signals regarding goal progress, with positive affect signaling sufficient and negative affect signaling insufficient progress (as indicated by smaller vs. larger discrepancies, respectively, between expected and current status within a regulatory feedback loop). As such, affect constitutes an informational signal intrinsic to behavior regulation (Schwarz, 1990). This section elaborates on this idea, but also touches on optimism as an instance of affect regulation, and on broader affective consequences that have received considerable research attention, including attitudes, aesthetics, humor, and depression.

Behavior-Oriented Affective Consequences

Negative Affect Fuels Behavior Change

The immediate default response to a disconfirmed expectancy is negative affect (Mandler, 1975; Olson et al., 1996). This primary affective consequence is best understood with regard to its implications for behavior regulation. Specifically, negative affect spurs greater behavioral effort aimed at problem solving (McDonald & Hirt, 1997; Schwarz, 1990; Taylor, 1991), as mediated by the cognitive consequences reviewed in the previous section. Studies of regret and disappointment have been particularly revealing as indications of how negative affect spurs alterations in behavior (Zeelenberg, 1999). This basic notion also extends through the *investment model* of interpersonal relationships, in which a relationship that is perceived to fall short of a generic expectation of relationship quality (termed a "comparison level") results in negative affect (disappointment) which in turn predicts relationship dissolution (Rusbult et al., 1998; Thibaut & Kelley, 1959).

Negative Affect Is the Default Response to Processing Fluency Disruption

On a more basic level, the disruption of processing fluency is experienced as affectively unpleasant (Reber, Winkielman, & Schwarz, 1998; Winkielman & Cacioppo, 2001). It must be emphasized that we are referring to the initial and default affective response to processing dysfluency; this affective response is by no means the only or the typical response. As detailed subsequently, inferential processing creates secondary affective responses, which may vary widely in valence as a function of the specific form of interpretational attribution (e.g., Whittlesea & Williams, 2001).

Expectancy Disconfirmation Shifts Evaluation via Contrast Effect

Reactions to success and failure hinge not only on the intrinsic quality of the outcome but also on how the outcome is framed by expectations. A negative outcome creates dissatisfaction when it is expected but is even more extremely dissatisfying when unexpected. By the same token, a positive outcome may taste sweet if expected but may be all the sweeter when it takes the perceiver by surprise.

The underlying mechanism for such effects is the perceptual contrast between the expectancy and outcome (Mellers, Schwartz, Ho, & Ritov, 1997; Roese, 1997). Specifically, an outcome that disconfirms an expectancy can be evaluated in part by comparison to the expected yet unattained outcome (i.e., a counterfactual comparison). The juxtaposition of an alternative that is either better (an upward counterfactual) or worse (a downward counterfactual) renders the evaluation of the factual outcome either more negative or more positive, respectively (Roese, 1994). By this analysis, the evaluative direction between the obtained outcome and its counterfactual alternative matters most in determining affect (Barthalow et al., 2002; Bettencourt, Dill, Greathouse, Charlton, & Mullhuland, 1997; Feather, 1969; Shepperd & McNulty, 2002). Hence, an unexpected failure evokes more negative affect than an expected failure by virtue of the contrast to an upward counterfactual (i.e., the expected, more positive outcome), and an unexpected success evokes more positive affect than an expected success because of the contrast to a downward counterfactual (i.e., the expected, less positive outcome). New evidence suggests a somewhat different interpretation, that both positive outcomes (that could have been better) and negative outcomes (that could have been worse) may be characterized not so much as contrast-effect shifted unitary affect experience but rather as an ambivalent experience of simultaneous mixed emotions (Larsen, McGraw, Mellers, & Cacioppo, 2004).

That expectancies shape evaluation is a rule of thumb assumed by many, as exemplified by the tactic of "lower-

ing the bar." For example, prior to the 2000 presidential election debates, candidate George W. Bush's handlers tried to "spin" low expectations for his performance, thereby ensuring that nearly any performance by Bush would exceed expectations and thus enhance voter satisfaction via a downward counterfactual comparisons ("He could have performed so much worse . . . "). Research indeed suggests that individuals sometimes strategically reduce expectancies of success (Shepperd et al., 2000; van Dijk, Zeelenberg, & van der Pligt, 2003). Corporate earnings statements have similarly been shown to be manipulated in such a way that they just barely exceed (but rarely just barely miss) earlier forecasts of performance, thereby ensuring that downward rather than upward counterfactuals are the most salient consequence (Burgstahler & Dichev, 1997).

Optimism and Affect Regulation

People are, on average, optimistic: When looking to their own future, they see more positives than negatives (Newby-Clark & Ross, 2003). In pondering future career moves, romantic encounters, financial deals, or vacations with family, people expect events to go well and dedicate relatively little thought to negative possibilities. Such optimism has been documented against several benchmarks, the most common of which is an interpersonally relative judgment in which individuals predict the likelihood of good and bad events befalling them compared to similar others. This form of judgment has often been called *unrealistic optimism* (Radcliffe & Klein, 2002; Weinstein, 1980; Weinstein & Klein, 1996). Another benchmark is the individual's current circumstance, and here again there is a general tendency to expect future improvement relative to current state; this judgment has been termed "upward temporal comparison" (Heckhausen & Krueger, 1993; McFarland & Alvaro, 2000; Wilson & Ross, 2001). Yet a third benchmark is actual outcomes, and again individuals tend to expect greater riches and rewards than objectively occur at a later time (Mitchell, Thompson, Peterson, & Cronk, 1997; Wirtz, Kruger, Scallon, & Diener, 2003). A variation of this observation has been termed the "planning fallacy": In judging the future accomplishment of personal or professional goals, projects, or assignments, people expect to complete more and in a briefer period of time than actually turns out to be the case (Buehler, Griffin, & Ross, 1994; Kruger & Evans, 2004; Newby-Clark, Ross, Boehler, Kohler, & Griffin, 2000). Although people are on average optimistic, variation across individuals (e.g., Norem & Cantor, 1986) and cultures certainly exists (e.g., Chang, Asakawa, & Sanna, 2001; Heine & Lehman, 1995), and one recent argument was that variability in optimism constitutes perhaps the most basic of personality factors (Haugen, Ommundsen, & Lund, 2004).

Why are people unrealistically optimistic? Several explanations have been offered, such as those rooted to an egocentric emphasis on self-relevant information (e.g., Karniol, 2003; Kruger & Burrus, 2004). But another ex-

planation running through several prominent theories is that optimistic expectancies produce positive affect, and therefore individuals are motivated to create them. According to this view, people are optimistic simply because it feels good to imagine a more positive future. Accordingly, people use optimistic expectancies for affect regulation; for example, optimism increases in a compensatory fashion in response to threat (McFarland & Alvaro, 2000). In the *theory of positive illusions* (Armor & Taylor, 1998; Taylor & Brown, 1988), unrealistic optimism was one of several self-aggrandizing cognitions that, as long as they remain moderate rather than extreme distortions of reality, serve to facilitate psychological well-being (Kaiser, Major, & McCoy, 2004; Scheier, Carver, & Bridges, 2001), aid in coping with misfortune (Aspinwall & Taylor, 1997), and even enhance physical health (Peterson & Bossio, 2001), as indexed, for example, by immune system functioning (Segerstrom, Taylor, Kemeny, & Fahey, 1998; Taylor, Lerner, Sherman, Sage, & McDowell, 2003). Benefits of optimism for productivity, persistence, and life satisfaction have also been observed (e.g., Gilham, 2000; Seligman, 1998). The ultimate negative expectancy is death, and according to *terror management theory* (Greenberg, Solomon, & Pyszczynski, 1997), recognition of one's own mortality unleashes a range of processes aimed at keeping this horrific expectancy cloaked in obscurity. The theory suggests that many beliefs and behaviors are aimed at providing distraction from mortality awareness; examples include religion, political ideology, nationalist or other ingroup identification, and even sexual ritual (e.g., Goldenberg, Pyszczynski, Johnson, Greenberg, & Solomon, 2000). All considered, the range of evidence to support the claim that optimism is motivated, or is sometimes recruited for affect regulation, is substantial. Contemporary theory moreover emphasizes the deep interplay (rather than opposition) between motivated and purely informationally based mechanisms (e.g., Kruglanski, 1996; Kunda, 1990).

Under the rubric of *affective forecasting*, the concept of *impact bias* has been used to describe people's tendency to exaggerate the emotional impact of future events (Wilson & Gilbert, 2003). In this literature, the focal judgment is magnitude of emotional consequence following a specified event, operationalized in terms of both intensity and duration of affective experience. Bias is defined in terms of a comparison between predicted and actual emotional experiences (Buehler & McFarland, 2001; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998); for both positive and negative emotions, people expect greater emotional intensity and duration than actually transpires. At first glance, findings on affective forecasting with regard to negative emotions seem to contradict the standard optimism finding: Research on affective forecasting suggests that people exaggerate the emotional impact of negative future events, whereas research on unrealistic optimism indicates that people skew their future view in a positive direction. This contradiction may be explained by recognizing that these two literatures, in tapping separate judgments (emotional impact

and event likelihood, respectively) have differentially emphasized two sides of the same regulatory balancing act. When viewed together, impact and likelihood judgments are negatively correlated: The more emotionally consequential the event, the more unlikely individuals deem it to be (Blanton, Axsom, McLive, & Price, 2001). The most threatening expectancy would be one in which a possible negative event (losing a job, losing a spouse) is deemed both very impactful ("It would really wipe me out") and highly likely. Individuals forced to confront the prospect of a high-impact negative event may thus strategically reduce its perceived likelihood ("Sure that would be awful, but it'll never happen to me"). Similarly, if individuals realize that a negative event is likely, they may strategically reduce its perceived impact ("I know it might happen, but it won't be so bad"; Kay, Jimenez, & Jost, 2002). Such a balancing act between judged impact and likelihood is a further instantiation of the principle that individuals regulate expectancies to defend against affectively threatening inferences (Jemmott, Ditto, & Croyle, 1986).

To summarize, people are on average optimistic when they look to their own personal future, and this optimism brings about positive affective consequences, which in turn bring a variety of further beneficial consequences for health and performance.

Broader Affective Consequences

The previously noted principles center on the immediate emotional offshoots of expectancy confirmation and disconfirmation. In the following examples of attitudes, aesthetics, humor, and depression, the importance of the expectancy construct may be seen across a broader range of psychological effects.

Attitudes Reflect the Intersection of Expectancy and Value

Attitudes involve evaluations of people, objects, or issues, and a popular means of conceptualizing them involves the varying contributions of information, prior behavior, and affect (Breckler, 1984). The expectancy-value approach to attitudes (and judgment more generally) is that an attitude is the summation of a set of beliefs about the value of particular properties along with the expected likelihood that the attitude object contains those properties (Ajzen & Fishbein, 1980). A positive attitude toward an automobile built by Volkswagen, for example, would be the result of a set of value and expectancy beliefs regarding attributes such as price (low price expected with high certainty), quality (high reliability expected with moderate certainty), and driving experience (solid handling expected with some uncertainty). Although this perspective has proven useful, a recent controversy has centered on whether statistically inappropriate techniques have rendered many past findings ambiguous, and which techniques are best suited for new research (French & Hankins, 2003; Haddock & Zanna, 1998).

Aesthetic Appreciation Hinges on Moderate Expectancy Disconfirmation

Although critics will debate the merits of great versus mediocre art along innumerable dimensions, psychologists have long noted that aesthetic appreciation at its simplest level reflects moderate surprise. That is, new art is appreciated by most individuals when it involves a moderate, neither weak nor extreme, deviation from expectancies. As revealed in early psychophysical studies of taste, brightness, heat, and form, previous experience creates an adaptation level or habituation point, against which experiential discrepancies are perceived in affective terms (Haber, 1958; Helson, 1964). Later studies of aesthetic judgments of painting, music, and literature confirmed that very small deviations from past experience are perceived as boring; large deviations as unpleasantly bizarre, but moderate deviations as sweetly intriguing (Berlyne, 1974). Reber, Schwarz, and Winkielman (2004) argued processing fluency to be a key mechanism underlying the relation between expectancy and aesthetic appreciation. These authors proposed that the more fluently the perceiver can process an object, the more positive the resulting aesthetic response. Small deviations from expectancy may well be processed fluently, but perhaps so fluently as to barely register as interesting, blending instead into the perceptual background. Large deviations from expectancy, on the other hand, interrupt fluency by stimulating deeper cognitive processing aimed at explaining the discrepancy. Between these two extremes are intermediate deviations that stimulate simpler (and more successful) explanations that rapidly restore fluency to its prior, faster rate. In this regard, it is perhaps not so much greater fluency in an absolute sense that evokes aesthetic appreciation but, rather, the contrast effect resulting from a rapid shift from relatively low to high fluency that creates the subjective feeling of aesthetic pleasure (cf. Whittlesea & Williams, 2001).

Humor Derives from Resolution of Incongruity

The experience of humor may be taken as a special case of the more general principle guiding aesthetic appreciation covered in the previous section. Humor appears to be the product of a two-step process: a surprise based on some sort of incongruity (e.g., between a particular utterance and the expected word usage; between a "funny walk" vs. the typical way of walking), followed very rapidly by an inferential resolution that renders the incongruity nonthreatening (Suls, 1983; Wyer & Collins, 1992). As explored in greater detail in the precursor to this chapter (Olson et al., 1996), the greater the initial surprise and the more satisfyingly comprehensive the resolution, the bigger the laughs.

Depression Involves Hopelessness Expectancies

Depression is a mental disorder marked by extreme negative affect, demotivation, and behavior deficits. A signature symptom of major depression is global expectancies

of future negative events, especially events that are personally important to the individual, and has been specified in detail within the *hopelessness theory of depression* (Abramson et al., 1989). The correspondence between biased negative expectancies and behavioral dysfunction (e.g., job performance deficits and interference in interpersonal relationships) marked by depression stands in contrast to biased positive expectancies that facilitate the behavior among healthy individuals, thus further underscoring the functionality of expectancies modestly biased toward optimism (Taylor & Brown, 1988).

Summary

This section reviewed three main classes of affective consequences of expectancies, with a particular emphasis on consequences of expectancy disconfirmation. First, affective responses to expectancies were considered in terms of their role as regulatory signals regarding goal progress. Second, expectancies for personal events tend to be optimistic, which in turn brings about positive affective consequences. Third, broader affective consequences were also discussed, with a spotlight placed on attitudes, aesthetics, humor, and depression.

CONCLUSION

To predict the future is to navigate it more effectively. The psychological literature on the construct of expectancy is enormous, yet we argue that the numerous findings uncovered by this literature are most parsimoniously described with regard to behavior regulation, effective action, and survival. We distinguished between semantic and episodic expectancies, arguing that both facilitate performance but in different ways. Semantic expectancies are summaries of multiple prior experiences; they tend to be relatively abstract, implicit, and efficiently deployed. Most expectancies are of this semantic sort: They provide a wealth of general background knowledge and "common sense" that silently guide construal and behavior with efficient accuracy. Episodic expectancies are derived from memories of particular past instances; they tend to be relatively concrete and explicit and provide deeper, more specific information.

Whereas semantic expectancies provide implicit and mundane guidelines (a clear morning in July is likely to become a hot day, so dress lightly . . .), consciously held episodic expectancies may range from small to large plans, from tonight's dinner to next summer's wedding. Episodic expectancies for desired goals may not only guide behavior but also facilitate performance, either by energizing motivation via hopeful imagery or by making specific step-by-step behavioral requirements salient. Anticipation of problems and obstacles results in proactive avoidance behavior. Such examples of expectancies creating their own reality are deeply useful to human beings, yet the potential for mishap, as when negative and inaccurate stereotypes fuel self-fulfilling prophecies, springs from the same basic mechanisms.

The usefulness of an expectancy in guiding behavior depends on its accuracy; hence expectancies must be revisable in light of disconfirmation. At the same time, overly capricious sensitivity to nonrepresentative or nondiagnostic disconfirmation would impair rather than facilitate expectancy accuracy. Generally speaking, disconfirmed expectancies demand attention and require cognitive capacity to support efforts at explanation and understanding, the inferential products of which result in one of four consequences: ignoring the discrepancy, tagging the discrepancy, bridging the discrepancy with new insight, or revising the expectancy on a deeper level. Although we have argued that expectancies tend to be accurate overall, this assertion masks a complicated set of conceptual issues (Judd & Park, 1993; Kruglanski, 1989). These issues were explored in depth in the precursor chapter by Olson and colleagues (1996); that discussion still stands as an authoritative overview of this difficult subject.

Expectancy is one of the bedrock constructs in the field of psychology. Although numerous areas of psychology use the expectancy construct, from developmental to clinical psychology, from animal conditioning to cognitive psychology, this chapter emphasized the voluminous research deriving from social psychology. We have attempted to conceptualize the expectancy construct in terms of its functional basis in effective behavior control, positioning cognitive and affective consequences as secondary to and supportive of this main function. Yet as we write, social psychology is evolving rapidly to embrace new advances in brain imaging technology. With the emerging discipline of social cognitive neuroscience gaining momentum in pinpointing brain structure associated with specific cognitive function, we have constructed our review with an eye to providing a function-oriented roadmap for this new research. Indeed, we have noted one central and highly promising mapping of function to structure in the form of a processing dysfluency detector localized at the anterior cingulate. This is an exciting development, but merely the beginning of what promises to be a sharp expansion of our understanding of the deep relation between brain structure and cognitive function, an expansion we anticipate eagerly.

NOTE

1. The origin of the term "pipeline" can be traced to the seminal paper by Jones and Sigall (1971), in which they observed that psychologists have long fantasized "about discovering a direct pipeline to the soul (or some nearby location)" (p. 349).

REFERENCES

Abela, J. R. Z., & Seligman, M. E. P. (2000). The hopelessness theory of depression: A test of the diathesis–stress component in the interpersonal and achievement domains. *Cognitive Therapy and Research, 24,* 3–18.

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358–372.

Ahn, W., Novick, L., & Kim, N. S. (2003). Understanding it makes it more normal. *Psychonomic Bulletin and Review, 10,* 746–752.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice Hall.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203–231.

Anderson, C. A. (1983). Imagination and expectation: The effect of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology, 45,* 293–305.

Armor, D. A., & Taylor, S. E. (1998). Situated optimism: Specific outcomes and expectancies and self-regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 309–379). San Diego, CA: Academic Press.

Aspinwall, L. G., & Taylor, S. E. (1997). A stitch in time: Self-regulation and proactive coping. *Psychological Bulletin, 121,* 417–436.

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin, 120,* 338–375.

Baillargeon, R. (2004). Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. *Developmental Science, 7,* 391–424.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5-month old infants. *Cognition, 20,* 191–208.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goals revisited. *Journal of Applied Psychology, 88,* 87–99.

Bargh, J. A., & Ferguson, M. L. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126,* 925–945.

Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology, 49,* 1129–1146

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197–204.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Oxford, UK: Macmillan.

Battle, E. S. (1965). Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology, 2,* 209–218.

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 680–740). New York: McGraw-Hill.

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Metacognition and implicit memory* (pp. 309–338). Mahwah, NJ: Erlbaum.

Berlyne, D. E. (Ed.). (1974). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation.* New York: Taylor & Francis.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain responses to reward. *Journal of Neuroscience, 21,* 2793–2798.

Bettencourt, B. A., Dill, K. E., Greathouse, S., Charlton, K., & Mullhuland, A. (1997). Evaluations of ingroup and outgroup members: The role of category-based expectancy violation. *Journal of Experimental Social Psychology, 33,* 244–275.

Blanton, H., Axsom, D., McClive, K. P., & Price, S. (2001). Pessimistic bias in comparative evaluations: A case of perceived vulnerability to the effects of negative life events. *Personality and Social Psychology Bulletin, 27,* 1627–1636.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Brandstätter, V., Lengfelder, A., & Gollwitzer, P. M. (2001). Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology, 81,* 946–960.

Breckler, S. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology, 47,* 1191–1205.

Brehm, J. (1966). *A theory of psychological reactance.* New York: Academic Press.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking.* Oxford, UK: Wiley.

Bryant, F. B., & Brockway, J. H. (1997). Hindsight bias in reaction to the verdict in the O. J. Simpson criminal trial. *Basic and Applied Social Psychology, 19,* 225–241.

Budescu, D. V., Kuhn, K. M., Kramer, K. M., & Johnson, T. (2002). Modeling certainty equivalents for imprecise gambles. *Organizational Behavior and Human Decision Processes, 88,* 748–768.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67,* 366–381.

Buehler, R., & McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin, 27,* 1480–1493.

Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *Journal of Language and Social Psychology, 12,* 30–48.

Burgstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24,* 99–126.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science, 304,* 1167–1170.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280,* 747–749.

Carver, C. S., Blaney, P. H., & Scheier, M. F. (1979). Reassertion and giving up: The interactive role of self-directed attention and outcome expectancy. *Journal of Personality and Social Psychology, 37,* 1859–1870.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* New York: Cambridge University Press.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chang, E. C., Asakawa, K., & Sanna, L. J. (2001). Cultural variation in optimistic and pessimistic bias: Do Easterners really expect the worse and Westerns really expect the best when predicting future life events? *Journal of Personality and Social Psychology, 81,* 476–491.

Chapman, G. B. (1996). Temporal discounting and utility of health and money. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22,* 771–791.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling nature of automatically-activated stereotypes. *Journal of Experimental Social Psychology, 33,* 541–560.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. (2005). Separating multiple processes in implicit social cognition: The Quad-Model of implicit task performance. *Journal of Personality and Social Psychology, 89,* 469–487.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Crawford, M. T., McConnell, A. R., Lewis, A. C., & Sherman, S. J. (2002). Reactance, compliance, and anticipated regret. *Journal of Experimental Social Psychology*, *38*, 56–63.

Crocker, J., Hannah, D. B., & Weber, R. (1983). Person memory and causal attributions. *Journal of Personality and Social Psychology*, *44*, 55–66.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 968–986.

DeSteno, D., Petty, R. E., Wegener, D. T., & Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion specificity. *Journal of Personality and Social Psychology*, *78*, 397–416.

Dougherty, M. R. P., Gettys, C. F., & Thomas, R. P. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes*, *70*, 135–148.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.

Downey, G., Freitas, A. L., Michaelis, B., & Khouri, H. (1998). The self-fulfilling prophecy in close relationships: Rejection sensitivity and rejection by romantic partners. *Journal of Personality and Social Psychology*, *75*, 545–560.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, *58*, 568–581.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt, Brace Jovanovich.

Eagly, A. H., Wood, W., & Diekman, A. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.

Einstein, G. O., & Hunt, R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 588–598.

Erez, A., & Isen, A. M. (2002). The influence of positive affect on the components of expectancy motivation. *Journal of Applied Psychology*, *87*, 1055–1067.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 161–202). San Diego, CA: Academic Press.

Feather, N. T. (1966). Effects of prior success and failure on expectations of success and subsequent performance. *Journal of Personality and Social Psychology*, *3*, 287–298.

Feather, N. T. (1969). Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. *Journal of Personality and Social Psychology*, *13*, 129–144.

Feather, N. T., & Saville, M. R. (1967). Effects of amount of prior success and failure on expectations of success and subsequent task performance. *Journal of Personality and Social Psychology*, *5*, 226–232.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fischhoff, B. (1976). The effect of temporal setting on likelihood estimates. *Organizational Behavior and Human Performance*, *15*, 180–194.

Fiske, S. T. (2003). Five core social motives, plus or minus five. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 233–246). Mahwah, NJ: Erlbaum.

Förster, J., Higgins, E. T., & Strack, F. (2000). When stereotype disconfirmation is personal threat: How prejudice and prevention focus moderates incongruency effects. *Social Cognition*, *18*, 178–197.

Förster, J., Higgins, E. T., & Werth, L. (2004). How threat from stereotype disconfirmation triggers self-defense. *Social Cognition*, *22*, 54–74.

French, D. P., & Hankins, M. (2003). The expectancy-value muddle in the theory of planned behaviour—And some proposed solutions. *British Journal of Health Psychology*, *8*, 37–55.

Gigerenzer, G., Todd, P. M., & ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford, UK: Oxford University Press.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science*, *15*, 346–350.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *59*, 617–638.

Gillham, J. E. (Ed.). (2000). *The science of optimism and hope: Research essays in honor of Martin E. P. Seligman*. Radnor, PA: Templeton Foundation Press.

Gilovich, T., Kerr, M., & Medvec, V. H. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology*, *64*, 552–560.

Goldenberg, J. L., Pyszczynski, T., Johnson, K. D., Greenberg, J., & Solomon, S. (2000). Fleeing the body: A terror management perspective on the problem of human corporeality. *Personality and Social Psychology Review*, *4*, 200–218.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Gollwitzer, P. M., & Kinney, R. F. (1989). Effects of deliberative and implemental mind-sets on the illusion of control. *Journal of Personality and Social Psychology*, *56*, 531–542.

Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *18*, 319–332.

Grant, H., & Higgins, E. T. (2003). Optimism, promotion pride, and prevention pride as predictors of quality of life. *Personality and Social Psychology Bulletin*, *29*, 1521–1532.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 61–139). San Diego, CA: Academic Press.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about self and others. *Journal of Personality and Social Psychology*, *59*, 1128–1139.

Haber, R. N. (1958). Discrepancy from adaptation level as a source of affect. *Journal of Experimental Psychology*, *56*, 370–375.

Haddock, G., & Zanna, M. P. (1998). On the use of open-ended measures to assess attitudinal components. *British Journal of Social Psychology*, *37*, 129–149.

Hamilton, D. L., & Mackie, D. M. (1990). Specificity and generality in the nature and use of stereotypes. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 3, pp. 99–110). Hillsdale, NJ: Erlbaum.

Hastie, R. (1984). Causes and consequences of causal attribution. *Journal of Personality and Social Psychology*, *46*, 44–56.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology*, *49*, 129–134.

Haugen, R., Ommundsen, Y., & Lund, T. (2004). The concept of expectancy: A central factor in various personality dispositions. *Educational Psychology*, *24*, 43–55.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgment of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327.

Heckhausen, J., & Krueger, J. (1993). Developmental expectations for the self and most other people: Age grading in three functions of social comparison. *Developmental Psychology*, *29*, 539–548.

Heine, S. J., & Lehman, D. R. (1995). Cultural variation in unrealistic optimism: Does the West feel more invulnerable than the East? *Journal of Personality and Social Psychology*, *68*, 595–607.

Helson, H. (1964). *Adaptation-level theory*. New York: Harper & Row.

Hewstone, M., Johnston, L., & Aird, P. (1992). Cognitive models of stereotype change: Perceptions of homogeneous and heterogeneous groups. *European Journal of Social Psychology*, *22*, 235–249.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (2000). Social cognition: Learning about what matters in the social world. *European Journal of Social Psychology*, *30*, 3–39.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hirt, E. R. (1990). Do I see only what I expect? Evidence for an expectancy-guided retrieval model. *Journal of Personality and Social Psychology*, *58*, 937–951.

Hirt, E. R., Erickson, G. A., & McDonald, H. E. (1993). Role of expectancy timing and outcome consistency in expectancy-guided retrieval. *Journal of Personality and Social Psychology*, *65*, 640–656.

Hirt, E. R., Kardes, F. R., & Markman, K. D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *Journal of Experimental Social Psychology*, *40*, 374–383.

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, *69*, 1069–1086.

Irwin, F. W. (1944). The realism of expectations. *Psychological Review*, *51*, 120–126.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.

Jemmott III, J. B., Ditto, P. H., & Croyle, R. T. (1986). Judging health status: Effects of perceived prevalence and personal relevance. *Journal of Personality and Social Psychology*, *50*, 899–905.

Jentzsch, I., & Sommer, W. (2002). The effect of intentional expectancy on mental processing: A chronopsychophysiological investigation. *Acta Psychologica*, *111*, 265–282.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*, 20–31.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Johnston, W. A., & Hawley, K. J. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin and Review*, *1*, 56–72.

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349–364.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, *100*, 109–128.

Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, *93*, 429–445.

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, *9*, 131–155.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Kaiser, C. R., Major, B., & McCoy, S. K. (2004). Expectations about the future and the emotional consequences of perceiving prejudice. *Personality and Social Psychology Bulletin*, *30*, 173–184.

Kanazawa, S. (1992). Outcome or expectancy? Antecedent of spontaneous causal attribution. *Personality and Social Psychology Bulletin*, *18*, 659–668.

Karniol, R. (2003). Egocentrism versus protocentrism: The status of self in social prediction. *Psychological Review*, *110*, 564–580.

Kawada, C. L., Oettingen, G, Gollwitzer, P. M., & Bargh, J. A. (2004). The projection of implicit and explicit goals. *Journal of Personality and Social Psychology*, *86*, 545–559

Kay, A. C., Jimenez, M. C., & Jost, J. T. (2002). Sour grapes, sweet lemons, and the anticipatory rationalization of the status quo. *Personality and Social Psychology Bulletin*, *28*, 1300–1312.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–240). Lincoln: University of Nebraska Press.

Kelly, G. (1955). *A theory of personality: The psychology of personal constructs*. New York: Norton.

Kiesler, C. A. (1971). *The psychology of commitment*. New York: Academic Press.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychological Review*, *109*, 306–329.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110*, 499–519.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118.

Kruger, J., & Burrus, J. (2004). Egocentrism and focalism in unrealistic optimism (and pessimism). *Journal of Experimental Social Psychology*, *40*, 332–340.

Kruger, J., & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology*, *40*, 586–594.

Kruglanski, A. W. (1989). The psychology of being "right:" The problem of accuracy in social perception and cognition. *Psychological Bulletin*, *106*, 395–409.

Kruglanski, A. W. (1996). Motivated social cognition: Principles of the interface. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263–283.

Kulik, J. A. (1983). Confirmatory attribution and the perpetuation of social beliefs. *Journal of Personality and Social Psychology*, *44*, 1171–1181.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*, 565–579.

Kunda, Z., & Oleson, K. C. (1997). When exceptions prove the rule: How extremity of deviance determines deviants' impact on stereotypes. *Journal of Personality and Social Psychology*, *72*, 965–979.

Larsen, J. T., McGraw A. P., Mellers, B. A., & Cacioppo, J. T. (2004).

The agony of victory and thrill of defeat: Mixed emotional reactions to disappointing wins and relieving losses. *Psychological Science*, *15*, 325–330.

Lewin, K., Dembo, T., Festinger, L., & Sears, P. S. (1944). Level of aspiration. In J. M. Hunt (Ed.), *Personality and the behavior disorders* (pp. 333–378). New York: Roland Press.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology*, *34*, 199–249.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1968–1980. *Psychological Bulletin*, *90*, 125–152.

Loewenstein, G. F. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal*, *97*, 666–684.

Loewenstein, G., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review*, *100*, 91–108.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization. *Personality and Social Psychology Bulletin*, *13*, 411–429.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

MacCoun, R. J. (1998). Biases in the interpretation and use of research results. *Annual Review of Psychology*, *49*, 259–287.

Mandler, G. (1975). *Mind and emotions*. New York: Wiley.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*, 437–458.

Matt, J., Lethold, H., & Sommer, W. (1992). Differential effects of voluntary expectancies on reaction times and event-related potentials: Evidence for automatic and controlled expectancies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 810–822.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McDonald, H. E., & Hirt, E. R. (1997). Motivated use of expectancy information in reconstructive memory. *Journal of Personality and Social Psychology*, *72*, 5–23.

McFarland, C., & Alvaro, C. (2000). The impact of motivation on temporal comparisons: Coping with traumatic events by perceiving personal growth. *Journal of Personality and Social Psychology*, *79*, 327–343.

McGill, A. L. (1993). Selection of a causal background: Role of expectation versus feature mutability. *Journal of Personality and Social Psychology*, *64*, 701–707.

McMullen, M. N., & Markman, K. D. (2000). Downward counterfactuals and motivation: The wake-up call and the Pangloss effect. *Personality and Social Psychology Bulletin*, *26*, 575–584.

McNulty, J. K., & Karney, B. R. (2002). Expectancy confirmation in appraisals of marital interactions. *Personality and Social Psychology Bulletin*, *28*, 764–775.

McNulty, J. K., & Karney, B. R. (2004). Positive expectations in the early years of marriage: Should couples expect the best or brace for the worst? *Journal of Personality and Social Psychology*, *86*, 729–743.

Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic athletes. *Journal of Personality and Social Psychology*, *69*, 603–610.

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*, 423–429.

Metcalfe, J., & Mischel, W. (1999). A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.

Miceli, M., & Castelfranchi, C. (2002). The mind and the future: The (negative) power of expectations. *Theory and Psychology*, *12*, 335–366.

Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. *Annual Review of Psychology*, *37*, 233–256.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Mitchell, T. R. (1974). Expectancy models of job satisfaction, occupational preferences, and effort: A theoretical, methodological, and empirical appraisal. *Psychological Bulletin*, *81*, 1053–1077.

Mitchell, T., Thompson, L. Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluation of events: The "rosy view." *Journal of Experimental Social Psychology*, *33*, 421–448.

Neuberg, S. L. (1989). The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies. *Journal of Personality and Social Psychology*, *56*, 374–386.

Newby-Clark, I. R., & Ross, M. (2003). Conceiving the past and future. *Personality and Social Psychology Bulletin*, *29*, 807–818.

Newby-Clark, I. R., Ross, M., Buehler, R., Koehler, D. J., & Griffin, D. (2000). People focus on optimistic and disregard pessimistic scenarios while predicting their task completion times. *Journal of Experimental Psychology: Applied*, *6*, 171–182.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Norem, J. K., & Cantor, N. (1986). Defensive pessimism: "Harnessing" anxiety as motivation. *Journal of Personality and Social Psychology*, *52*, 1208–1217.

Norem, J. K., & Illingworth, K. S. S. (1993). Strategy dependent effects of reflecting on self and tasks: Some implications of optimism and defensive pessimism. *Journal of Personality and Social Personality*, *65*, 822–835.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53–79.

Oettingen, G., & Mayer, D. (2002). The motivating function of thinking about the future: Expectations versus fantasies. *Journal of Personality and Social Psychology*, *83*, 1198–1212.

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.

Pennington, G. L., & Roese, N. J. (2003). Regulatory focus and temporal perspective. *Journal of Experimental Social Psychology*, *39*, 563–576.

Peterson, C., & Bossio, L. M. (2001). Optimism and physical well-being. In E. C. Chang (Ed.), *Optimism and pessimism: Implications for theory, research, and practice* (pp. 127–145). Washington, DC: American Psychological Association.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 323–390). New York: McGraw-Hill.

Pham, L. B., & Taylor, S. E. (1999). From thought to action: Effects of process- versus outcome-based mental simulations on performance. *Personality and Social Psychology Bulletin*, *25*, 250–260.

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.). New York: International Universities Press.

Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and Attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology*, *80*, 876–893.

Radcliffe, N. M., & Klein, W. M. P. (2002). Dispositional, unrealistic, and comparative optimism: Differential relations with knowl-

edge and processing of risk information and beliefs about personal risk. *Personality and Social Psychology Bulletin*, *28*, 836–846.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*, 45–48.

Roediger, H. L. III. (1990). Implicit memory. *American Psychologist*, *45*, 1043–1056.

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*, 805–818.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*, 133–148.

Roese, N. J. (2001). The crossroads of affect and cognition: Counterfactuals as compensatory cognitions. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 307–316). Mahwah, NJ: Erlbaum.

Roese, N. J. (2004). Twisted pair: Counterfactual thinking and the hindsight bias. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 258–273). Oxford, UK: Blackwell.

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, *114*, 363–375.

Roese, N. J., & Olson, J. M. (1997). Counterfactual thinking: The intersection of affect and function. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 1–59). San Diego, CA: Academic Press.

Rosenthal, R., & Jacobson, LO. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston.

Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, *35*, 817–829.

Rothbart, M. (1981). Memory processes and social beliefs. In D. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 145–181). Hillsdale, NJ: Erlbaum.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, B. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology*, *14*, 237–255.

Rothbart, M., & Lewis, S. (1988). Inferring category attributes from exemplar attributes: Geometric shapes and social categories. *Journal of Personality and Social Psychology*, *55*, 861–872.

Rotter, J. B. (1954). *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice Hall.

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The investment model scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships*, *5*, 357–391.

Sagan, C. (1997). *Billions and billions: Thoughts on life and death at the brink of the millennium*. New York: Ballantine.

Sanna, L. J. (1999). Mental simulations, affect, and subjective confidence: Timing is everything. *Psychological Science*, *10*, 339–345.

Sanna, L. J. (2000). Mental simulation, affect, and personality: A conceptual framework. *Current Directions in Psychological Science*, *9*, 168–173.

Sanna, L. J., & Schwarz, N. (2004). Integrating temporal biases: The interplay of focal thoughts and accessibility experiences. *Psychological Science*, *15*, 474–481.

Sanna, L. J., & Turley, K. J. (1996). Antecedents to spontaneous counterfactual thinking: Effects of expectancy violation and outcome valence. *Personality and Social Psychology Bulletin*, *22*, 906–919.

Schank, R. C. (1982). *Dynamic memory*. Cambridge, UK: Cambridge University Press.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (2001). Optimism, pessimism, and psychological well-being. In E. C. Chang (Ed.), *Optimism and pessimism: Implications for theory, research, and practice* (pp. 189–216). Washington, DC: American Psychological Association.

Schkade, D. A., & Kilbourne, L. M. (1991). Expectation-outcome consistency and hindsight bias. *Organizational Behavior and Human Decision Processes*, *49*, 105–123.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527–561). New York: Guilford Press.

Sedikides, C., & Green, J. D. (2000). On the self-protective nature of inconsistency/negativity management: Using the person memory paradigm to examine self-referent memory. *Journal of Personality and Social Psychology*, *79*, 906–922.

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 209–269). San Diego, CA: Academic Press.

Segerstrom, S. C., Taylor, S. E., Kemeny, M. E., & Fahey, J. L. (1998). Optimism is associated with mood, coping, and immune change in response to stress. *Journal of Personality and Social Psychology*, *74*, 1646–1655.

Seligman, M. E. P. (1998). *Learned optimism* (2nd ed.). New York: Pocket Books.

Shepperd, J. A., Findley-Klein, C., Kwavnick, K. D., Walker, D., & Perez, S. (2000). Bracing for loss. *Journal of Personality and Social Psychology*, *78*, 620–634.

Shepperd, J. A., & McNulty, J. K. (2002). The affective consequences of expected and unexpected outcomes. *Psychological Science*, *13*, 85–88.

Sherman, J. W. (2001). The dynamic relationship between stereotype efficiency and mental representation. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 177–190). Mahwah, NJ: Erlbaum.

Sherman, J. W., & Bessenoff, G. R. (1999). Stereotypes as source monitoring cues: On the interaction between episodic and semantic memory. *Psychological Science*, *10*, 106–110.

Sherman, J. W., Conrey, F. R., & Groom, C. J. (2004). Encoding flexibility revisited: Evidence for enhanced encoding of stereotype-inconsistent information under cognitive load. *Social Cognition*, *22*, 214–232.

Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype-relevant information under cognitive load. *Personality and Social Psychology Bulletin*, *26*, 26–34.

Sherman, J. W., Klein, S. B., Laskey, A., & Wyer, N. A. (1998). Intergroup bias in group judgment processes: The role of behavioral memories. *Journal of Experimental Social Psychology*, *34*, 51–65.

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology*, *75*, 589–606.

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and stereotyping: Cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology*, *11*, 145–175.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*, 211–221.

Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology*, *17*, 142–158.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review*, *94*, 439–454.

Showers, C. (1992). The motivational and emotional consequences of considering positive or negative possibilities for an upcoming event. *Journal of Personality and Social Psychology*, *63*, 474–484.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, *68*, 247–260.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis confirmation strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, *22*, 93–121.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3–21.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*, 656–666.

Spaniol, J., & Bayen, U. T. (2002). When is schematic knowledge used in source monitoring? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 631–651.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605–632.

Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 316–345.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, *96*, 58–63.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*, 42–61.

Stern, L. D., Marrs, S., Millar, M. G., & Cole, E. (1984). Processing time and the recall of inconsistent and consistent behaviors of individuals and groups. *Journal of Personality and Social Psychology*, *47*, 253–262.

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. *Psychological Bulletin*, *130*, 324–340.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Suls, J. (1983). Cognitive processes in humor appreciation. In P. E. McGhee & J. H. Goldstein (Eds.), *Handbook of humor research* (Vol. 1, pp. 43–60). New York: Springer-Verlag.

Swann, W. B., Jr., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, *46*, 1287–1302

Swann, W. B., Jr., Stein-Seroussi, A., & Giesler, B. (1992). Why people self-verify. *Journal of Personality and Social Psychology*, *62*, 392–401.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457–482.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*, 67–85.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Taylor, S. E., & Gollwitzer, P. M. (1995). Effects of mindset on positive illusions. *Journal of Personality and Social Psychology*, *69*, 213–226.

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003). Are self-enhancing cognitions associated with healthy or unhealthy biological profiles? *Journal of Personality and Social Psychology*, *85*, 605–615.

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters*, *8*, 201–207.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Thorpe, S. J., Rolls, E. T., & Maddison, S. (1983). The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Experimental Brain Research*, *49*, 93–115.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 401–421.

Tuckey, M. R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied*, *9*, 101–118.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Clarendon Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., & Houle, S. (1994). Novelty encoding networks in the human brain: Positron emission tomography data. *NeuroReport*, *5*, 2525–2528.

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, *49*, 577–585.

van Dijk, W. W., Zeelenberg, M., & van der Pligt, J. (2003). Blessed are those who expect nothing: Lowering expectations as a way of avoiding disappointment. *Journal of Economic Psychology*, *24*, 505–516.

von Hippel, W., Jonides, J., Hilton, J. L., & Narayan, S. (1993). Inhibitory effect of schematic processing on perceptual encoding. *Journal of Personality and Social Psychology*, *64*, 921–935.

Vorauer, J. D., & Miller, D. T. (1997). Failure to recognize the effect of implicit social influence on the presentation of self. *Journal of Personality and Social Psychology*, *73*, 281–295.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition*, *93*, 167–198.

Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 548–573.

Weinstein, N. D. (1980). Unrealistic optimism about life events. *Journal of Personality and Social Psychology*, *39*, 806–820.

Weinstein, N. D., & Klein, W. M. (1996). Unrealistic optimism: Present and future. *Journal of Social and Clinical Psychology*, *15*, 1–8.

Whittlesea, B. W. A. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 37, pp. 211–264). San Diego, CA: Academic Press.

Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, *49*, 62–79.

Whittlesea, B. W. A., & Williams, L. D. (2001). The discrepancy–attribution hypothesis, Part II: Expectation, uncertainty, surprise and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 14–33.

Wilcox, T., Nadel, L., & Rosser, R. (1996). Location memory in healthy preterm and full-term infants. *Infant Behavior and Development*, *26*, 529–568.

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology*, *80*, 572–584.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 345–411). San Diego, CA: Academic Press.

Wilson, T. D., & LaFleur, S. J. (1995). Knowing what you'll do: Effects of analyzing reasons on self-prediction. *Journal of Personality and Social Psychology*, *68*, 21–35.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Wilson, T. D., Lisle, D. J., Kraft, D., & Wetzel, C. G. (1989). Preferences as expectation-driven inferences: Effects of affective expectations on affective experience. *Journal of Personality and Social Psychology*, *56*, 519–530.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Wirtz, D., Kruger, J., Scallon, C. N., & Diener, E. (2003). The role of predicted, online, and remembered experience on future choice. *Psychological Science*, *14*, 520–524.

Wong, P. T. P., & Weiner, B. (1981). When people ask "why" questions and the heuristics of attributional search. *Journal of Personality and Social Psychology*, *40*, 650–663.

Wyer, R. S., & Collins, J. E. (1992). A theory of humor elicitation. *Psychological Review*, *99*, 663–688.

Zeelenberg, M. (1999). The use of crying over spilled milk: A note on the rationality and functionality of regret. *Philosophical Psychology*, *12*, 325–340.

Zeelenberg, M., Beattie, J., van der Pligt, J., & de Vries, N. K. (1996). Consequences of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes*, *65*, 148–158.

The Principles of Social Judgment

Arie W. Kruglanski
David Sleeth-Keppler

The topic of social judgment has received considerable attention from researchers over the last several decades. To our knowledge, the first volume with the term "social judgment" in its title was the classic work by Muzafer Sherif and Carl Hovland (1961). Since then, a number of books (e.g., Eiser, 1990; Forgas, 1991; Forgas, Williams, & Von Hippel, 2003; Martin & Tesser, 1992; Nisbett & Ross, 1980; Wyer, 2004) and a much larger number of research articles (e.g., Beauregard & Dunning, 1998; Bodenhausen, Kramer, & Süsser, 1994; Dunning & Cohen, 1992; Fiedler, Armbruster, Nickel, Walther, & Asbeck, 1996; Gilovich, Medvec, & Savitsky, 2000; Innes-Ker & Niedenthal, 2002; Jackson, Sullivan, & Hodge, 1993; Lambert, 1995; Moskowitz & Roman, 1992; Mussweiler, 2003; Vallacher, Nowak, & Kaufman, 1994; Wegener & Petty, 1995) have referred to *social judgment* in their titles.

For the most part, these contributions have addressed a plethora of diverse phenomena and did not attempt to delineate an integrative theory of social judgment, or to articulate the general principles that govern the formation and/or change of social judgments (for exceptions, see Eiser, 1990; Wyer, 2004). Instead, each relevant article or volume addressed a specific category of judgmental contents (e.g., attitudes, attributions, stereotypes, numerical estimates, social comparison, and impression formation) and/or dwelt on a specific domain of variables affecting that phenomenon. For example, the volume by Forgas (1991) focused on the impact of *affective states* on social judgments, whereas Nisbett and Ross

(1980) examined mechanisms that may lead to *biased* judgments, and so on. The various contributions to the social judgment literature thus contributed empirical fragments to the complete "puzzle" without piecing them together for an overall perspective on the social judgment process.

Our aim in this chapter is twofold: (1) to provide a general overview of the topics (phenomena and mechanisms) that fit under the general umbrella of "social judgment," and (2) to provide an overall conceptual framework for organizing these topics and affording a set of principles governing the social judgment process. To that end, we explore commonalities between judgmental phenomena that have not been heretofore treated in similar terms, such as attitudes, stereotypes, attributions, heuristics, and biases, and cultural differences in judgment.

Our approach follows from our prior analyses of the epistemic process (see Chun & Kruglanski, 2006; Erb, Kruglanski, et al., 2003; Kruglanski, 1989, 2004; Kruglanski, Erb, Pierro, Mannetti, & Chun, in press; Kruglanski et al., 2005; Kruglanski & Thompson, 1999a, 1999b). In the pages that follow we first outline that process and the principles it suggests. We then apply these notions to heterogeneous judgmental phenomena and compare our framework to alternative conceptualizations, in particular to the various dual-process and dual-systems models of social judgment. Space considerations preclude a truly exhaustive review of the voluminous literature on social judgments. Ideally, however, we cover a

sufficient amount of diverse theoretical notions and research programs to illustrate the integrative potential of our approach.

THE PROCESS OF JUDGMENT FORMATION: THE "HOW" OF SOCIAL JUDGMENT

We are assuming that at a deep level all judgments, regardless of topic, are formed in the same way. Specifically

1. Judgments are derived from evidence in accordance with specific inference rules.
2. Such inferences may be consciously represented or carried out without conscious awareness.
3. They may be general or context specific.
4. They may be held commonly by group of people, or be unique to given individuals.
5. They may be rendered with ease and alacrity or via a tantalizingly laborious and time-consuming process.

Crucial to our overarching analysis of human judgment are several continuous parameters whose intersections may characterize different judgmental *contexts* and determine the impact of information on judgmental outcomes. Those parameters are described subsequently, after we have laid forth a few general assumptions.

The Concept of Evidence

As already noted, we are assuming that judgments are based on information the knower regards as pertinent "evidence." For instance, in forming an attitudinal judgment one might consider information about the positive and negative consequences mediated by an attitude object (Ajzen & Fishbein, 1980); these serve as evidence for the attitude object's goodness (or badness). Similarly, in forming a causal attribution, one may consider information about the covariation of an effect with an entity, and covariation (along with temporal precedence and contiguity) may serve as evidence for causality. In forming a personal impression of someone, one might treat as evidence that individual's category membership (e.g., her or his gender, profession, race, age, or religion), and the characteristics stereotypically attached to that category. In forming a forecast about one's affective reaction to a future event, one's evidence might be one's present feeling state plus one's subjective estimate of how long such feelings may endure (Gilbert, Gill, & Wilson, 2002; Wilson, Centerbar, Kermer, & Gilbert, 2005).

Some authors (in particular Lieberman, Gaunt, Gilbert, & Trope, 2002; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004) have argued recently that in addition to the rule-following process implicit in the notion of "evidence," individuals may also reach judgments via a more "automatic, associative" process. We discuss such processes at a later juncture and examine whether in fact they qualitatively differ from rule following, hence whether they constitute a valid alternative to the general inferential process we are postulating.

The Logical Structure of Reasoning from Evidence to a Conclusion

To function as evidence, information has to fit an inference rule of the "if X then Y" type (Erb, Kruglanski, et al., 2003; Kruglanski & Thompson, 1999a, 1999b; Pierro, Mannetti, Kruglanski, & Sleeth-Keppler, 2004). For example, information that an actor succeeded at a difficult cognitive task could serve as evidence that he or she is intelligent, given the prior assumption that "if someone succeeds at difficult cognitive tasks, he or she is intelligent." More formally speaking, the reasoning from evidence to conclusions is syllogistic. It involves a major premise, the "if X then Y" conditional rule mentioned earlier, and a minor premise that instantiates the antecedent term of the major premise (X) for a given entity (event, etc.) P. Specifically, such minor premise asserts that P is X, and hence that Y is to be expected. For instance, a person may subscribe to the stereotypic belief "if college professor then absent-minded." Upon encountering a person known to be a college professor (minor premise instantiating the antecedent term of the major premise), the knower would be subjectively justified in inferring that he or she must, therefore, be absent-minded.

Issues of Relativity or Invariance

In some cases, the major premises an individual subscribes to may appear to hold across contexts; in other cases they may be context-specific. For example, the statement "if mammal then nursed" applies to all mammals. On the other hand, the statement "if weighing over 7 lbs. then heavy" may be assumed to hold true in some contexts only, and not in others. For instance, a steak weighing 7 lbs. is heavy, but a bicycle weighing 7 lbs. is light. Similarly, a prison term of 7 years imposed by a given judge for a given crime may be judged as lenient in context A of other judges imposing considerably *longer* terms, and as harsh in context B of other judges imposing considerably *briefer* terms. In other words, one could adopt the premise "if 7 then lenient" in context A and "if 7 then harsh" in context B (Higgins & Stangor, 1988). As we shall see, major topics in the psychology of judgment (e.g., issues of assimilation and contrast, dispositional inferences, judgmental anchoring, and estimation tasks) are fundamentally related to the contextual dependence of pertinent inference rules to which individuals may subscribe.

Inferential and Enabling Factors in Judgment Formation

Inferential Factors

Persons may differ in the inference rules (major premises) to which they subscribe. Consequently, the very same information may lead them to different, and occasionally opposite, conclusions. For instance, one person may hold a stereotype whereby "all college professors are absent-minded," whereas another person may believe that "all college professors are well organized." These two individuals are, therefore, likely to form quite differ-

ent impressions (i.e., judgments) of a target person upon learning that he or she is a professor.

The derivation of conclusions from subjectively relevant evidence may appear to represent a highly "rational," deliberative, and conscious process. That is probably why authors have often juxtaposed rule following to automatic processes represented by the associationistic "route" to judgments (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004). Appearances can be misleading, however. Thus, some inferences may be routinized to the point of automaticity, and hence occur without much conscious awareness. For instance, Uleman's work on spontaneous trait inferences (Uleman, 1987, 1989; Uleman, Winborne, Winter, & Shechter, 1986) suggests that lawful (i.e., rule following) inferences presumably may occur without explicit inferential intentions, and without conscious awareness of performing an inference: The spontaneous trait inference that John is "clumsy" on basis of the information that "he stepped on Stephanie's foot while dancing" surely requires the inference rule "if stepping on a dancing partner's foot, then clumsy," or some variant thereof. A person who did not subscribe to that premise would not probably reach that conclusion. Unconscious inferences from mood states are treated in these terms by Schwarz and Clore's (1996) "mood as information" model.

As already implied, the major premise from which an individual draws her inferences need not be widely shared, or universally proclaimed as rational. For instance a premise whereby a rain dance by a shaman will engender rain (i.e., "if rain dance, then rain") might appear as "irrational" to sophisticated Westerners yet as quite valid to members of a tribe. Thus, the general mechanism of drawing conclusions from evidence is assumed to be the same irrespective of the apparent veridicality or "rationality" of one's premises to other persons (Cole & Scribner, 1974).

The foregoing analysis suggests that individuals' lay theories may determine what they regard as compelling evidence for a given judgment. Individuals who subscribe to an "entity" theory of intelligence (Dweck, 1999) may regard the observation that an individual failed an intelligence test as evidence for her or his low intelligence because of their assumptions that intelligence is an invariant personal trait tapped by intellectual tasks. By contrast, individuals subscribing to an "incremental" theory (Dweck, 1999), that is, ones who believe that intelligence is malleable, may view failure as irrelevant to an inference of intelligence.

Enabling Factors

TASK DEMANDS

Beyond inferential factors, such as the apparent relevance of evidence, judgment formation is affected by several enabling factors that do not typically enter into the judging individual's calculations. A major such factor is the cognitive demandingness of the information-processing task. This may be determined by the length and complexity of the information given, accessibility of major premises retrieved from memory (of the "if X then Y" type) and accessibility or saliency of information constituting the minor premises (X).

Accessibility of events from memory may affect judgments, without the judges' explicit awareness of it being so, and without such awareness entering into the judges' explicit inferential process. For instance, participants in a study by Lichtenstein, Slovic, Fischoff, Layman, and Combs (1978) estimated the relative frequencies of death from various causes without realizing that their personal experiences involving each class of lethal event (determining the events' accessibility) was correlated with their estimates. Accessibility of such personal experiences may have made their utilization *easy* in deriving an overall estimate of death causes. As shown subsequently, ease of applying an inference rule should play a particularly important role in judgments if one's processing resources were limited due to various motivational and cognitive factors.

MOTIVATION

Different motivational categories may determine (1) the extent of effort invested in information processing en route to judgment, and (2) the weights assigned to different informational items as a function of their desirability. The former effect has been classified in the category of nondirectional motivation and the latter in that of directional motivation (see Kruglanski, 1996, for a review).

Nondirectional exertion of effort is important in relation to task demands. In cases in which task demands are high (e.g., where the information was lengthy, complex or relatively inaccessible), considerable effort might be required for it to be properly utilized. For instance, with insufficient motivation and a demanding task, highly relevant information (e.g., high-quality arguments) might go underappreciated, whereas information of relatively low relevance (e.g., low-quality arguments) might be overappreciated (Petty, Wells, & Brock, 1976).

Differential weight given to informational items as function of their correspondence with a given "directional" motivation may result in bias such that the judgmental conclusions would be congruent with the individual's wishes and desires. For instance, an individual informed that coffee drinking is healthy or unhealthy may distort his or her memories of coffee-drinking instances (Kunda & Sanitioso, 1989) to reach a desirable conclusion that his or her coffee drinking habit is health preserving (see also Dunning, 1999; Kruglanski, 1999; Kunda & Sinclair, 1999).

The impact of a directional motivation may also depend on task demands. Where these are considerable, individuals' processing resources might be consumed by labors of decoding the information given, with less resources being left for suppression and/or alternative construals consistent with the directional motivation. Indirect evidence that processing "ease" facilitates motivational distortion comes from a number of sources. Festinger (1957), for example, argued that the *attitude change* occasioned by cognitive dissonance would take

place because it is an *easier* avenue of restoring consonance than is altering one's cognition about the dissonant behavior. Similarly, Wood, Kallgren, and Preisler (1985) found that participants with greater access to attitude-relevant information in memory, that is, participants for whom it should have been *easier* to generate motivationally congruent construals, generated more negative thoughts and were less persuaded by counterattitudinal, hence motivationally undesirable, messages than were less knowledgeable participants. Finally, Kunda (1990) as well as Hsee (1996) argued that motivational biases are more likely when implications of the information given are ambiguous versus clear-cut presumably because clear-cut implications are more *difficult* to distort. According to the present perspective, however, with sufficient motivational resources even information with relatively clear implications (i.e., clear to outside observers) might be distorted as illustrated—for example, by the occasional denials by widows and widowers of the death of their spouses (Stroebe, Hansson, Stroebe, & Schut, 2001), or the denials of terminally ill patients of their own impending death (Kübler-Ross, 1969).

COGNITIVE CAPACITY

The individual's cognitive capacity may play a similar role to nondirectional motivation in determining the extent of information processing en route to judgment. Thus, an individual whose cognitive capacity is taxed (e.g., by cognitive busyness or load) may be unable to thoroughly process and hence fully appreciate the relevance of the information given under high task demand conditions. Again, this may reduce the difference in impact of highly judgmentally relevant versus less judgmentally relevant information. In this vein, work by Petty and colleagues (1976) demonstrated that under distraction (vs. no distraction) conditions, individuals differentiated less between high- and low-quality arguments.

SEQUENCE AND CARRYOVER EFFECTS

Early information may bias the processing of subsequent information especially if the subsequent information is ambiguous. For instance, if the early information implied a given position, subsequent arguments may be interpreted in its light. Such early information may exist in the form of prior expectancies or attitudes to which the knower subscribes (Olson, Roese, & Zanna, 1996). In an early study demonstrating this effect, Lord, Ross, and Lepper (1979) found that participants differed in their perceived validity of research studies depending on the results' consistency with these persons' attitudes. Participants opposed to capital punishment regarded the methodology of studies supportive of capital punishment as inferior to the methodology of studies opposed to it. By contrast, participants in favor of capital punishment viewed studies supportive of capital punishment as methodologically superior.

In the realm of persuasion, biasing effects of heuristic cues on the processing of ambiguous message arguments were reported by Chaiken and Maheswaran (1994). Spe-cifically, under high-motivation conditions the same arguments attributed to a credible source were deemed of a better quality than when they were attributed to a noncredible source. Erb, Pierro, Mannetti, Spiegel, and Kruglanski (2003) showed further that early message arguments can bias the evaluation of subsequent message arguments, and also the subsequent perception of the communicator's expertise, determining their impact on the ultimate judgments.

This concludes our introductory analysis of judgmentally relevant factors. We now apply this framework to several classical issues in social judgment, beginning with the often-invoked notions of *assimilation* and *contrast*.

CLASSIC ISSUES IN SOCIAL JUDGMENT

Assimilation and Contrast

The concepts of *assimilation* and *contrast* were featured prominently in the early volume on social judgment by Sherif and Hovland (1961). These specific terms and the research paradigm from which they derived were heavily influenced by research in psychophysics. From this perspective, assimilation is said to occur when a stimulus is judged as closer to a standard (on some dimension, e.g., weight) than objectively warranted. Contrast, on the other hand, occurs when a stimulus is judged as more distant from a standard than objectively warranted. One may think about it in terms of the *range* of possible values introduced by the anchor. A heavy anchor, way above the stimulus series, creates a much larger range than a relatively light anchor close to the stimulus series. The wider the range the more its end points appear as different from one another. Such differentiation may be exaggerated resulting in contrast. For instance if the high end of the range is 200 grams then 5 grams appears *extremely different* from that value and the notion of "extreme difference" may lead to its *underestimation* leading to a judgment of say 3.5 grams. Assimilation on the other hand may lead to an exaggeration of similarity. Thus, if the difference between a stimulus and a standard was very slight (say, the standard was 5 grams and the stimulus was 4.5 grams) the difference might not be even discerned; the two stimuli may be perceived as "extremely similar," and the stimulus may be misjudged as even closer to the standard than actually is the case.

The same logic may apply to social judgments. A given stimulus, say, uttering a curse, would be seen as innocuous when the considered range of values included *genocide* and as much more hostile/aggressive when the range was much more restricted (e.g., with *shoving* as its high end). Occasionally opposite ends of a range can be activated or brought to mind by a primed social stimulus. Thus, Herr (1986) found opposite contrast effects through priming different person concepts: Priming the concept of Shirley Temple resulted in labeling a moderately hostile behavior as much more aggressive than priming the concept of Adolf Hitler. Presumably the Temple prime activated the low end of the aggressive range with regard to which the target stimulus (moder-

ately hostile behavior) was *contrasted* as aggressive. Similarly, the Hitler prime might have activated a high end of the aggressive range to which the target behavior was *contrasted* as nonaggressive.

It is instructive to consider the assimilation and contrast phenomena in terms of what a given stimulus, say, a given weight or a given behavior, *means* in a given context. The context, in other words, sets up an *inference rule* linking a given stimulus value with a given judgment. Consider contrast. In the context of highly aggressive behaviors a moderately aggressive behavior may be classified as mild in accordance with the inference rule "if mass murder is what *aggressive* means, then shoving is *mild.*" Similarly in the context of relatively unaggressive behaviors a moderately aggressive behavior may be classified as *very* aggressive according to a rule such as "if sulking is what *aggressive* means then shoving is *very aggressive.*" As noted earlier, these inferences could be mediated by the judgment that the stimulus and a comparison standard are *very different*, with the attached labels or numerical values (as in the case of weights) reflecting this notion of *difference*. Thus, if Gulliver's height is judged in the context of Lilliputs—his dissimilarity from them establishes him as a giant. If, however, it is judged in the context of Giants, his dissimilarity establishes him as a dwarf.

Whereas contrast is mediated by the perception of pronounced *dissimilarity* between a stimulus and a standard, assimilation is mediated by the perception of their *similarity* on some judgmental dimension. Such perceived similarity affords the inference of equivalence between the stimulus and the standard. If the category "adventurous," for instance, includes risky behavior, and if such category is activated, then any behavior with a risky component (say playing the stock market) may be considered an instance of the "adventurous" category (Higgins, Rholes, & Jones, 1977). The inference rule in such a case might be "if risky means adventurous, and if playing the stock market is risky then playing the stock market is adventurous." Thus, similarity between a stimulus and a category on a defining dimension may mediate assimilation in accordance with the general inference rule "if a stimulus is similar to a category on its defining property, then stimulus if a member of the category." This notion also underlies the classical representativeness heuristic discussed by Tversky and Kahneman (1974).

If assimilation and contrast effects are driven by exaggerations of perceived similarity and dissimilarity, it is possible that people whose cognitive style requires clear-cut distinctions (for instance, persons high on the need for closure dimension—Kruglanski, 2004; Kruglanski & Webster, 1996) would exhibit more pronounced assimilation and contrast effects. That is, they would judge relatively close stimuli as closer to each other, glossing over their possible differences, than people low on the need for closure and would judge relatively distant stimuli as more distant from each other, glossing over their possible similarities, than would the lows. These notions are akin to Sherif and Hovland's (1961) hypothesis that the strength of assimilation and contrast effects, expressed in polarized perception of others' attitudes, would be posi-

tively related to the judges' ego involvement in the issue. If one could assume that ego involvement determines commitment to one's position, and if commitment induces a desire for clarity and an intolerance of ambiguity, this could result in a polarization of one's assessment of other people's positions. Various studies have confirmed this proposition in the context of racial attitudes in the United States (see Eiser, 1990, for a review).

Needs for Specific Closures

Our analysis suggests that setting up an inference rule determining assimilation and contrast effects should be influenced by two processes: (1) identification of the standard and (2) assessment whether the judged stimulus is similar on a relevant dimension to (assimilation) or different from the standard (contrast). Among other factors, this may depend also on individuals' directional motivations, or needs for specific closure (Kruglanski, 1989, 2004). For example, a committed group member may be motivated to view the group norm as a standard, an aspiring worker may be motivated to view the boss' attitudes as a standard, and mediocre students may be motivated to view a failing grade as a standard (from which they may wish to differentiate themselves), and so on.

Given that a certain standard had been identified, individuals may be motivated to perceive themselves as close (i.e., similar) or as distant (i.e., different) from it. If the standard was positively valued (e.g., if it represented success, health, or happiness), one might be motivated to perceive oneself as similar to it (resulting in assimilation). If, however, the standard was negatively valued (representing failure, illness, or unhappiness), one might well wish to distance oneself from the standard (resulting in contrast). For instance, in Brewer's (1991) optimal distinctiveness model, individuals might wish to identify with others under some circumstances; such motivation could foster an exaggerated perception of similarity between themselves and the group standard. In other circumstances, however, they might wish to assert their uniqueness. This may lead to exaggerated perceptions of a difference, and hence to a contrast, between themselves and other people.

Inclusion and Exclusion Processes

The comparison standard may be placed within the category (or, within the stimulus series) or outside it. This may affect how the category as a whole is evaluated. Inclusion increases the perception of similarity between the category and the standard whereas exclusion increases the perception of dissimilarity. Bless and Schwarz (1998) had German participants consider Richard Von Weizsäcker (a highly respected German politician) either as president of the Republic (which excludes him from his party, Christian Democratic Union [CDU]) or as a member of the party (signifying inclusion). When he was included, an assimilation effect occurred and the CDU was evaluated more positively than in a control condition, whereas when he was excluded (i.e., placed outside the group) the CDU was evaluated more negatively than

in the control condition. A different party, the Social Democratic Party of Germany, was not affected by these manipulations.

Latitudes of Acceptance and Rejection

An intriguing concept introduced by Sherif and Hovland (1961) concerned the distinction between latitudes of *acceptance* and *rejection*. The latitude of acceptance denotes the range of attitudinal positions with which an individual *agrees*. These are typically clustered around the individual's own attitudinal position. The latitude of rejection is a range of positions with which the individual *disagrees*. These are typically clustered at some distance from the individual's own position. Sherif and Hovland (1961) hypothesized that the size of the latitudes of acceptance and rejection depends on the degree of one's ego involvement in the attitude at issue. When ego involvement is high, the latitude of acceptance narrows. In other words, the individual would agree with positions closely resembling her or his own, and disagree with other positions (see Hovland & Sherif, 1952).

As noted earlier, ego involvement may represent a motivationally based commitment to, or "freezing" on, one's position (Kruglanski, 2004; Kruglanski & Webster, 1996). This may induce a reluctance to change one's position in direction of other discrepant positions. In other words, the wider latitude of rejection evinced by ego-involved respondents might reflect a motivationally based refusal to move toward alternative positions.

But note also that ego involvement may reduce the latitude of noncommitment by interpreting items in this category as *supportive* of one's own position. This leads to the prediction that the relative size of the latitudes of acceptance and rejection might depend on the inclusion of ambiguous items (those defining the noncommitment category). The more such items are included, the greater should be the number of items in the latitude of acceptance category, hence the larger that particular category relative to the latitude of rejection. Similarly, the fewer such items the larger should be the latitude of rejection relative to that of acceptance.

In short, ego involvement may instigate both the freezing on one's own specific position, increasing the latitude of rejection, and the tendency to absorb ambiguous items (originally in the latitude of noncommitment) in one's latitude of acceptance, increasing the acceptance latitude through item distortion. Future research is needed to identify the moderator variables that might determine the relative strength of the "freezing" and distortion tendencies and, hence, the conditions under which ego involvement might lead to the narrowing or the widening of the latitudes of acceptance or rejection.

In summary, various findings in the assimilation and contrast literature may be interpreted in terms of the present conceptual analysis. Assimilation and contrast may depend on inference rules regarding the meaning attached to a given stimulus (e.g., a "shove") in a given context. This, in turn, should depend on this stimulus's similarity or dissimilarity to other stimuli whose mean-ings are known to the perceiver. Assimilation and contrast may also depend on participants' motivations, determining their tendency to stick to ("freeze upon") their positions, and determining their latitudes of acceptance and rejection.

Persuasion

Background Knowledge and Informational Impact

Background knowledge affects the degree to which an argument, delivered by some source, will be persuasive to individuals. For instance, a person who assumed (as part of her background knowledge) that military service is essential for an effective discharge of presidential duties will consider the information that a candidate served (or did not serve) in the military as relevant to whether he or she would be a good president. By contrast, a person who did not assume a relation between military service and presidential efficacy would not be affected much by information concerning a candidate's military service and would consider it irrelevant to the candidate's fitness for the job.

To be influential in persuasion settings, background knowledge need not be restricted to the *content* of presented arguments but may include assumptions related to the *context* in which they are presented to recipients. These may have to do, for instance, with expertise, or epistemic authority, of the communicator (Kruglanski et al., 2005), the degree of consensus about a given position, and so on (Chaiken, 1980). For instance, one's background knowledge may include the assumptions that "experts are correct," or "majorities can be trusted," lending persuasive power to information that a given communicator is an expert, or that a majority advocates a given position.

Background knowledge may have other effects as well. Thus, people with a great deal of such knowledge are likely to have well-formed and relatively fixed judgments and attitudes on the issues at hand. Such attitudes should allow them to make decisions quickly and effortlessly. For example, research by Fazio, Blascovich, and Driscoll (1992) as well as Blascovich and colleagues (1993) has shown that persons who rehearsed their attitudes toward sets of abstract paintings (vs. just focusing on the color of the paintings) showed less autonomic reactivity when requested to evaluate the paintings.

It appears then that possessing well-developed attitudes as part one's store of knowledge frees up cognitive resources for other tasks. On the other hand, having well-developed attitudes has a potential downside in effecting fixation and rigidity (Kruglanski, 2004). For example, individuals with highly (vs. less) accessible attitudes about an object appear to show greater difficulties in perceiving changes that it may undergo (Fazio, Ledbetter, & Towles-Schwen, 2000).

Relatedly, background knowledge may facilitate counterarguing against persuasive communications assaulting one's position, which may assist individuals in maintaining their original views. A recent meta-analytic review of the literature on forewarning found statistical support

for the notion (Wood & Quinn, 2003) that persons cognitively bolster their existing attitudes by recruiting available knowledge in support of their position.

Task Demands, Cognitive Capacity, and Processing Motivation

Information presented to recipients in a persuasive context may be easier or more difficult to process depending on the situation. For instance, the information may be lengthy or complex or fraught with irrelevant detail. All these may make it difficult to process. By contrast, if presented succinctly and pointedly information may be relatively easy to process. Furthermore, the inference rules that lend the information its relevance may be relatively inaccessible in memory, hence difficult to retrieve. Alternatively, they may be quite accessible and easy to retrieve (Higgins, 1996). If the information is difficult to process, discernment of its judgmental relevance may require considerable processing resources, consisting of cognitive capacity and processing motivation. Thus, if the judgmental task demands are high (the information is difficult to process, and/or pertinent inference rules are difficult to retrieve), persuasive impact of the information given will manifest itself to a greater extent under ample (vs. limited) processing motivation and cognitive capacity. Similarly, where cognitive capacity and motivation are limited, easy-to-process information would have a relatively greater persuasive impact than difficult-to-process information.

The aforementioned considerations are pertinent to our recent reinterpretation of evidence cited in support of the dual-process models of persuasion, the elaboration likelihood model (ELM) (see Petty & Cacioppo, 1986) and the heuristic systematic model (HSM) (Chaiken, Liberman, & Eagly, 1989; see, e.g., Erb, Kruglanski, et al., 2003; Kruglanski et al., 2006; Kruglanski & Thompson, 1999a, 1999b; Kruglanski, Thompson, & Spiegel, 1999). Basically, both dual-process models propose two qualitatively distinct ways of being persuaded. One is via the message arguments, the other by aspects of the context in which these arguments are embedded (e.g., characteristics of the communicator, one's mood, degree of consensus about the advocated position). Changing one's attitudes on basis of the message arguments has been referred to as persuasion by the central route (in the ELM) or via the systematic mode (in the HSM). Changing one's attitudes on basis of the contextual elements has been referred to as persuasion via the peripheral route (in the ELM) or via the heuristic mode (in the HSM). A great deal of empirical support for the dual-mode models of persuasion rested on evidence that contextual information is persuasive under different conditions than information contained in the message arguments: Contextual information was typically persuasive under conditions of low cognitive capacity and/or motivation, whereas message or issue information was typically persuasive under high capacity and motivation conditions.

Whereas these results appear consistent with the dual-mode view of persuasion, they are subjects to a general

alternative interpretation based on a possible confounding of task demands with informational type (Kruglanski & Thompson, 1999a, 1999b). In a typical (dual-mode) persuasion experiment, the peripheral or heuristic cues are presented up front, and the message arguments are presented subsequently. Moreover, the message arguments are typically lengthier and more complex than the (heuristic or peripheral) cues (see Kruglanski & Thompson, 1999b, for a review).

The foregoing features could have rendered the message arguments more difficult to process than the heuristic cues. That could be the reason why the cues typically exerted their persuasive effect under low processing motivation or cognitive capacity, whereas the message arguments typically did so under high motivation and capacity. Indeed, research by Kruglanski and his associates found in several studies that controlling for the difficulty of processing eliminates the previously found interactions between the type of information (peripheral/heuristic cues vs. message arguments) and processing resources as far as attitude change is concerned (Erb, Pierro, et al., 2003; Kruglanski & Thompson, 1999a, 1999b; Pierro et al., 2004; Pierro, Mannetti, Erb, Spiegel, & Kruglanski, 2005).

Subjective Relevance of Information: The "Override" Effect

A puzzling phenomenon in persuasion research has been that "peripheral" or "heuristic" cues, have appreciable impact on recipients' attitudes and opinions under limited (motivational or cognitive) processing resources (conditions of low "elaboration likelihood" in the ELM terminology), yet they typically fail to exert persuasive impact under ample processing resources, where the (subsequently presented) message arguments carry the persuasive brunt. One of the explanations for this effect was Petty's (1994) *cue-weighting* hypothesis whereby "when motivation and ability are high . . . the cues are in essence discounted as irrelevant at the time of attitude judgment" (p. 234). Indeed, in an analysis of the relevant research Pierro and colleagues (2004) discovered that (positive) cues are typically perceived by participants as less relevant to the attitudinal judgments than the (strong) message arguments used in that research. Based on this finding, Pierro and colleagues restated Petty's cue-weighting hypothesis in more general terms as a *relevance override* hypothesis whereby in a confrontation between a more subjectively relevant and a less subjectively relevant piece of information, *any* more relevant versus less relevant information (not necessarily message arguments vs. "cues"), the less relevant information would lose some of the persuasive impact it might otherwise exert.

Pierro and colleagues (2004) argued further that if the less relevant information came first and was followed by the more relevant information—the less relevant information should have greater impact under limited processing resources, and the more relevant information should have greater impact under ample processing resources, just as found in the preponderance of persua-

sion studies (see Kruglanski & Thompson, 1999a, 1999b, for reviews). If, however, the order was reversed and the more relevant information came first, it should have the greater impact irrespective of processing resources. Pierro and colleagues obtained strong support for the foregoing *relevance override* hypothesis across three different studies that differed in the type of information presented early ("cues" or "message arguments") and late ("cues" or "message arguments") in the sequence.

Taken as a body, the research to date suggests that the content or type of information (e.g., represented in the distinction between cues and message arguments) does not matter as far as persuasive impact is concerned. What matters are the judgmental parameter values (e.g., on subjective relevance, task demands, or presentation sequence parameters). When these are controlled for, prior pervasive *differences* between the peripheral or heuristic processing of "cues" and the central or systematic processing of issue- or message-related information disappear, questioning the need to regard the two modes of persuasion identified in the ELM or the HSM as *qualitatively distinct*.

Causal Attributions

Background Knowledge about the Meaning of Causality

Causal attribution requires that individuals use their background knowledge as to the nature of causality, namely, that it involves *temporal precedence* of the cause relative to the effect (see Kassin & Pryor, 1985, for a review), as well as *covariation* between the two (Hilton & Slugoski, 1986; Kelley, 1967, 1971; Kruglanski, 1989). Detection of temporal precedence of *X* with respect to *Y* and covariation between the two is evidence that *X* is cause of *Y*. So, for example, in Kelley's (1967) well-known scheme (the so-called Kelley's cube) an *entity attribution* is made if the effect covaried with a given entity but not with other entities (the distinctiveness criterion), and if it did not covary with persons (the consensus criterion), time, or modality (the consistency criterion), attesting that neither the specific person nor the time or the modality was a cause of the effect. Similarly, a *person attribution* would be made if the effect covaried with the person but not with time, modality, and entity (e.g., the individual was polite to all other individuals, at all times and in all situations).

Background knowledge, or the lay theories to which one happens to subscribe, may also determine what specific causal hypotheses are considered plausible in the first place. For instance, a person experiencing the symptoms of a common cold may generate the alternative hypotheses that this was caused by exposure to an inclement weather, an encounter with others suffering the same ailment, or insufficiently warm clothing, because this individual may have subscribed to lay theories causally linking these factors with the effect at hand. Individuals with different lay theories may generate different causal hypotheses to explain the same effect. In the realm of achievement motivation, for instance, individuals with an entity theory (Dweck, 1999) may account for failure on a task in terms of a lack of ability, whereas individuals with an incremental theory (Dweck, 1999) may account for it in terms of insufficient mastery.

Behavior Identifications versus Dispositional Attributions

The importance of background knowledge (about the concept of causality in general, and/or about specific "candidate causes") is not at all unique to causal attributions and is shared by all instances of judgment. A question whether this is actually so was raised in a set of intriguing papers by Yaacov Trope and his colleagues (Trope, 1986; Trope & Alfieri, 1997; Trope & Liberman, 1996). Trope (1986) proposed an influential dual-process model wherein the *contextual-constraint* information (i.e., whether powerful situational pressures existed impelling most people to enact the behavior under the circumstances) impacts judgments of behavior *identification* and *dispositional* attributions in qualitatively different ways.

Trope's (1986) analysis assumes a sequence of phases wherein behavior identification precedes dispositional attribution. It assumes further that at the behavior identification stage, the incorporation of contextual constraints is effortless, automatic, and independent of cognitive resources. By contrast, at the dispositional-inference stage the influence of context was assumed to be controlled, deliberative, and capacity demanding (Trope & Alfieri, 1997, p. 663).

In support of these notions, Trope and Alfieri (1997, experiment 1) found that assimilating a behavior to the context appears to constitute a resource-independent, effortless task that anyone can carry out, irrespective of cognitive capacity, whereas dispositional inference is effortful and resource-dependent and people can perform it only if endowed with adequate cognitive capacity.

But from a more general perspective, both behavior identification and dispositional attribution constitute judgments that differ in contents but presumably are governed by the same process. Behavior identification tasks turn on the question "*What is it?*" and How can it be classified? Dispositional attributions revolve about the different question of "*What caused it?*" According to our analysis, contextual information, to the extent that it is subjectively relevant, may constitute evidence relevant both to the behavioral identification and to the dispositional attribution judgments.

We assume as well that the more difficult the judgmental task, the more resources should be required to carry it out. In those terms, it is quite possible that Trope and Alfieri's (1997) findings pertain to a situation wherein, for some reason, using the contextual information to answer the *identification* question was relatively easy, and, hence, was unaffected by cognitive load, whereas using that same information to answer the dispositional-*inference* question was relatively difficult and hence sensitive to load.

That the dispositional inference question can be independent of load was demonstrated in a series of studies by Trope and Gaunt (2000). As Trope and Gaunt summarized it, "cognitive load eliminated discounting when situational [or contextual] information was low in salience,

low in accessibility, or low in specificity. However, when situational information was made more salient, accessible, or specific, it produced strong discounting effects even under cognitive load . . . " (p. 344). These results seem at odds with a dual-process model that portrays dispositional inferences (in counterdistinction to behavioral identifications) as *inherently* exigent of resources and, hence, sensitive to load.

Chun, Spiegel, and Kruglanski (2002) conducted three separate experiments investigating the parallel question, namely, that incorporating the context into *behavior identifications* would require resources if this was made sufficiently difficult. Consistent with this analysis, participants' perception of the ambiguous behavior was independent of cognitive load where the behavior identification task was made easy (by increasing the saliency of the behavioral or the contextual information), but it was significantly dependent on load where this task was difficult. From this perspective, it does not seem necessary to posit *qualitatively distinct* judgmental processes for the phases of behavior identification and dispositional inference. The extant evidence suggests that when the parameter of processing difficulty is controlled for the putative processing, differences between these phases disappear.

Correction Processes

Whereas nondirectional processing motivation (e.g., an involvement in an issue) may afford the discovery of an otherwise difficult to *discern* inference rule (e.g., embedded in a lengthy and complex message argument), directional motivation may affect individuals' *readiness* to use that rule. This issue has been addressed by research on correction processes in social cognition (Martin, 1986; Martin, Seta, & Crelia, 1990; Petty & Wegener, 1993; Wegener & Petty, 1995). Thus, Martin's (1986) set–reset model posits that assimilation of judgment to a prior context (such as a prime) is the default tendency for most people. For instance, if after being primed with the word "reckless" one finds out that "Donald shot the rapids in a kayak" one might apply the rule "If pursuing dangerous activities (such as rapid-shooting) then reckless" (representing assimilation). However, if the priming was particularly blatant and or if alternative information became available suggesting that a given bit of information was inappropriate and should not be used, people may suppress such information. The motivation to suppress the "illegitimate" information may fail in its quest if this required greater processing resources than the individual possessed or was ready to commit. In this vein, a study by Martin and colleagues (1990) showed that only people who were high but not those who were low in the need for cognition *contrasted* their impressions of a target person (along a dimension ranging from "stubborn" to "persistent") away from an implication of a prior prime.

Is assimilation invariably the default option? Petty and Wegener (1993) argued, alternatively, that this may depend on individuals' *metacognitive* inference rules about how their judgments are affected by prior information. Under some conditions, contrast rather than assimilation may be the default. Aware of this fact, individuals

motivated to correct for what they perceive as an illegitimate "contrast bias" may strive to eliminate it. In this vein, Petty and Wegener showed that under correction instructions, participants rated several target cities equally, regardless of whether they previously thought of a desirable vacation spot (i.e., Hawaii) or a more neutral spot (i.e., St. Louis). When no correction instructions were given, however, participants in the desirable location condition rated the targets as less desirable, compared to participants in the neutral city condition. These results suggest that assimilation can be the corrective strategy in situations in which the default bias seems to be to contrast an average target away from a desirable target. In other words, the rule "if Paris exemplifies a desirable vacation spot then Lawrence, Kansas is rather undesirable" was corrected for by participants who were specifically motivated to do so.

Wegener and Petty (1995) proposed, more generally, that people may hold either assimilative or contrasting inference rules (hence, that neither is necessarily the *default*). If motivated to correct they might do so by "bending over backwards" and executing contrasting or assimilative procedures depending on the perceived initial source of bias. In support of this hypothesis, Wegener and Petty found that if people held an assimilative theory they tended to use a contrasting correction strategy, following instructions to correct for the bias. However, people who initially subscribed to a contrast theory reacted with an assimilative correction strategy when instructed to avoid bias.

In summary, people may be aware of some of the inference rules that mediate their judgments. If led to believe that a given application of those rules is inappropriate and biased, and if motivated to correct for such a bias, they may either suppress the rules in question (Martin, 1986) or apply compensatory metacognitive rules (e.g., "If I overestimated the job satisfaction of someone in the Bahamas [because of contrasting or assimilative tendencies], I need to reduce my estimate.")

Heuristics and Biases

Base Rate Neglect and the "Representativeness" Heuristic

For several decades now, instigated by the seminal contributions of Amos Tversky and Daniel Kahneman (1974, Kahneman, 2003; Kahneman & Tversky, 1973), investigators of judgment and decision-making behavior have been exploring people's tendency to render judgments consistent with normative statistical models, versus relying on a variety of suboptimal shortcuts referred to as heuristics. A major appeal of this work was that it challenged the presumption of human rationality and questioned the rational-choice paradigm, popular in the various social sciences (Cohen, 1979, 1981; Hastie & Dawes, 2001; Scott, 2000). Whereas the normative statistical models (the Bayes theorem in particular) were seen to offer optimal algorithms for likelihood estimates from quantitative information, research has shown that people's actual behavior often strays from the dictates of

such models and seems governed by simplistic rules of thumb rather than more normatively appropriate statistical information.

Yet, if our analysis is correct, the claims that persons generally neglect normative statistical principles in favor of suboptimal heuristics have to be qualified by taking into account (1) people's background knowledge of statistical concepts, and their accessibility as well as (2) their background knowledge and accessibility of given heuristic rules. One way of doing so is to impart to individuals a rule, which some subsequent information might fit. In this manner, individuals who have been taught the rule may have it available in their mental repertory, while other individuals deprived of such training may not. Consistent with this notion, Ginossar and Trope (1980) found that people vary in their tendency to endorse the sampling rule linking base rates to likelihood estimates, and that the endorsers utilize the base rate to a greater extent than do the nonendorsers.

Other research (Nisbett, Fong, Lehman, & Cheng, 1987; Sedlmeier, 1999) has established that statistical reasoning can be taught and that it can result in the increased use of statistical information. Training effects were also invoked by Gigerenzer (1996; Gigerenzer & Hoffrage, 1995, 1999).

Furthermore, the perceived relevance of the individuating information can also be readily manipulated, for instance, by making it more or less fitting to a given stereotype. Ginossar and Trope (1987, Experiment 1), for example, found that when the individuating information was made relevant to the lawyer versus the engineer categories, it was used that way and the base rates were essentially ignored. However, when it was irrelevant it had no impact on judgments and the base rates were markedly utilized.

Beyond the availability of particular (individuating or statistical) rules that lend relevance to the information given, their utilization may depend on their accessibility. For instance, Sedlmeier and Gigerenzer (2001) showed drastic improvements in participants' ability to apply Bayesian inference rules, simply by reframing Bayesian conditional problems in terms of frequencies, a statistical notion more *accessible* to the average person than ratios (Sedlmeier & Gigerenzer, 2001). Erb, Kruglanski, and colleagues (2003) similarly indicated that making statistical constructs accessible through priming increases the use of base rates and reduces the reliance on stereotypical, "representativeness" information.

Processing Difficulty and Processing Resources

One might still wonder why relevant individuating information typically *overrides* the base rate information, even though the latter too is undeniably perceived as relevant and is generally used in the absence of relevant individuating information (Ginossar & Trope, 1987). In a recent paper, Chun and Kruglanski (2006) explored the possibility that this may be so in part because of a confounding in typical base rate neglect experiments between *type of information* (e.g., individuating vs. statistical) and its *processing difficulty*, recalling a similar confounding in per-

suasion studies (Erb, Kruglanski, et al., 2003; Kruglanski & Thompson, 1999a, 1999b; Pierro et al., 2005). In a typical "lawyer/engineer" study described in the literature (e.g., Kahneman & Tversky, 1973), the base rate information was presented via a single sentence appearing up front. The individuating (or representativeness) information usually followed and was conveyed via a relatively lengthy vignette. As a consequence, the base rate information may have been relatively easy to process, compared to the individuating information. If one assumes that participants in the typical base rate neglect studies had sufficiently high degrees of processing motivation and cognitive capacity, it is possible that they focused their processing efforts on the more complex individuating information and hence were able to appreciate its relevance to the judgment at hand. Furthermore, the focus on the later and more complex information may have reduced individuals' attention to the earlier information, resulting in its "neglect," just as in persuasion studies the early and easy-to-process "cue" information is typically neglected (under high motivation and capacity conditions) (e.g., Chaiken et al., 1989; Kruglanski & Thompson, 1999a, 1999b; Petty & Cacioppo, 1986).

But if processing difficulty (i.e., task demandingness) matters, we should be able to increase or decrease the use of statistical *or* individuating information, by appropriately varying its processing difficulty, and the participants' resources. Chun and Kruglanski (2006) found strong support for these notions across several experiments. For example, in their first experiment, these authors replicated the typical lawyer/engineer paradigm (Kahneman & Tversky, 1973) in one condition by presenting brief and upfront base rate information followed by lengthier individuating information. In a contrasting condition, these relations were *inversed* by presenting brief individuating information first followed by lengthier and more complex base rate information. As expected, the former condition replicated the typical finding of base rate neglect whereas the latter condition evinced significant base rate utilization. In a subsequent study (Chun & Kruglanski, 2006, study 2), imposition of cognitive load actually ironically increased the ("normative") use of base rate information when it was presented briefly and upfront; hence it was easy to deal with.

The considerable research accorded base rate neglect and other failures to use statistical information attest that these phenomena are nonunique and that they are governed by the same judgmental parameters that determine the utilization of any other information.

Other Judgmental Heuristics

AVAILABILITY

The availability heuristic pertains to judgments of frequency or likelihood of objects and events, in reference to reliance on the ease with which instances of such object or event categories come to mind (Kahneman & Tversky, 1973, 1982). Whereas *superficially* the availability heuristic may appear quite different from the representativeness heuristic—the two share a *deep* similarity in

that both refer to inferential rules, albeit of different contents. The representativeness heuristic refers to a rule whereby "if *X* (an object or an event) exhibits the defining properties of category *Y* (or is "similar" to the category) then it is member of that category." The availability heuristic refers to the metacognitive rule, "if an exemplar of a category comes to mind easily, there must be many instances of it" (Ruder & Bless, 2003, p. 21).

In recent years, the ease-of-retrieval interpretation of the availability heuristic was juxtaposed to the amount of instances (or content) interpretation. Thus, Schwarz and colleagues (1991) found that fewer versus more numerous instances of recalled instances of some trait (e.g., assertiveness) led to higher ratings on the corresponding trait (see also Schwarz, 1998; Wänke, Bless, & Biller, 1996). Nonetheless, the *amount* of instances (or content of the information) has also been found to mediate likelihood or frequency inferences. Thus, Rothman and Schwarz (1998) reported that individuals with low issue involvement relied on ease of retrieval whereas highly involved individuals relied on the content that came to mind (see also Aarts & Dijksterhuis, 1999). Assuming that the ease-of-retrieval rule is easier to apply than the content rule (that at the very least requires the operation of counting instances), these latter findings seem to constitute additional instances of the principle (illustrated in our earlier discussion of base rate utilization) that for it to be applied in judgment, the difficulty or demandingness of a rule must be matched by the individuals' processing resources.

Obviously, the ease-of-retrieval rule would not be used if the instantiating information about ease (the minor premise in the syllogism) was lacking. Thus, in research by Wänke and colleagues (1996) when yoked participants received a list of arguments generated by others, they were more influenced by the large (vs. the small) number of arguments, probably because the experience of ease was unavailable to them. Also, when the "ease rule" was rendered contextually inapplicable via a misattribution of difficulty to background music participants tended to use the "amount of instances" rule and to associate many (vs. few) instances with higher ratings on the focal trait (of assertiveness) (Schwarz et al., 1991).

Research by Fishbach, Igou, and Kruglanski (2005) suggested that participants are capable of using either the ease or the amount or the content rules and that they would use them when their corresponding instantiating information was rendered salient. These authors' research demonstrates that the ease-of-retrieval rule need not be easier to apply than the amount of instances or the content rules, suggesting that there is no universal connection between the tendency to apply the ease rule (or, for that matter, any other rule) and processing difficulty. For example, participants seem to be quite selective in their application of the ease-of-retrieval rule, and they realize that their own ease of retrieval is irrelevant to assessing another person's properties. Furthermore, when the ease-of-retrieval rule is made more accessible, participants rate themselves as more assertive in a fewer (vs.

more) recalled instances condition. However, when the amount-of-instances rule is made more accessible, participants rate themselves as more assertive in the more (vs. fewer) instances condition (Fishbach et al., 2005).

The foregoing research suggests that the availability heuristic pertains to inference rules (of either the ease or the amount variety) whose functioning and application are governed by similar principles (related to relation between the difficulty of accessing an inference rule and processing resources, perceived rule relevance to given judgments, etc.) as those underlying the use of other inference rules of different contents.

Anchoring Effects

One of the most interesting heuristics identified by Tversky and Kahneman (1974) is that of *anchoring* and *adjustment*. Empirically, this phenomenon is typically illustrated in a paradigm that includes two sequential questions. The first question requires participants to estimate whether a judgmental target, or target feature, is higher or lower than some numerical comparison standard, or anchor. This is referred to as the *comparative* question. The second question requires participants to provide an *absolute* estimate pertaining to the same judgmental target. In one of the most well-known studies employing the classic anchoring paradigm, Tversky and Kahneman's (1974) participants were first presented with a *comparative* question, which asked them to indicate whether the percentage of African nations in the United nations was higher or lower than 65% or 10%. Afterwards, participants were asked to provide an *absolute* estimate of the percentage of African nations in the United Nations. Despite the fact that the anchors were generated randomly (by spinning a wheel of fortune), results revealed that median estimates of the absolute percentage, 45% and 25%, respectively, clearly were impacted by those anchor values.

Recently, Strack and Mussweiler (1997) proposed that the anchoring effect may be explained in terms of a selective knowledge activation mechanism (SA), which assumes that judges, comparing a given target to an anchor value, may positively test the hypothesis that the target in question is identical to the anchor value, generating different contents of knowledge consistent with the magnitude implications of a given anchor (Chapman & Johnson, 1999; Klayman & Ha, 1987; Strack & Mussweiler, 1997). Anchoring effects essentially obtain, according to Strack and Mussweiler's (1997) model, because people tend to assimilate their final judgments to previously activated, anchor-biased target knowledge.

One of the implications of this model involves the degree to which the target knowledge, activated in response to the positive hypothesis-testing process, is deemed *relevant* to a subsequent absolute estimation task (see also Higgins, 1996; Martin & Achee, 1992; Wegener & Petty, 1995, 1997). The selective activation (SA) model posits that activated target knowledge has to be *applicable* to a subsequent absolute judgment and *representative* of the target of the absolute judgment (Strack & Mussweiler,

1997). Consistent with the notion that activation of specific target knowledge mediates anchoring, in one study (Strack & Mussweiler, 1997, experiment 1) an anchoring effect was only obtained in a standard anchoring condition in which the judgmental dimension in question (height of the Brandenburg Gate in Berlin) was held constant between the comparative and absolute judgments, whereas no anchoring effect was obtained in a condition involving a dimension change between the comparative and absolute judgments (from height to width of the gate), supporting the knowledge applicability constraint of the SA model (Strack & Mussweiler, 1997). Thus, the *subjective relevance* of the information generated during the comparative judgment task appears to affect the degree to which the anchoring effect obtains under conditions wherein the target of the absolute judgment differs from the target of the comparative judgment.

The hypothesis that people generate different contents of knowledge, or evidence, during the comparative phase, as a function of the anchor value considered during the comparison phase, is perhaps best supported by studies that show a reaction time advantage to anchor-consistent words in a lexical decision task (Mussweiler & Strack, 2000, study 1).

Despite the potential of the selective accessibility model to explain a wide range of anchoring effects, involving different targets and differentially extreme anchor values, there exists a class of anchoring phenomena, known as *basic anchoring effects* (Wilson, Houston, Etling, & Brekke, 1996), that are not easily explained by the semantic activation model. For example, anchoring effects appear to obtain under conditions in which the target of the initial comparison is completely unrelated to the target of the absolute judgment. In one study Wilson and colleagues (1996) asked one group of participants whether the number of nations in the United Nations was more than, less than, or equal to 1930 and another group whether the number of physicians in the local phonebook was more than, less than, or equal to 1930. In a subsequent absolute judgment task, all participants were asked to estimate the number of nations in the United Nations, showing anchoring effects with comparable effect sizes in both groups. As Wilson and colleagues have argued, anchoring effects can be explained by invoking a simple short-term memory representation of the numerical anchor itself. Because numerical values may be represented in short-term memory independent of the context in which they were initially encountered, anchors may influence a large variety of subsequent absolute judgments without the relevance constraints imposed by the selective accessibility model.

To unify the perspectives of semantic activation and basic numeric anchoring, Sleeth-Keppler (2004) recently proposed a simplified semantic activation model that has the potential to account for anchoring effects obtained under the classic paradigm as well as under conditions of changed judgmental targets. Unlike Strack and Mussweiler (1997), who include in their activation hypothesis primarily *target* semantics (e.g., selectively activated knowledge *about a given target*), Sleeth-Keppler has argued that the semantics generated during the comparative judgment task may consist of simple, semantic representations of the *numerical anchor value* itself (i.e., nontarget-specific information about quantity such as "large" or "small," "tall" or "short"). For example, in one study, testing this model, Sleeth-Keppler varied the target of the comparative judgment task in a between-subjects design, while holding the numerical anchor value constant. Thus, half of the participants were asked to indicate whether pop singer Britney Spears was older or younger than 45 years, whereas the other half was asked to indicate whether Hollywood actor Clint Eastwood was older or younger than 45 years. In the Britney Spears condition, participants were expected to be primed with the concept "old," because 45 years constituted a high standard for her age. However, in the Clint Eastwood condition, participants were expected to be primed with the concept "young," because 45 years represented a low standard for his age. After the comparative task, all participants were asked to rate the age of a third, unrelated target person (Senator Bennett, R-Utah). The results revealed higher estimates of the target's age in the Britney Spears compared to the Clint Eastwood condition. Furthermore, results from a lexical-decision task, which was conducted as part of the same study, showed a significant reaction-time advantage to target words related to the concept "old" in the Spears condition (compared to words related to the dimension "young"). This pattern of results was reversed in the Eastwood condition, where participants were faster at identifying "young" compared to "old" target words.

Note, that both Strack and Mussweiler's (1997) SA model and Sleeth-Keppler's (2004) general semantic model assume that manipulation of the anchor serves to generate information *relevant* to the judgmental task engaged in during the absolute phase of the procedure. Such information may activate matching aspects in the target, which then are used as evidence for the absolute judgment. For example, activation of the concept "old" by the comparative question may focus the individual's attention on corresponding features of the target (e.g., aspects of a person that signify old age). In turn, those activated features may serve as *evidence* that the target is in fact old. Alternatively, accessibility of the semantic term activated by the anchor may itself be taken as evidence that the target is characterizable by this term. Thus, if "old" comes to mind when thinking of a target person, this might suggest to one (i.e., be taken as "evidence," or intuitive hunch) that he or she in fact is old. In short, recent interpretations of the anchoring phenomenon suggests that it is mediated by inferences from evidence based on the appropriate "if–then" rules, just as are other judgments.

Finally, there is also evidence that when participants are motivated to engage in extensive informational search, and the relevant information is available in their environment, a high degree of accuracy motivation may lead these individuals to search for alternative evidence beyond that activated by the anchor, thus mitigating the anchoring effects (Kruglanski & Freund, 1983).

Additional Judgmental Domains

Social Comparison Processes

The present scope does not allow a detailed review of all additional domains of social judgment, but there are reasons to believe that the presently identified judgmental parameters play the same role in those other areas as well. Consider the well-tilled domain of social comparison processes. Recently Mussweiler (2003) provided an integrative review of this realm of phenomena, viewing social comparison as a special case of *hypothesis testing*. Obviously, such testing represents a "rule-based' process departing from the assumption that "if data D are observed, then hypothesis H is valid." What is special about social comparison is the comparative content of the hypotheses being tested. Thus, one may test the hypothesis that John has greater ability than Paul, or that Paul's age is less than John's. In cases discussed by Mussweiler persons are assumed to generate from memory data relevant to testing the various hypotheses of interest, though in other cases of social comparison data obtained from external sources, or by observation, may also be used. Individuals' cognitive capacity and (nondirectional and directional) motivation may play the same role in social comparisons as they play in other judgmental domains in determining the extent and thoroughness of information processing en route to judgment, as well as the amount of bias toward desirable (and away from undesirable) comparative conclusions (see Kruglanski & Mayseless, 1990, for a review).

Impression Formation

Impression formation has constituted a central domain of dual-process theorizing. Two influential models in this area have been those of Fiske and Neuberg (1990; Fiske, Lin, & Neuberg, 1999) and of Brewer (1988). Though these conceptualizations differ in important regards, they share several central features. Most important, they both distinguish between category-based and attribute-based processing and view them as *qualitatively different*. Category-based processing is assumed to be "top down" and attribute-based processing is assumed to be "bottom up" (Brewer, 1988, p. 4; Fiske & Neuberg, 1990, p. 60).

Both models also assume that impression formation follows a fixed order commencing with an automatic identification of the target in terms of some general categories. According to Fiske and Neuberg (1990), "the category label is more likely to be a social grouping (demographic category, role, job) than a single personality trait" (p. 10). Subsequently, if the incoming information and the knower's self-involvement warrant it, he or she may continue processing information and address individuating or personal-attribute information as well. In Brewer's (1988) model, for example, personalization requires a sufficient degree of self-involvement, which allows "attributes and behaviors . . . inconsistent with previously established expectancies (to) be processed extensively and incorporated into the person representation" (p. 23).

Fiske and Neuberg (1990) as well as Brewer (1988) are undeniably correct in pointing out that social category information may often appear highly relevant to various requisite judgments (e.g., the information that a target is a woman my seem relevant to the judgment of this person's submissiveness or nurturance by someone subscribing to a particular gender stereotype). But it is also true that such information might appear irrelevant to other judgments, to which individuating information may appear more relevant (e.g., the information that the target is a woman may appear irrelevant to a judgment of her wealth, whereas the information that her name happens to be Rockefeller may appear highly relevant to such judgment).The critical role of subjective relevance in this connection was noted by Medin (1988), who observed that "the response should depend on what information is *needed* and how well that information can be predicted from knowledge about who the person is, what groups they can be classified into, or what situation is instantiated" (p. 122, emphasis added). Earlier, the same point was implied by Bruner (1957) namely that "a primary determinant of category activation is the search requirements imposed by the perceiver's needs, objectives and task goals . . . " (cited in Brewer, 1988, p. 18).

Critics also questioned whether the terms "social category" and "personal attribute" may be meaningfully distinguished from each other as far as the judgmental process is concerned. In the statement "Judy is a nurse," the term "nurse" (i.e., a social category) can be readily considered one of Judy's attributes, as may the category "friendly person" in the statement "Judy is friendly." As Klatzky and Andersen (1988) noted, "Even individualized person concepts have associated attributes and . . . these attributes are themselves social categories" (p. 98). Medin (1988) expanded on this point to state that "if every person were treated as absolutely unique, than there would be no basis at all for generating expectations. This would be analogous to the situation of a physician being confronted with a totally new disease unlike any other . . . " (p. 122).

To summarize then, the information used on an encounter with a social stimulus may indeed represent a "social" *category* or a "trait/attribute" *category*, all depending on its subjective relevance ("search requirements," "task goals," or "informational needs") to the judgment at hand. As far as the judgmental process is concerned, "social categories," "traits," or "attributes" seem functionally equivalent in constituting *evidence* for requisite judgments. A detailed examination of dual-process models of impression formation in terms of the present judgmental parameters is given by Erb, Kruglanski, and colleagues (2003).

Dual-System Models of Social Judgment: Associative versus Rule-Following Accounts

Various dualistic social judgment models examined earlier revolved about *content differences* in inputs (to) or outputs from, the judgmental process. For instance, the ELM distinguished between the processing of *message*

arguments versus *peripheral cues* (Petty & Cacioppo, 1986) as inputs into attitude formation. Similarly, the HSM contrasted *message arguments* and *heuristic cues* as inputs. The work on biases and heuristics (Kahneman, 2003; Tversky & Kahneman, 1973) distinguished between the use of *statistical* (e.g., base rate) and *nonstatistical* (e.g., representativeness) inputs into likelihood estimates. Fiske and Neuberg (1990) as well as Brewer (1988) distinguished between inputs based on *categorical* and *individuating* information, and Trope's dual-process model of attributional inferences (Trope & Gaunt, 2000) distinguished between the judgmental *outcomes* of behavioral *identifications* versus dispositional *attributions*.

Sloman's Case for Two Systems of Reasoning

A different group of dualistic models distinguished between qualitatively distinct *systems* of reaching judgments. An early such model, described by Sloman (1996), distinguished between *associative* and *rule-based* systems of judgment. Sloman's primary attempt was to identify substantial criteria for building a case for the systemic-distinctiveness argument. For example, Sloman asked whether the speed with which a judgment is formed (e.g., in the context of engaging in simple vs. complex arithmetic problems) provides a demarcation for two systems of reasoning. But the *quickness* with which an answer comes to mind need not be compelling evidence for a qualitative difference in judgmental "systems." Instead, it could reflect the *difficulty* of the arithmetic task in the two instances. This may depend in part on the task itself, and in part on the degree to which it was routinized for a given individual (Schneider & Shiffrin, 1977). Thus, for a young child who is just learning to count, figuring out that 10 − 9 = 1 may be quite arduous and deliberative, whereas for a cashier who routinely sells a certain product for $27.35, figuring out the change for a $50 bill might be extremely rapid.

A second demarcation criterion, *awareness of the judgmental process*, was quickly discarded by Sloman (1996) himself. In his words, "Awareness provides only a fallible heuristic for identifying systems not a necessary or sufficient condition" (p. 6). For instance, "some reasoning is not obviously associative and yet apparently occurs without conscious awareness" (Nisbett & Wilson, 1977).

The S Criterion

Sloman (1996) finally settles on the one demarcation criterion that he viewed as satisfactory in warranting a qualitative distinction in process. This is his Criterion S (for simultaneity) described as follows: "A reasoning problem satisfies *Criterion S* if it causes people to simultaneously believe two contradictory responses" (p. 11). Because of the key importance that Sloman attaches to this particular criterion, it is well to submit it to a careful scrutiny. Take Sloman's own example, the statement that a "whale is a mammal" (p. 11). Whales are commonly perceived to resemble fish more than typical mammals like a cow or a horse. Thus, a knower may need to deal in this case with two contradictory

beliefs, one derived from the whale's outward similarity to fish and one derived from the "academic" knowledge that classifies whales as mammals.

Yet, it is quite possible that we may have here are two distinct *rules* yielding opposite conclusions. One rule is based on similarity, or the "representativeness" heuristic (and heuristics, after all, constitute rules *by definition*) (e.g., "if *X* looks like a fish, swims like a fish, and lives like a fish, *X* is a fish"). The other rule may be based on other criteria for classification in the mammal category (e.g., "breast feeding of offspring"), or, indeed the source *heuristic* "If a biology text claims *X* (e.g., that whales are mammals) then *X* is the case."

Perhaps the most striking of Sloman's (1996) examples concerns the Müller–Lyer illusion. Here, perception provides the answer that the lines are of unequal length, and a ruler furnishes an incompatible answer, that they are equally long. Once again, however, it is easy to understand this phenomenon in terms of two rules in which the individual happens to strongly believe, and that happen to yield disparate conclusions. One of these rules is that one's visual perceptions are valid ("if my eyes inform me that *X* then *X* it is"); the other, that application of a ruler yields valid answers.

Note that not every single person necessarily upholds both rules. For instance, individuals hampered by limited eyesight may probably harbor considerable doubt about the veracity of their perceptual experiences. Similarly, members of primitive cultures absent length measuring instruments may not trust the ruler much. Suppose, finally, that on two different occasions an individual measured the same line with different rulers, one of which was biased (e.g., having an inch represented by 3cm, rather than the normal 2.3cm). In one instance, the conclusion might be that the line's length is less than some *X*, and in the other instance, that it is more than *X*. Clearly then the use of the very same rule-based "reasoning system" (here application of a ruler) may lead to two contradictory conclusions, satisfying Sloman's Criterion S for systems distinctiveness.

In summary then, Sloman's (1996) criteria for a duality of the judgmental process seem readily explicable in terms of a single process determined by several quantitative parameters. Rapidity, lack of awareness, or automaticity more generally (cf. Bargh, 1996) may reflect degrees of routinization (hence processing difficulty) rather than a qualitative duality (as Sloman himself recognizes). And the Criterion S (of incompatible, strongly held beliefs) is compatible, in fact, with the notion that different rules (major premises) applied to the same evidence (minor premises) may yield completely different conclusions. None of this seems to warrant the assumption of a qualitative difference in the reasoning processes. That is not to say that associationistic processes do not exist but, rather, to question their qualitative distinction from rule-based processes. In fact, associations can be thought of as conditional rules of the "if *X* then *Y*" variety that may come to mind very rapidly and effortlessly because of their strength, that is, the degree to which the individual believes that *X* attests to *Y*, and their accessibility (Higgins, 1996).

Evidence from Conditioning Research

Conditioning phenomena (both classical and instrumental) are often viewed as the prototypes of associative learning. Yet, compelling evidence exists (Holyoak, Koh, & Nisbett, 1989; Rescorla, 1985; Rescorla & Holland, 1982; Rescorla & Wagner, 1972) that they are fundamentally rule based, in fact. Associative learning is assumed to constitute learning by *contiguity*, and *repeated pairing* of a conditioned stimulus (CS) and an unconditioned stimulus (US). Yet, evidence from the animal learning literature suggests that neither is necessary or sufficient for conditioning. That temporal contiguity is not necessary is attested by the fact that conditioning can occur when an interval of minutes or even hours elapses between the stimuli. Thus, if a rat ingested a novel substance and was made ill minutes or hours later, it will form a strong aversion to that substance (e.g., Garcia, McGowan, Ervin, & Koelling, 1968).

That pairing is not sufficient for conditioning has been further demonstrated (1) by Rescorla's (1968) results that if the probability of the reinforcer is the same in the absence of the CS as in its presence, no appreciable conditioning will take place, and (2) by conditioned inhibition effects whereby a higher-order stimulus (CS1) paired on some trials with another CS stimulus that on other trials was paired with some US (say, a shock) comes actually to inhibit the reaction associated with the US (e.g., crouching) even though a strict associationistic interpretation (of second-order conditioning) would suggest that this stimulus (i.e., the CS1) should evoke that very reaction. It thus appears that an animal, rather than responding mechanistically to contiguous pairings of stimuli over repeated occasions (representing the associationistic account) is attempting to learn environmental contingencies in which the occurrence of one event (e.g., shock) is conditional upon the occurrence of another (e.g., noise).

In the same vein, and based on their extensive review of pertinent conditioning studies, Holyoak and colleagues (1989) conclude that "representations of the environment take the form of . . . rules that compose mental models . . . the rat's knowledge about the relation between tones and shocks might be informally represented by a rule such as 'if a tone sounds in the chamber then a shock will occur, so stop other activities and crouch' " (p. 320).

The notion that "automatic" phenomena in the domain of (motor, or cognitive) skill acquisition involve a routinization of *if–then* sequences has been central to Anderson's (1983) adaptive control of thought model. Smith (1984, 1989; Smith & Branscombe, 1988; Smith, Branscombe, & Bormann, 1988) has extended this conception to the realm of social judgment. Specifically, Smith's research has demonstrated that social judgments represent a special case of procedural learning based on practice that strengthens the *if–then* components resulting in increased efficiency (or "automaticity). In summary, there are reasons to believe that the distinction between associative and rule-based processes, the former being "automatic" and the latter controlled or delibera-tive, confuses surface manifestations and deep structure, or as Bargh (1996) put it, constitutes a "false dichotomy" (p. 1706).

Reflexion versus Reflection

In a recent paper Lieberman and colleagues (2002) stated that "the idea that automatic processes are merely faster and quieter versions of controlled processes is theoretically parsimonious, intuitively compelling, and wrong" (p. 205). This conclusion is founded on the notion that different brain structures seem to be activated in automatic versus controlled behavior. More specifically, they propose that (what they refer to as) the X system, including the lateral temporal cortex, amygdala, and basal ganglia, is involved in automatic processing. By contrast, the C system, related to activity in anterior cingulate, prefrontal cortex, and hippocampus, seems activated when deliberative, or controlled processing takes place.

But the fact that different brain structures have been involved in instances of automatic versus controlled processing need not constitute compelling evidence that automatic and controlled processes are not "faster and quieter versions of controlled processes." The continuum from controlled to automatic processes is uncontroversial (Bargh, 1996; Schneider & Shiffrin, 1977), and the rule-like ("if–then") nature of "associationistic" as well as controlled inferences (Holyoak et al., 1989) has been generally recognized in the domain of classical conditioning. It is possible, for instance, that the different "brain systems" (i.e., X and C) simply respond to difficulty of processing such that beyond some threshold of difficulty, system X processing might not suffice and other brain structures (e.g., system C) need to kick in. This would be analogous to additional muscles, or stronger muscles, getting involved with increases in the weight that one tried to lift. Thus, involvement of different brain structures in the processing of more versus less practiced (and hence, efficient) *if–then* rules (Bargh, 1996) might merely signify the fact that the brain is responsive to *resource requirements* of information processing.

Lieberman and colleagues (2002) advanced an additional claim for their two anatomically distinct brain system. Namely, system X is assumed to process *identity information* (e.g., identification of a given behavior as member of a given category) whereas the C system is assumed to process causality information, and hence to be related to the behavior's causal origins. In other words, the X system is assumed to be involved both in the processing of identify information *and* in automatic processing. Similarly, the C system is assumed to be involved both in the processing of causality information and in controlled processing. All would be well if all identity processing was automatic, and all causality processing was controlled. There are reasons to believe, however, that this is not so. In work referred to earlier, Trope and Gaunt (2000), by making the causal attribution task easy, showed that it is insensitive to cognitive load, which would put it in the category of "automatic" processing (in terms of its efficiency).

On the other hand, Schneider and Shiffrin's (1977) classic task was one of identifying letters and digits in an array (representing an identity task par excellence), yet it took participants months to "automatize" it so that the speed of identification was not affected by the row length any more. Furthermore, in a study mentioned earlier, Chun and colleagues (2006) rendered the identification task "controlled" and hence resource dependent by decreasing the saliency of information pertinent to the "identity" inference. Thus, if system X is involved in automatic processing and not in controlled processing it could not be generally be involved in identity processing because, as we have seen, such processing can be "automatic" in some circumstances (in the efficiency sense) and "controlled" in other circumstances. Similarly, if system C is involved in controlled processing it could not be similarly involved in causal processing, which again could be "automatic" or "controlled." This reasoning leads to the conclusion that the dichotomy between the X and C systems does not offer strong evidence for a dual-process model of attributional inferences.

The Two Memory Systems Model

Smith and DeCoster (2000) proposed the existence of two qualitatively different memory systems: the slow learning and the fast learning systems. The slow learning system is assumed to be associative, and to learn general regularities slowly. The fast learning system is assumed to be rule based and to form representations of unique or novel events quickly. Presumably, the existence of this duality reflects the need to meet two conflicting demands: "One demand is to record information slowly and incrementally so that the total configuration in memory reflects a large sample of experiences. This is important so that general expectancies and long-term stable knowledge can be based on the average, typical properties of the environment. . . . A second demand is for rapid learning of new information so that a novel experience can be remembered after a single occurrence. After all, people can at least sometimes learn things by being told once" (Smith & DeCoster, 2000, p. 109).

Smith and DeCoster (2000) define rules as symbolically represented and structured by language and logic. "Symbolic rules may constitute a formal system such as the laws of arithmetic or of logical inference that is accepted by social consensus in a way that goes beyond its inherent persuasiveness for any given individual" (p. 114). Note that this definition is narrower than the general "if–then" conception of rules as contingencies (see, e.g., Anderson, 1983; Holyoak et al., 1989; Rescorla & Wagner, 1972; Tolman, 1932). There is nothing inherently wrong in defining rules the way Smith and DeCoster do, yet doing so has substantial implications. Thus, if to qualify as a rule a cognitive relation needs to be stated in symbolic terms, and be part of a formal system such as arithmetic or logic, then conditioning (of animal or human behavior) could not possibly be rule based contrary to recent agreements that *classical conditioning* is a form of signal or expectancy learning, involving "if–then" anticipations, generally viewed as "rules" (for re-

views, see Baeyens & De Houwer, 1995; De Houwer, Thomas, & Baeyens, 2001; Holyoak et al. 1989).

Furthermore, in Smith and DeCoster's (2000) framework, heuristics such as "expertise implies correctness" or "careful manner of dressing and persuasive argumentation are typical of lawyers" that may vary in their inherent persuasiveness to different individuals (Trope & Ginossar, 1988) could not qualify as rules even though they are pervasively defined as "rules of thumb" in the social judgment literature (Chaiken et al., 1989; Tversky & Kahneman, 1973).[1] To summarize then, the widely accepted definition of the rule concept in terms of its conditional, "if–then" structure is broader than Smith and DeCoster's definition that characterizes the rule concept in terms of its contents (symbolic, formal) as well as in the degree of social consensus it commands (see above).

Beyond definitional matters, the properties of rules proposed by Smith and DeCoster (2000) raise questions as to their generality. Thus, (symbolic) rules are assumed to be learned fast yet we know from experience how slow, difficult, and arduous can be the learning of statistical or logical rules (Evans, 1982; Kahneman, 2003; Tversky & Kahneman, 1973; Wason & Johnson-Laird, 1972). On the other hand, the learning of what Smith and DeCoster would define as "associations" could be exceedingly fast. Taste aversion is one striking instance of such rapid learning. Furthermore, evidence exists that evaluative conditioning (EC) can also occur in a minimal number of trials, even one (Baeyens, Crombez, Hendrickx, & Eelen, 1995; Martin & Levey, 1978, 1987, 1994; Stuart, Shimp, & Engle, 1987). Thus, apparently it is not necessary for associations to be built slowly over time, and/or for rules to be acquired very quickly.

According to Smith and DeCoster (2000):

> Rule based processing . . . tends to be analytic, rather than based on overall or global similarity, for example a symbolic rule may single out one or two specific features of an object to be used in categorization, based on conceptual knowledge of the category. In contrast, associative processing categorizes objects non analytically on the basis of their overall similarity to category prototypes or known exemplars. (p. 112)

The question, however, is whether phenomena such as classical conditioning, which have been typically regarded as quintessentially associative, actually are based on perceptions of *overall* similarity versus on specific features that seem to predict a given event (e.g., the onset of the US). In this vein, Holyoak and colleagues (1989) noted:

> Unless a feature is included in a candidate rule, nothing can be learned about its relation to other features or to appropriate behaviors. Moreover, a complex environment may contain many features, few of which are likely to be cues that would help form useful rules. For example, a rat may receive a shock while listening to an unfamiliar tone, scratching itself, looking left, and smelling food pellets. Intuitively, we might expect that the rule "if tone, then expect shock" will more likely be generated in this situation than the rule "If looking left, scratching, and smelling pellets, then shock" . . . unusual features of the environment are favored as candidates to build the conditions of new rules. (pp. 320–321)

In a similar vein, Murphy and Medin (1985) noted:

> A . . . major complication derives from the fact that no constraints have been provided on what is to count as a feature or property in analyses of similarity . . . any two entities can be arbitrarily similar or dissimilar by changing the criterion of what counts as a relevant attribute. Unless one can specify such criteria, then the claim that categorization is based on attribute matching is almost entirely vacuous. (p. 292).

The philosopher Nelson Goodman (1972) expressed this emphatically: "Similarity, ever ready to solve philosophical problems and overcome obstacles, is a pretender, an imposter, a quack. It has, indeed, its place and its uses, but is more often found where it does not belong, professing powers it does not possess" (p. 437).

Whereas Holyoak et al. (1989) emphasized that what determines the formation of associative rules is the principle of "unusualness" (i.e., salience) it is likely that other factors may also play a role. For instance, accessibility, which is different from salience (Higgins, 1996), may also determine the stimulus feature attended to and that one associates with another feature. Presumably then, if the construct "drink" was accessible in one's mind, one might more readily form an association between "drink" and "accident" if a person was seen to have a drink and subsequently be involved in an accident, compared to the case wherein "speed" was accessible in one's mind in which case an association between "speed" and "accident" might form more readily.

From the present perspective then, an association can be thought of as an expectancy, or as a conditional, "if–then," rule whereby given an antecedent term the consequent term may be expected. The contents of the consequent term may vary widely in the different rules. The kind of rules involved in conditioning phenomena are *predictive* in their content (Holyoak et al., 1989), for instance, an animal may learn to predict an important event such as the occurrence of a shock on the basis of a previously sounded tone. Thus, in conditioning studies the consequent of the conditional refers to a future occurrence of some event (e.g., the onset of the U.S.).

Other rules may have different consequent terms. Consider a semantic association between the terms "doctor" and "nurse." Clearly, encountering a doctor does not imply that a nurse will soon appear. The rule here is based on individuals' lay theories of what the "doctor" concept means, including, for example, the predicate "works with nurses" (representing the rule "if doctor then works with nurses"). In the same way that a tone may signal to an animal, and in this sense constitute evidence, that a shock will soon appear, knowledge that someone is a doctor may constitute evidence that this person works with nurses.

In summary, the distinction between slow associative learning and fast rule learning (Smith & DeCoster, 2000) is more complex than meets the eye. What has been traditionally viewed as associations can be learned relatively fast, whereas what has been characterized as rules may be acquired with difficulty and rather slowly. The notion that "associations" are learned on the basis of global simi-

larity and that "rules" are based on specific features too has been questioned (Holyoak et al., 1989; Murphy & Medin, 1985). Ultimately, associations can be thought of as "if–then" rules in which some term or category is conditionally linked with another. These arguments address the several dual systems models whose foundational assumption concerns the distinction between associative and rule-following (or propositional) processes (Gawronski & Bodenhausen, 2006; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004).

CONCLUDING COMMENTS

Vast Judgmental Diversity

The domain of social judgments is exceedingly diverse. First, judgments vary a great deal in terms of their contents: Causal attributions are judgments, but so are category assignments, stereotypical characterizations, behavioral identifications, likelihood estimates, personal impressions, attitudes, evaluations, and so on. Second, judgments vastly vary in speed and immediacy. Some are based on extended deliberations and a laborious examination of available evidence. Others have a "pop up" quality—they come to mind instantly, and they feel thoroughly natural, almost inevitable, under the circumstances. Judgments also may vary on the process-awareness dimension. In some cases (e.g., of attorneys in a court of law, or of academics arguing a point), the judgments rendered are exquisitely conscious and explicit. They often come in a written form wherein the details of the argumentation are fully transparent. In other cases, individuals may be hard pressed to explain what drove them to a given judgment, and their explanations might be wrong at that (Nisbett & Wilson, 1977; Wilson, Dunn, Kraft, & Lisle, 1989).

Small wonder then that a plethora of different models and theories were advanced to explain different judgmental phenomena. Some of these addressed specific content domains (e.g., models of stereotyping, of attribution, or of attitudes), others drew distinctions along the amount of processing continuum (Fiske & Neuberg, 1990; Petty & Cacioppo, 1986), yet others stressed the difference between "automatic" or "associative" and "deliberative" or "rule-following" ways of judging, including aspects of efficiency as well as awareness (Bargh, 1996; Devine, 1989). As a consequence of these various distinctions, the field of human judgment is highly fragmented these days, with different bodies of literature making little contact with each other, and the disparate research programs resembling those blind individuals in the adage, groping different parts of the same elephant.

But is there a deep-seated "elephant" behind the vast surface heterogeneity of judgmental phenomena, and how is it all put back together? (Anderson et al., 2004, p. 1036). It is in this vein that Newell (1990) argued for "the necessity of a theory that provides the total picture and explains the role of the parts and why they exist" (p. 18). In the present chapter we attempted to sketch an outline of such theory emphasizing the dimensional pa-

rameters along which different judgmental situations may vary.

The Principles of Social Judgment

Our conceptual framework affords the identification of several fundamental principles applicable to all of judgment. These are briefly elaborated below.

- *Principle 1*: *All judgments are rule based* (see also Anderson et al., 2004; Holyoak et al., 1989). This assertion is based on the notion that associations too are rule-like in the sense of denoting conditional "if–then" statements that may or may not be relevant to a judgmental question of interest. Nor is the concept of "pattern recognition" incompatible with rule-driven judgment, in the sense that a given stimulus pattern (e.g., of facial features, or of pathological symptoms) is assumed to warrant (or constitute a condition that justifies) a given judgment (e.g., that this face is Paul's, or that this patient has the flu) (for discussion, see Kruglanski, Erb, Pierro, Mannetti, & Chun, in press).
- *Principle 2*: *Judgmental task demands determine the amount of motivational and cognitive resources needed to appreciate the ("if–then") relevance of a given bit of information to a given judgment*. In some cases, the rule that lends the information its relevance may be highly accessible and, be activated immediately upon the individuals' encounter with the stimulus (Fazio, Sanbomatsu, Powell, & Kardes, 1986). In such circumstances the judgment (e.g., an evaluation) would be generated with considerable alacrity following the stimulus presentation. In other cases, the search for a relevant inference rule may be quite laborious. As a consequence, under limited attentional or motivational resources a less subjectively relevant, yet easier to access, rule may be utilized instead of a more relevant, yet more difficult to access, rule (Pierro et al., 2004).
- *Principle 3. The process of judgment formation is independent of the content of the inference rules being used*. Individuals may have at their disposal a wide variety of rules in different domains of interest. For instance, some of the rules may have to do with social stereotypes ("if X is member of category A, then X has properties a, b, c"), other rules may have to do with causal attributions (e.g., "If X preceded Y was contiguous to Y in time and space, and covaried with Y over time, then X is the cause of Y"), yet other rules may have to do with statistical inferences (e.g., "if X is chosen on the basis of its extreme standing in some distribution, then on a repeated observation X is likely to be less extreme"), and so on, yet the conditional ("if–then") structure is common to all inference rules, as is the relation between the difficulty of discerning or retrieving the rule (or task demandingness) and the role of individuals' cognitive and attentional resources in basing judgments on the information given. In other words, Principles 1 and 2 are assumed to apply across the multifarious content domains of judgment.
- *Principle 4*: *The rendition of all judgments is mediated by the same continuous parameters*, described earlier. Though the rendition of specific judgments may appear to differ

vastly from one instance of judgment to the next (some judgments may be rendered very fast, others rather slowly, some may be rendered spontaneously and/or unconsciously, others deliberately and in full awareness) the same general variables are involved in all instances of judgment formation. Thus, the speed of forming a judgment may depend on the degree to which the pertinent "if–then" rule has been routinized, In turn, the more routinized a rule is, the less conscious attention may be required for its application and as a consequence the less conscious an individual might be of its operation. From this perspective, judgmental speed and a lack of awareness do not necessarily indicate the operation of qualitatively different processes but may instead betoken routinization defining the level of task difficulty and the amount of resources necessary for appreciating the potential relevance of the information given to the requisite judgment.

NOTE

1. Indeed, Smith and DeCoster (2000) admonish Chaiken et al. (1989) for using the term "rules" in describing the heuristic processing system. According to their recommendation "if the representations used in heuristic processing were described as well learned associations rather than as rules, the distinction would be clearer" (p. 120).

REFERENCES

Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica*, *103*, 77–89.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood-Cliffs, NJ: Prentice Hall.

Anderson, J. R. (1983). *The structure of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S. Lebiere, C., & Quin, Y. (2004). An integrated theory of mind. *Psychological Review*, *111*, 1036–1060.

Baeyens, F., Crombez, G., Hendrickx, H., & Eelen, P. (1995). Parameters of human evaluative flavor-flavor conditioning. *Learning and Motivation*, *26*, 141–160.

Baeyens, F., & De Houwer, J. (1995). Evaluative conditioning is a qualitatively distinct form of classical conditioning: A reply to Davey (1994). *Behaviour Research and Therapy*, *33*, 825–831.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169–183). New York: Guilford Press.

Beauregard, K. S., & Dunning, D. (1998). Turning up the contrast: Self-enhancement motives prompt egocentric contrast effects in social judgments. *Journal of Personality and Social Psychology*, *74*, 606–621.

Blascovich, J., Ernst, J. M., Tomaka, J., Kelsey, R. M., Salomon, K. L., & Fazio, R. H. (1993). Attitude accessibility as a moderator of autonomic reactivity during decision making. *Journal of Personality and Social Psychology*, *64*, 165–176.

Bless, H., & Schwarz, N. (1998). Context effects in political judgment: Assimilation and contrast as a function of categorization processes. *European Journal of Social Psychology*, *28*, 159–172.

Bodenhausen, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology*, *66*, 621–632.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cog-*

nition: A dual process model of impression formation (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*, 123–152.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752–766.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance in attribute judgment. *Journal of Personality and Social Psychology*, *66*, 460–473.

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation and the construction of values. *Organizational Behaviors and Human Decision Processes*, *79*, 115–153.

Chun, W. Y., & Kruglanski, A. W. (2006). The role of task demands and processing resource in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*, *91*, 205–217.

Chun, W. Y., Spiegel, S., & Kruglanski, A. W. (2002). Assimilative behavior identification can also be resource dependent: A unimodel perspective on personal-attribution phases. *Journal of Personality and Social Psychology*, *83*, 542–555.

Cohen, J. (1979). On the psychology of prediction: Whose is the fallacy? *Cognition*, *7*, 385–407.

Cohen, J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*, 317–370.

Cole, M., & Scribner, S. (1974). *Culture and thought: A psychological introduction.* Oxford, UK: Wiley.

Conrey, F.R., Sherman, J.W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad-model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469–487.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*, 853–869.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Dunning, D. (1999). A newer look: Motivated social cognition and the schematic representation of social concepts. *Psychological Inquiry*, *10*, 1–11.

Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology*, *63*, 341–355.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* New York: Psychology Press.

Eiser, J. R. (1990). *Social judgment.* Belmont, CA: Brooks/Cole.

Erb, H.-P., Kruglanski, A. W., Chun, W. Y., Pierro, A., Mannetti, L., & Spiegel, S. (2003). Searching for commonalities in human judgment: The parametric unimodel and its dual mode alternatives. *European Review of Social Psychology*, *14*, 1–47.

Erb, H.-P., Pierro, A., Mannetti, L., Spiegel, S., & Kruglanski, A. W. (2003). *Persuasion according to the unimodel.* Unpublished manuscript, University of Bonn, Germany.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

Fazio, R. H., Blascovich, J., & Driscoll, D. M. (1992). On the functional value of attitudes: The influence of accessible attitudes upon the ease and quality of decision making. *Personality and Social Psychology Bulletin*, *18*, 388–401.

Fazio, R. H., Ledbetter, J. E., & Towles-Schwen, T. (2000). On the costs of accessible attitudes: Detecting that the attitude object has changed. *Journal of Personality and Social Psychology*, *78*, 197–210.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Fiedler, K., Armbruster, T., Nickel, S., Walther, E., & Asbeck, J. (1996). Constructive biases in social judgment: Experiments on the self-verification of question contents. *Journal of Personality and Social Psychology*, *71*, 861–873.

Fishbach, A., Igou, E., & Kruglanski, A. W. (2005). *Ease of retrieval as an inference rule.* Unpublished manuscript, University of Maryland.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231–254). New York: Guilford Press.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Forgas, J. P. (1991). *Emotion and social judgments.* Elmsford, NY: Pergamon.

Forgas, J. P., Williams, K. D., & von Hippel, W. (2003). *Social judgments: Implicit and explicit processes.* New York: Cambridge University Press.

Garcia, J., McGowan, B. K., Ervin, F., & Koelling, R. (1968). Cues: Their relative effectiveness as reinforcers. *Science*, *160*, 794–795.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, *103*, 592–596.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Gigerenzer, G., & Hoffrage, U. (1999). Overcoming difficulties in Bayesian reasoning: A reply to Lewis and Keren (1999) and Mellers and McGraw (1999). *Psychological Review*, *106*, 425–430.

Gilbert, D. T., Gill, M. J., & Wilson, T. D. (2002). The future is now: Temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, *88*, 430–444.

Gilovich, T., Medvec, V. H., & Savitsky, K. (2000). The spotlight effect in social judgment: An egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of Personality and Social Psychology*, *78*, 211–222.

Ginossar, Z., & Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology*, *16*, 228–242.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and social Psychology*, *52*, 464–474

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects* (pp. 437–447). Indianapolis, IN: Bobbs-Merrill.

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making.* Thousand Oaks, CA: Sage.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*, 141–154.

Higgins, E. T., & Stangor, C. (1988). A "change-of-standard" perspective on the relations among context, judgment, and memory. *Journal of Personality and Social Psychology*, *54*, 181–192.

Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, *93*, 75–88.

Holyoak, K. J., Koh, K., & Nisbett, R. E. (1989). A theory of conditioning: Inductive learning within rule-based hierarchies. *Psychological Review*, *96*, 315–340.

Hovland, C. I., & Sherif, M. (1952). Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology*, *47*, 822–832.

Hsee, C. K. (1996). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes*, *66*, 122–129.

Innes-Ker, Å., & Niedenthal, P. M. (2002). Emotion concepts and emotional states in social judgment and categorization. *Journal of Personality and Social Psychology*, *83*, 804–816.

Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology*, *65*, 69–84.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, *11*, 123–141.

Kassin, S. M., & Pryor, J. B. (1985). The development of social attribution processes. In J. B. Pryor & J. D. Day (Eds.), *The development of social cognition*. New York: Springer-Verlag.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–238). Lincoln: University of Nebraska Press.

Kelley, H. H. (1971). Attribution in social interaction. In E. E. Jones, D. E. Kanause, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1–26). Morristown, NJ: General Learning Press.

Klatzky, R. M., & Andersen, S. M. (1988). Category-specificity effects in social typing and personalization. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition: A dual process model of impression formation* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.

Kruglanski, A. W. (1996). Motivated social cognition: Principles of the interface. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Kruglanski, A. W. (1999). Motivation, cognition, and reality: Three memos for the next generation of research. *Psychological Inquiry*, *10*, 54–58.

Kruglanski, A. W. (2004). *The psychology of closed mindedness*. New York: Psychology Press.

Kruglanski, A. W., Chen, X., Pierro, A., Mannetti, L., Erb, H.-P., & Spiegel, S. (2006). Persuasion according to the unimodel. *Journal of Communication*, *56*(Suppl.), 510S–512S.

Kruglanski, A. W., Erb, H. P., Pierro, A. , Mannetti, L., & Chun, W. Y. (in press). On parametric continuities in the world of binary either ors. *Psychological Inquiry*.

Kruglanski, A., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects of impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, *19*, 448–468.

Kruglanski, A. W., & Mayseless, O. (1990). Classic and current social comparison research: Expanding the perspective. *Psychological Bulletin*, *108*, 195–208.

Kruglanski, A. W., Raviv, A., Bar-Tal, D., Raviv, A., Sharvit, K., Ellis, S., et al. (2005). Says who?: Epistemic authority effects in social judgment. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 343–392). New York: Academic Press.

Kruglanski, A. W., & Thompson, E. P. (1999a). The illusory second mode or, the cue *is* the message. *Psychological Inquiry*, *10*, 182–193.

Kruglanski, A. W., & Thompson, E. P. (1999b). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Kruglanski, A. W., Thompson, E. P., & Spiegel, S. (1999). Separate or equal?: Bimodal notions of persuasion and a single-process "unimodel." In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 293–313). New York: Guilford Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263–283.

Kübler-Ross, E. (1969). *On death and dying*. New York: Macmillan.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kunda, Z., & Sanitioso, R. (1989). Motivated changes in the self-concept. *Journal of Experimental Social Psychology*, *25*, 272–285.

Kunda, Z., & Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application, and inhibition. *Psychological Inquiry*, *10*, 12–22.

Lambert, A. J. (1995). Stereotypes and social judgment: The consequences of group variability. *Journal of Personality and Social Psychology*, *68*, 388–403.

Lichtenstein, S., Slovic, P., Fischoff, B., Layman, M., & Coombs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 551–578.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflexion and Reflection: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). New York: Academic Press.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy*, *1*, 57–102.

Martin, I., & Levey, A. B. (1987). Learning what will happen next: Conditioning, evaluation and cognitive processes. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 57–82). New York: Wiley.

Martin, I., & Levey, A. B. (1994). The evaluative response: Primitive but necessary. *Behaviour Research and Therapy*, *32*, 301–305.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 195–216). Hillsdale, NJ: Erlbaum.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, *59*, 27–37.

Martin, L. L., & Tesser, A. (1992). *The construction of social judgments*. Hillsdale, NJ: Erlbaum.

Medin, D. (1988). Social categorization: Structures, processes, and purposes. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition: A dual process model of impression formation* (Vol. 1, pp. 119–126). Hillsdale, NJ: Erlbaum.

Moskowitz, G. B., & Roman, R. J. (1992). Spontaneous trait inferences as self-generated primes: Implications for conscious social

judgment. *Journal of Personality and Social Psychology*, *62*, 728–738.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.

Mussweiler, T., & Strack, F. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology*, *78*, 1038–1052.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, *238*, 625–631.

Nisbett, R. E., & Ross L. (1980), *Human inference: Strategies and shortcomings of social judgement*. Englewood Cliffs, NJ: Prentice Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *87*, 231–259.

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.

Petty, R. E. (1994). Two routes to persuasion: State of the art. In G. d'Ydewalle, P. Eelen, & P. Berteleson (Eds.), *International perspectives on psychological science* (Vol. 2, pp. 229–247). Hillsdale, NJ: Erlbaum.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). San Diego, CA: Academic Press.

Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology*, *29*, 137–165.

Petty, R. E., Wells, G. T., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, *34*, 874–884.

Pierro, A., Mannetti, L., Erb, H.-P., Spiegel, S., & Kruglanski, A. W. (2005). Informational length and order of presentation as determinants of persuasion. *Journal of Experimental Social Psychology*, *41*, 458–469.

Pierro, A., Mannetti, L., Kruglanski, A. W., & Sleeth-Keppler, D. (2004). Relevance override: On the reduced impact of "cues" under high motivation conditions of persuasion studies. *Journal of Personality and Social Psychology*, *86*, 251–264.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*, 1–5.

Rescorla, R. A. (1985). Conditioned inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299–326). Hillsdale, NJ: Erlbaum.

Rescorla, R. A., & Holland, P. C. (1982). Behavioral studies of associative learning in animals. *Annual Review of Psychology*, *33*, 265–308.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Procasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton-Century-Crofts.

Rothman, J. A., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, *24*, 1053–1064.

Ruder, M., & Bless, H. (2003). Mood and the reliance on the ease of retrieval heuristic. *Journal of Personality and Social Psychology*, *85*, 20–32.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, *2*, 87–99.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195–202.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford Press.

Scott, J. (2000). Rational choice theory. In G. Browning, A. Halcli, N. Hewlett, & F. Webster (Eds.), *Understanding contemporary society: Theories of the present* (pp. 25–39). Thousand Oaks, CA: Sage.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*, 380–400.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Sleeth-Keppler, D. (2004). *The effects of anchor-based semantic priming on judgmental anchoring*. Unpublished doctoral dissertation, University of Maryland, College Park.

Sloman, S. A. (1996). The empirical case of two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.

Smith, E. R. (1984). Model of social inference processes. *Psychological Review*, *91*, 392–413.

Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology*, *25*, 500–523.

Smith, E. R., & Branscombe, N. R. (1988). Category accessibility as implicit memory. *Journal of Experimental Social Psychology*, *24*, 490–504.

Smith, E. R., Branscombe, N. R., & Bormann, C. (1988). Generality of the effects of practice on social judgment tasks. *Journal of Personality and Social Psychology*, *54*, 385–395.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*, 437–446.

Stroebe, M. S., Hansson, R. O., Stroebe, W., & Schut, H. (2001), *Handbook of bereavement research: Consequences, coping and care*. Washington, DC: American Psychological Association Press.

Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research*, *14*, 334–49.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology*, *73*, 662–674.

Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology*, *79*, 344–354.

Trope, Y., & Ginosar, Z. (1988). On the use of statistical and nonstatistical knowledge: A problem-solving approach. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 209–230). Cambridge, UK: Cambridge University Press.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 239–270). New York: Guilford Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1130.

Uleman, J. S. (1987). Consciousness and control: The case of spontaneous trait inferences. *Personality and Social Psychology Bulletin*, *13*, 337–354.

Uleman, J. S. (1989). A framework for thinking intentionally about unintended thoughts. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 425–449). New York: Guilford Press.

Uleman, J. S., Winborne, W. C., Winter, L., & Shechter, D. (1986). Personality differences in spontaneous personality inferences at encoding. *Journal of Personality and Social Psychology*, *51*, 396–403.

Vallacher, R. R., Nowak, A., & Kaufman, J. (1994). Intrinsic dynamics of social judgment. *Journal of Personality and Social Psychology*, *67*, 20–34.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin*, *22*, 1105–1115.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in correction for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). San Diego, CA: Academic Press.

Wilson, T. D., Centerbar, D. B., Kermer, D. A., & Gilbert, D. T. (2005). The pleasures of uncertainty: Prolonging positive moods in ways people do not anticipate. *Journal of Personality and Social Psychology*, *88*, 5–21.

Wilson, T. D., Dunn, D. S. Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 287–343). San Diego, CA: Academic Press.

Wilson, T. D., Houston, C., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 4*, 387–402.

Wood, W., Kallgren, C. A., & Preisler, R. M. (1985). Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology*, *21*, 73–85.

Wood, W., & Quinn, J. M. (2003). Forewarned and forearmed? Two meta-analytic syntheses of forewarnings of influence appeals. *Psychological Bulletin*, *129*, 119–138.

Wyer, R. S., Jr. (2004). *Social comprehension and judgment: The role of situation models, narratives, and implicit theories*. Mahwah, NJ: Erlbaum.

Automatic Thought

SUSAN M. ANDERSEN
GORDON B. MOSKOWITZ
IRENE V. BLAIR
BRIAN A. NOSEK

Much of everyday life is guided by psychological processes that are unintended and unobservable. The nature of such processes has been the focus of experimental social psychology for well over two decades. Theories and methods capturing automatic aspects of human thought include those addressing trait concepts that guide person judgment and decisions, attitudes that guide evaluations, stereotypes and prejudice that guide intergroup interaction, significant others that guide interpersonal relations, and overall, motives and goal states that guide action. Controlled processes are now considered in terms of exactly how they direct, alter, bypass, or override automatic cognition. Indeed, controlled processes may become automatic with repetition. The curious interplay between automatic and controlled processes is thus one source of the widespread interest in automatic thought.

In this chapter, we consider what is known about automatic processes across differing research domains, and do so by focusing on each of four basic processes: the availability of social knowledge, the activation of this stored knowledge, its application to a new person or stimulus, and the question of whether and how it is self-regulated (Higgins, 1996a, 1996b). These processes involve varying degrees of automatic thought—collectively, they inform how social knowledge is used. The research domains through which we conceptualize the automaticity of each basic process cover much of social psychology. These include trait inferences, attitudes, stereotyping, significant personal relationships, and motives. Our goal in reviewing these research domains is to identify common themes and emergent trends that speak to these basic processes. We do not offer a comprehen-

sive review of each subfield but address pertinent work on which we may derive a metatheoretical characterization of automatic thought.

We begin with some brief definitions—the four components of automaticity, the four basic processes that comprise it, and the five research domains in which it is studied. We then present a characterization of five basic principles of automaticity that cut across these research domains and highlight questions and gaps in our knowledge that we have identified. These basic principles offer a framework for thinking about automatic thought and research that remains to be done. Following this presentation we briefly review the evidence across the four research domains to illustrate emergent themes by focusing on each process in turn (i.e., availability, activation, application, and regulation). Finally, we summarize with some concluding remarks.

DEFINITIONS

Automaticity: Four Components

Early conceptions of automaticity equated automatic processing with lack of conscious control (e.g., Hasher & Zacks, 1979; Hayes-Roth, 1977; Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). A process capable of running to completion without conscious monitoring was deemed automatic, even if initiated consciously. Given this definition, it is no surprise that automatic processes have often been equated with those that are *unconscious* and with methods that permit assessment of unconscious processing (e.g., subliminally

presented stimuli). If a person is unaware of a process, it occurred without conscious control. But automaticity can also be inferred when a process simply proceeds efficiently, or is involuntary, unstoppable, or unalterable even if conscious (Bargh, 1989, 1994). If initial stimuli are perceived consciously but a provoked process occurs outside awareness, or even despite attempts to prevent it, this is also automatic. Accordingly, automaticity is granted if the perceiver *lacks awareness* of the process, *does it with efficiency* (i.e., with minimal use of cognitive resources), has *no intention* to do it, or *cannot control* it. This approach frames automaticity as a continuous variable rather than as a discrete class of processes and as a matter of degree rather than kind. One or more aspects of automaticity may be in play in one circumstance but not in another, thus varying the form of automaticity (Bargh, 1994).

With the varieties of automaticity, interesting theoretical and social implications arise. Consider, for example, the implications of an attitude that influences judgment without one's awareness versus an attitude that has the same effect, *despite* one's awareness and attempts to avoid it. In practice, research tends not to delineate which type of automaticity is assumed or assessed, and few attempts are made to sort this out systematically. One consequence of this lack of specificity is the potential loss of information about how automatic processes operate in human judgment and behavior. Consistent with the literature, we nonetheless take an inclusive perspective on automaticity here, allowing that many procedures can be considered sufficient in demonstrating its presence. We hope that this inclusive approach will stimulate theoretical consideration of the nature of and variation in automaticity.

Four Basic Processes and How They Are Assessed

A useful and well-supported framework organizes psychologists' understanding of the automatic processes that guide the use of social knowledge (Higgins, 1996b; see also Higgins & King, 1981). First, social knowledge must be *available* in memory if it is to guide subsequent processing, and thus *availability* of social knowledge is a precursor to its use and understanding knowledge acquisition is thus relevant. Second, social knowledge must be accessible. *Accessibility* refers to a construct's readiness to be used, commonly defined as the degree to which it can be automatically *activated*. Third, accessible social knowledge may or may not be applied. *Applicability* refers to how well social knowledge matches attended-to features of a stimulus or situation. This match, or the usability of accessible social knowledge, determines in part whether or not the knowledge is applied. *Self-regulation* covers a heterogeneous set of processes that affect the accessibility and application of social knowledge such as inhibition, suppression, adjustment, or enhancement. Self-regulation can thus short-circuit activation at the outset, or redirect it once it has occurred, and can also prevent application or introduce a postapplication correction.

Given measurement challenges, there are often fuzzy boundaries between the processes of *accessibility* and *ap-*

plication. For example, a common method of assessing construct activation is one in which perceivers are exposed to a construct (e.g., hostile) as part of a perception task or comprehension task (as a prime), supposedly activating the construct. Activation is later assessed in a separate judgment task where impressions of a target are to be formed (e.g., hostile or not) based on a description of the target's (ambiguous) behavior (e.g., Srull & Wyer, 1979, 1980). In this method the automatic *application* of a construct is also serving as a proxy for automatic *activation* of the construct. The inference of automatic activation based on application is strengthened when it is clear that the individual is *unaware* of the activation and does not *intend* to apply such knowledge. For example, in Devine's (1989) classic study of automatic stereotypes, the activation of stereotypes was revealed through the judged aggressiveness of a target person following subliminal exposure to African American cues.

In distinguishing activation and application, the most direct measures of activation are those that are not necessarily "about" any particular target. Reaction time procedures, such as a lexical decision task, assess the comparative speed of judging whether or not a letter string is a word. When this judgment follows a prime in a priming task using a brief exposure of the prime stimulus, such as a Black or White face, the construct presented in the lexical decision task is primed/activated. Differential reaction times to different kinds of words, such as the word "good," as a function of the prime are thus taken as evidence of the automatic *activation* of evaluative knowledge in response to racial group members. These activation measures contrast with measures of application, which are quite directly "about" the stimulus cue (i.e., the person, group, concept, behavior in question), revealing the extent to which the social knowledge is (or is not) applied to the stimulus. This is termed "the aboutness principle" (Higgins, 1998) in which people seldom see an outcome or response as being "about" nothing (i.e., as "just occurring" at random). Increases in accessibility that are detected by the person (as a feeling of fluency) concern a particular concept and are seen as being "about" whatever target is salient in the context, and are thus applied to that target.

Beyond the conceptual relevance of these methods to the activation–application distinction, there is methodological promise in the realm of functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs) for illuminating automatic processes through the linkage to related brain structures and to physiological responses. These methods offer the potential to observe brain activation prior to or in conjunction with these measures of mental processes that assess behavioral responses.

Differing Knowledge Representations across Five Research Domains

Bruner (1957) argued that categorization serves the function of prediction. Knowledge that an object is a diamond informs us that it will be shiny and hard. The mere presence of cues linked to a category facilitates the trig-

gering of that category, while the activation of the category heightens perception of category-relevant information (and even broadens the scope of what is seen as category relevant). This pragmatism/functionalism was embraced by social psychologists at the dawn of the cognitive revolution in the field (e.g., Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979), and it remains central to research on social cognition and automatic processing. Indeed, to study how a person will behave, think, and feel in the context of social forces requires the assumption that the person will categorize for the purpose of prediction. Is this new person a threat? Is the person friendly and likely to behave as such?

Consideration of linguistic noun and adjective categories clarifies Bruner's fundamental insight into the relationship between categorization and prediction. Noun categories label the person or category (jock, politician, professor, Hispanics), while adjective features are featural descriptions of the person (tall, old, brown haired, athletic, educated, aggressive, kind, adventurous). Social psychologists study noun categories to identify the adjective concepts linked to them and to predict when such adjectives will be automatically activated and applied to make behavioral predictions about category members (Wyer & Srull, 1986). "Proper noun" categories, such as one designated by a significant other's name, are also of interest for their relevance to understanding automaticity in the self and in significant relationships (e.g., Higgins & King, 1981).

This noun–adjective distinction is of some interest because it helps in conceptualizing the distinction between the study of stereotypes (typically labeled by nouns) and the study of trait-labeled categories in simple linguistic terms. Nouns are used to tell us what an object is as a whole, and who a person is as a whole, while adjectives are used to identify features (i.e., parts), for example, of a person presumed. And this is presumed by laypersons to be pertinent to predicting what a relevant person will do. The meaning of a trait adjective can also be modified quite automatically by an activated category (e.g., a stereotype), and a trait can also modify the meaning of a category because of relevant linkages in memory that become differentially activated (Kunda, Sinclair, & Griffin, 1997). The boundaries and mutual influence between the two linguistic categories highlights the dynamism of social information processing (Blair, 2002).[1]

The social psychology of automaticity exists in multiple literatures, herein considered as five central research domains, each concerned with a specific form of knowledge representation and how it is used in inference and prediction. We define each research domain below. Afterward, we present our understanding of the basic principles of automaticity and follow with an integrative review of the evidence in each domain that speaks to these basic principles and processes of automatic thought as we conceive them.

TRAIT CONCEPTS AND TRAIT INFERENCES

People are not merely interested in *who* a person is (noun categories) but also in *what* qualities the person possesses (adjective categories) that guide what he or she will do. Thus, with automatic categorization follows automatic inference. Predictions about others are based, in part, on inferences about their underlying dispositions—personality traits imputed to the person—which are assumed to influence actions (e.g., Gilbert, 1989; Jones, 1979; Ross, 1977; Uleman, Newman, & Moskowitz, 1996; see Heider, 1944).

At its inception, social cognition was strongly identified with research on the adjective constructs used in perceiving others (Higgins & King, 1981). This research was typically pursued using priming paradigms in which incidental exposure to cues signifying a particular trait were shown to influence later social judgments (e.g., Higgins et al., 1977). Investigations of spontaneous trait inferences have also assessed how behavior is classified spontaneously using trait language—with little conscious awareness or intent (Winter & Uleman, 1984; Uleman, Newman, & Moskowitz, 1996). This research is typically pursued using memory paradigms in which incidental recall for behavioral information is assessed as aided by supposedly inferred adjectives as memory cues. Such inferences presumably explain "why" the individual did the behavior, reflecting what is known as the fundamental attribution error (Ross, 1977).

This literature forms one of the oldest bases of research on the role of automatic processes in social information processing. To varying degrees, the other four research domains owe a debt to this work.

ATTITUDES AND EVALUATION

Research on automatic evaluation has been situated in the domain of attitudes and has offered powerful evidence for its pervasiveness in social perception. An attitude is defined most simply as the association between an object and an evaluation (Fazio, 1986). Applied to a person or group, an evaluation, such as the word "good" (or "bad") is an adjective. When an attitude object (or its symbolic equivalent) is perceived, the object–evaluation association arises easily, unintentionally, and even outside awareness. For example, in a priming paradigm, words designating attitude objects are presented just prior to target words that match the evaluation of the word (or not). With few exceptions, people are faster at judging the target words when they are congruent in evaluative tone with the prime words, implying that the evaluation associated with the primed word was automatically activated and facilitated processing of the subsequent word. More broadly, the automatic evaluation effect has been shown to occur regardless of whether or not the attitude is strong—as it occurs even for weak attitudes (Bargh, Chaiken, Govendor, & Pratto, 1992), as well as for novel objects (Duckworth, Bargh, Garcia, & Chaiken, 2002).

STEREOTYPES

Given that social cognition as a field began with work on adjective or trait concepts, it is noteworthy that there has also been a focal interest in noun-labeled categories that

classify types of persons (Wyer & Srull, 1986). Stereotyping can involve a whole range of noun-labeled categories ("geek," "goth," "rapper," "politician," "professor," "male"), nominalized adjectives ("Black," "elderly") or noun phrases (e.g., "fundamentalist Christian," "liberal elitist," "corporate mogul," and "kid from the projects") (Andersen, Klatzky, & Murray, 1990; see also Andersen & Klatzky, 1987).

A stereotype is defined in terms of the attributes associated with a social category (and such a category may be dissociable from its evaluation; Amodio & Devine, 2006). Research in this domain typically exposes participants to a social category cue and examines whether subsequent judgments are influenced by attributes associated with that category. Perceivers are typically not aware that the cue is affecting their responses, suggesting the automatic activation and application of the stereotype. In a paradigmatic study on automatic stereotyping, participants were subliminally exposed to words associated with African Americans (Devine, 1989), and they subsequently came to judge another person as more hostile (compared to participants in a control condition). If a stereotype is available in memory, relevant cues can activate it without need of awareness, intention, effort, or control.

SIGNIFICANT-OTHER AND RELATIONAL REPRESENTATIONS

One-of-a-kind representations of persons that are noun labeled are denoted by a name and thus "proper" nouns (Higgins & King, 1981). Exemplars designating a specific person (Smith & Zarate, 1992) are *N*-of-one representations, and have been shown to function much like other social knowledge in activation and use. Research involving such highly familiar and significant mental representations in the life of the perceiver has shown that they can act as a lens through which new others are perceived (e.g., Andersen & Cole, 1990; Andersen, Glassman, Chen, & Cole, 1995).

In this work, the degree of match between cues encountered about a new person and a significant-other representation activates the latter, which is then used to make sense of the new person (Chen, Andersen, & Hinkley, 1999). Of primary interest is the fact that people appear to be unaware of doing this (e.g., Andersen & Glassman, 1996). In addition, significant-other representations do not stand alone in memory as they are linked with the self and, once triggered, alter the self and self-regulation (Baldwin, 1992; Baldwin, Carrell, & Lopez, 1990; Chen, Fitzsimons, & Andersen, in press). Contextual variability is thus central to what is known about these representations and the relational self (Andersen & Chen, 2002), in parallel with other social-cognitive work on the self (Mischel & Shoda, 1995).

MOTIVES AND GOAL STATES

Motives and goal states involve ends to be attained and behaviors that can achieve these ends. Goals studied by social psychologists are often interpersonal and involve states that a person seeks to achieve, or in fact "features" the person hopes to attain, and are thus examined in

terms of adjectives. Adjectives such as *free*, *thin*, *smart*, *cool*, *great*, or *egalitarian* can reflect the goals one has. Goals need not be adjective categories, however, as the goal of writing five pages a day is not an adjective. Goals can be defined by the actions that achieve them, suggesting that they can be conceived as verbs too. They are thought to differ from other representations in how much they energize behavior (Aarts et al., 2005; Bargh, 1990; Carver & Scheier, 1981; Gollwitzer & Moskowitz, 1996; Kruglanski, 1996). Goal states can be conceived as mental representations in memory (Bargh, 1990), and typically research on this conception incidentally exposes participants to situational cues associated with goals and assessing relevant behavior (Bargh & Barndollar, 1996; Bargh & Chartrand, 1999; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001; see also Aarts & Dijksterhuis, 2000). Planning in advance can also set the stage for automatic goal pursuit through implementation intentions (Gollwitzer, 1993, 1999), that is, by means of conscious commitment to a goal and the specification of *when*, *where*, and *how* one will respond to relevant situational cues. Implementation intentions promote goal attainment by removing the need for conscious reflection or thought of any kind at the moment of encountering the cues and initiating the action (e.g., Bayer, Moskowitz, & Gollwitzer, 2004 [as cited in Moskowitz, Li, & Kirk, 2004]; Brandstätter, Lengfelder, & Gollwitzer, 2001). In these ways, people regulate their behavior automatically to fulfill accessible goals (e.g., Bargh, 1990; Bargh et al., 2001).

WHAT WE KNOW AND WHAT WE DON'T KNOW: FIVE BASIC PRINCIPLES OF AUTOMATICITY

Automatic thought is best understood in terms of cognitive processes that transpire without effort, control, awareness, or intention (Bargh, 1989). The precise manifestations of these elements of automatic thought will depend in part on the social knowledge being used—and the research domains examined. Embracing the variability across these domains, however, enables consideration of emergent themes that span literatures and further illuminates a common core of the five basic principles of automatic processing.[2]

Automatic and Controlled Processes Are Not Discrete Opposites: Whither the Duel in Dual-Process Models?

In classic dual-process models of social cognition, automatic processes are assumed to occur first in the stream of information processing, and any altering of these processes to occur afterward, by overriding or correcting for the automatic process through a more controlled response (Chaiken & Trope, 1999). In automatic trait inference, an inference can be followed by correction, and if it is not, then correspondence bias (Gilbert, 1989) or priming effects (Thompson, Roman, Moskowitz, Chaiken, & Bargh, 1994) may emerge. When correction occurs, it is

often done inexactly, leading to contrast effects (Martin, Seta, & Crelia, 1990; Moskowitz & Skurnik, 1999; Strack & Hannover, 1996). In automatic stereotype activation too, relevant inferences can be followed by correction, and if not, stereotyping emerges (Devine, 1989). When correction occurs, stereotype control may emerge, but even this in itself can lead to stereotyping in rebound effects (Macrae, Bodenhausen, Milne, & Jetten, 1994; Monteith & Voils, 1998). In the attitude domain, correction can follow automatic belief in an assertion, and if it does not, an illusion of truth may emerge in which blatantly false information is believed to be true (Gilbert, Tafarodi, & Malone, 1993). Based on correction, information may correctly be identified as false. One may still be inaccurate by overcorrecting for this. Although there is no work we know of directly addressing correction or contrast effects in the study of significant-other representations, the consistency across other domains suggests that such effects would occur under the right conditions.

Dual-process models thus continue to capture a fundamental distinction in social cognition, but while they were once the reigning approach to studying automatic knowledge activation and use, they are now considered to be just one way that cognition is regulated. The simple temporal progression from automatic to controlled does not capture all that occurs. Contrast effects are often interpreted as deliberate attempts to correct for automatically activated constructs that are deemed to be biased. However, contrast can also occur automatically via simpler inhibitory and comparison processes, and some researchers have argued that they may occur exclusively automatically (Moskowitz & Skurnik, 1999; Stapel & Koomen, 2001b; Stapel, Koomen, & van der Pligt, 1997). Thus, evidence interpreted in terms of dual processing can at times be explained in simpler ways by considering the differing stimulus properties across the conditions of a study. One condition may simply activate different representations than another. Thus, differences across conditions might well be attributable to a change in the specific pattern of cues presented (an intricacy and multiplicity in cues that varies across the conditions). This suggests that different contexts may activate a different constellation of social knowledge and may do so automatically without dual processes.

Automatic Thought Is Context Dependent and Situation Specific: "It's the Context, Stupid"

That automatic thought is context dependent is better known than understood. A large body of work shows that subtle shifts in cues can substantially alter the automatic activation of social knowledge. If the situational specificity of thought is ignored, this central insight of social cognition is lost.

This is not to imply, of course, that baseline differences between people or between chronic tendencies to experience some social knowledge as more accessible than others, or particular self-regulatory patterns as more accessible, are unimportant. But even study of chronic differences in automatic processes between people requires examining variation in people's responses across different stimulus cues. Evidence shows that the if–then relation between individual and context is fundamental to automaticity, which has an intriguing parallel in social cognitive approaches to self and personality as well (Mischel & Shoda, 1995), suggesting that basic mechanisms of automatic thought are broadly relevant to understanding the person.

In short, the accessibility of available knowledge varies across contexts. Both chronic and transient sources of accessibility exist and play a role in increasing the relative accessibility of social knowledge with chronicity deriving from frequency of activation and transient accessibility deriving from recency (Higgins, 1996c). Knowledge that has been used frequently has a chronic readiness to be used (Bargh, 1999; Bargh & Thein, 1985; Higgins & King, 1981; Higgins, King, & Mavin, 1982), while transient accessibility arises based on triggering cues in the environment that temporarily increase social knowledge accessibility (Higgins et al., 1977; Srull & Wyer, 1979). Transient accessibility arises from a match between a cue to which one is exposed and an available representation when that cue is seen *before* a stimulus person is perceived (Higgins, 1989) or *while* the stimulus is perceived (Andersen et al., 1995). Context pulls from memory a representation and associated constructs, thus making the basic process of activation a shared venture between the representational system of the individual and the contextual cues that shape accessibility and retrieval. How this occurs remains under investigation.

Spreading activation models assume that knowledge is organized in memory through links or connections between isolated bits of information, or "nodes," in a network with the linkages reflecting the memory. When contextual cues trigger a memory, one or more nodes in the network are activated, and because these are linked through associations, spreading activation proceeds rapidly from one node to another as a function of the strength of association between the nodes. An act of kindness should thus facilitate the processing of associated concepts, with activation starting at the node directly triggered and traveling along whatever paths are present, flowing most easily and quickly along the paths that represent stronger associations. If the momentary accessibility of an available construct rises above threshold in any instance, the associated construct(s) will also be activated. This should occur automatically and thus quickly, "within 200 or 300 milliseconds [as] . . . [c]onscious processes take longer to develop—at least 500 or 600 milliseconds—and require considerable attentional resources" (Wegner & Bargh, 1998, p. 462).

Connectionist models, by contrast, tend to assume parallel distributed processing rather than activation that spreads serially across adjacent nodes (Rumelhart & McClelland, 1986). For example, Anderson's "ACT-R" model (Anderson, 1993) describes a systemwide set of associations that does not primarily emphasize nodes in a network organized as features and categories. Rather, it focuses on distributed activation that largely occurs all at once in parallel throughout the memory system (Kunda & Thagard, 1996; Smith, 1996), although concepts are still learned. To the degree that simultaneous activation

occurs, the dual-process models that dominated the 1980s and 1990s thus capture social cognition less well than competing models, about which we say more later.[3] For ease of discussion, we speak of categories and mental representations in this chapter, but we acknowledge that there is persistent debate about whether or not "representations" are in fact (entity-like) "things" in memory.

Indeed, regardless of the underlying structure and operation of the representational system, two basic principles emerge from this discussion. First, context provides a starting point for the rumblings of the representational system, determining the departure point from which some subset of available knowledge will be made relatively accessible. Second, because of the malleability of the process, the triggering and use of a concept alters the memory structure in a way consistent with the situational specificity. Thus, the construct triggered by and used in a specific context is altered accordingly in the process, such that whatever is stored is stored in modified form (see also Ross, 1989).

Intricacy in the Surround: Multiple Cues, Variable Representations

Cues are embedded in complex social contexts, which means there is rarely just one simple set of triggering cues that activate a representation without competition from others (Logan, 1989). Multiple knowledge categories often fit a given stimulus and context, and the final categorization may thus involve the concept that is retrieved most quickly (cf. Kunda & Thagard, 1996). This is why the strength of association between the stimulus and the relevant social knowledge, and also the relation between the stimulus and the social context, are so important. An Asian woman holding a pair of chopsticks elicits different associations than the same woman holding a makeup brush (Macrae, Bodenhausen, & Milne, 1995). The trait "aggressive" connotes one thing for an attorney in a courtroom, while this same trait connotes different associations for a "thug" in a back alley (Kunda et al., 1997). Numerous cues in the social context are processed simultaneously, thus constraining or enabling particular interpretations by inhibiting or facilitating concept activation.

Contexts Also Provoke Application Directly

Application occurs when knowledge is used to act on a goal or to draw inferences (Bruner, 1957; Higgins, 1989, 1996b; Higgins & King, 1981; Kunda & Sinclair, 1999; Kunda & Spencer, 2003). Although accessibility of social knowledge is a necessary condition for knowledge application, it is not sufficient. Delineating the conditions that lead to application has been a central concern to the field.

To a large degree, *applicability* of activated knowledge depends on social context. Knowledge that has been activated is more likely to be applied if it is *relevant*. When "aggression" is accessible, but seen as irrelevant to "Jane" due to gender stereotypes, it is not used in judgments of Jane (Banaji, Hardin, & Rothman, 1993). Similarly, when

people are able to identify why accessible information happens to be accessible, and it is also inapplicable, they may not use the accessible knowledge. This was demonstrated in a classic study showing that people offer more positive evaluations of their lives on sunny than on rainy days, but not when asked about the weather before making the global rating (see Schwarz & Clore, 1988).

Application also depends on the response. Activated racial bias may not be applied in explicit judgments of a group member but may alternatively be applied if assessed in nonverbal behavior (Amodio & Devine, 2006; Dovidio, Kawakami, & Gaertner, 2002; McConnell & Leibold, 2001). Activation of the elderly stereotype may speed the processing of stereotype-consistent versus stereotype-inconsistent information but simultaneously have a slowing effect on the respondent's behavior (Kawakami, Young, & Dovidio, 2002). The significance or "selection" of each response likely depends on the immediate context.

Another form of application that may be constrained by context is the misattribution people make about ease of processing, using it as information in their subsequent inferences. The ease with which information comes to mind influences various judgments (Schwarz et al., 1991; Schwarz & Clore, 1988) and inferences (Higgins & King, 1981; Tversky & Kahneman, 1973). People unconsciously misattribute perceptual fluency (induced by prior exposure) to how "famous" they think a stimulus person is (Jacoby, Kelley, Brown, & Jasechko, 1989) or to the confidence they have in their own beliefs, imputing more truth value to their own beliefs (Kelley & Lindsay, 1993; Skurnik, Moskowitz, & Johnson, 2006). Confidence in one's own thoughts increases with practice (Tormala, Petty, & Briñol, 2002), as does certainty in attitudes (Haddock, Rothman, Reber, & Schwarz, 1999) and even in negative expectancies about the future (Miranda & Andersen, 2005).

The relevance of contextual cues to stored social knowledge is likely to be calculated based on some kind of similarity metric (e.g., Tversky, 1977), reflecting how well the knowledge maps onto the stimulus (see Medin, Goldstone, & Gentner, 1993). This general principle may apply both to whether or not activation occurs and to whether or not such activated knowledge is applied.

Intentionality and Automaticity Are Not Opponent Processes: Don't Think, Regulate

People like to think of themselves as in control of their everyday actions. As such, the most intriguing questions about automaticity are those that concern the interaction of automaticity with intentional, controlled, effortful, and aware processing. What impact do the conscious processes that we treasure as essential to human nature actually have on thought and action? The accumulated evidence suggests a complicated interplay between automatic and controlled processes. As an example, while reading is automatic for literate adults, mundane behaviors such as picking up a book (or putting it down) are commonly subjected to intentional control. Likewise, even the dynamic process of physical balance or the

behavior of walking or of playing basketball reflect an interaction between automatic processes and volitional planning and execution. It is a dynamic interplay and is complex.

An especially complex issue is thus whether automatic thought can at times be considered intentional (Bargh & Gollwitzer, 1994). Automaticity is a heterogeneous construct that is inclusive of events with one or more of the four components of automaticity we have described—unintentionality, lack of awareness, uncontrollability, and efficiency in processing (Bargh, 1994). Because only one feature is necessary to assume automaticity, there is ample room for a wide variety of actions to be labeled "automatic" but to simultaneously possess one or more features of automaticity. Persons with obsessive–compulsive disorder may, for example, be perfectly aware of the thoughts relating to their obsessions or compulsions and yet still pursue them automatically because the operation of these acts is unintended and uncontrollable. Similarly, the notion of "intentional automaticity" occurs following practice and routinization of goals in recurring contexts. And, intended automaticity can occur when people establish an intention to pursue a goal when a particular situation occurs at a future time. The goal-relevant behavior can then be initiated automatically when relevant situational cues are encountered (Gollwitzer, 1999).

The interactive quality of automatic and controlled processes has shifted the understanding of self-regulation in social psychology. The answer to the question of whether or not self-regulation is conscious appears to be that "the field has evolved from a tentative answer of 'yes' to a firmer 'no' " (Vohs & Baumeister, 2004, p. 2). While dual-process models remain relevant to understanding the interplay between automatic activation and conscious intention to control the effects of that activation, it is clear that regulation also has automatic components. For example, self-regulation of stereotyping can involve deliberate corrective processes that overturn activated stereotypes (Fiske & Neuberg, 1990), or automatic processes that prevent the activation of stereotypes to begin with (for reviews, see Blair, 2001; Moskowitz, Li, & Kirk, 2004). Furthermore, there may be little use in discussing self-regulation as either automatic or controlled, because regulatory processes themselves may emerge from automatic processes (Dijksterhuis, 2005; Wegner, 2002).

Control in Self-Regulation

The preceding discussion does not dismiss the fact that intentions are of relevance for effective functioning and psychological health. What is it that predicts whether a person will formulate a behavioral intention that would be of value for him or her? How do automatic processes feed conscious intention? Evidence suggests that deliberating on and setting a goal is more effortful than implementing it once the goal has been set (Gollwitzer, 1999). People who have a goal that they want to accomplish over a period of time can automate that goal pursuit. For example, people who wish to avoid being prejudiced can develop a strategy that will automate nonprejudiced processes (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). Those motivated to be egalitarian also appear to be sensitive to progress in reaching this goal, even experiencing guilt when faced with shortcomings in their own responses (Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; Plant & Devine, 1998). And, when they are made aware of their shortcomings, they compensate for them automatically (Moskowitz et al., 2004). A possible explanation is that a feedback loop exists that sends implicit signals provoking the person to try to change course (see Carver & Scheier, 1981). However, if people are too conflicted to formulate an intention to do something (e.g., quit smoking), it will prevent establishing such a link between the goal and the relevant situations—accordingly, this intention will not be pursued in automatic behavior. Thus, when behavior change would be adaptive for a person, the question of how people formulate new, more adaptive goals that oppose habitual ones becomes important, as does the question of how the individual can become less conflicted enough to hold the new goals as focal, and this warrants further research attention.

This is still more complicated because conscious and effortful suppression quite often does not work, and instead backfires, as the classic studies asking individuals not to think of a white bear show in revealing rebound effects (Wegner, 1994). In these studies, people think more of a white bear than they otherwise would (for a different example, see Macrae, Bodenhausen, et al., 1994). Suppression of emotional expression also leads to more (not fewer) problematic consequences in physiological reactivity, again suggesting conscious efforts to regulate may not work. At this point, the evidence in social cognition greatly emphasizes the effectiveness of automatic self-regulation to the exclusion of effective conscious and effortful forms. Although the complex interplay between automatic and nonautomatic processes can introduce the homunculus problem of who is the controller and who the controlled (Ansfield & Wegner, 1996; Bargh, 1990; Wegner & Bargh, 1998), it remains important to better understand how and when mindfulness can be effective in self-regulation, a matter that is not yet well understood.

Automaticity at the Various Stages of Self-Regulation

Self-regulation is a process that guides both the activation and application of social knowledge. A formal analysis of regulation typically begins "by distinguishing two features of control—a *control action* (the influence) and a *control criterion* (the direction). Control involves acting upon something until a certain criterion is reached" (Wegner & Bargh, 1998, p. 450). The process involves (1) activation of relevant end states; (2) activation of beliefs about appropriate opportunities and means to successfully pursue these end states; (3) actions taken to remove whatever (negative) discrepancy exists between one's current standing and the desired state; and (4) monitoring of how close one is to attaining the end state (Carver & Scheier, 1999). Control theories in general can apply

to inanimate mechanisms such as thermometers, as well as to nonconscious biological homeostatic processes. Thus, these processes need not be accompanied by consciousness, nor do they need even to be specific to human thought.

The first steps of regulation are the activation and inhibition of goals, and these can both occur implicitly. Conceivably, other stages can as well. For example, monitoring processes that keep unwanted thoughts out of consciousness (Förster & Higgins, 2005; Wegner, 1994) and provide emotional feedback about goal pursuit or failure (Strauman & Higgins, 1987) can occur implicitly. Selective attention to goal-relevant stimuli and means (Moskowitz, 2002; Shah & Kruglanski, 2002), and then compensatory processes intended to bring one closer to desired end states (Moskowitz, 2001) are also implicit. Indeed, neural processes associated with such compensatory control are engaged within milliseconds of the activation of bias, well before conscious awareness that this might be needed is possible (Amodio et al., 2004). Such compensatory processes include inhibition of newly activated goals that are incompatible with a current goal that is already active (Fishbach, Friedman, & Kruglanski, 2003). Thus, the goal can be inhibited or its effects redirected so that the activated knowledge is not used (Kruglanski et al., 2002). Regulation is thus both "lateral"—with several competing constructs activated and "racing" to the top to influence perception (see Bodenhausen & Macrae, 1998; Kunda & Thagaard, 1996)—and "vertical" in that activation of an abstract goal may inhibit the activation of an incompatible, specific goal.

Changes in Knowledge Representation Based on Disuse

What happens to knowledge representations that were at one time habitual or automatic once they become regularly suppressed, replaced, or overridden by other responses? Do these knowledge representations decay with disuse, or do they simply become less accessible? While answers to such questions are far from clear in the literature, research on attitudes suggests that prior knowledge may well be retained when competing knowledge is formed. Some theories suggest that automatic evaluations are products of prior knowledge that are overridden by newer representations when self-regulation is successful (Wilson, Lindsay, & Schooler, 2000). Connectionist and learning theories do not conceive of new representations as replacing old ones that are then eliminated but, rather, see experience as continuously shaping the distributed representation system, allowing the possibility that "old" representations may be recovered quickly with reexposure because the neural network reflects the evolving course of the organism's entire learning history (McClelland & Rumelhart, 1985; Rescorla & Wagner, 1972). The social psychological understanding of the evolution of use and disuse of social knowledge may thus profit from insights stemming from classic and modern learning theories, and continuing research on such matters is of importance.

Automaticity Is Influenced by Practice and by What Social Knowledge Forms in Development

Although acquisition of automaticity is less commonly the focus of social psychological research, there is little doubt of its importance. We thus highlight this neglected aspect of automaticity as a needed area of research focus. This area of research may illuminate the processes underlying social concept formation and how the use of such concepts may become routinized.

Any representation that is available in memory can, in principle, be activated and used automatically. However, the existing literature suggests that the likelihood of such automization is moderated by multiple factors, most notably practice and repetition. Early work on automaticity drew from research on skill acquisition and equated automaticity with mental habits (Wegner & Bargh, 1998). From this perspective, mental tasks once performed with conscious direction can become habitual and autonomous with practice (Jastrow, 1906; Logan, 1988; Smith & DeCoster, 1999). Mental operations that once consumed attentional resources and were processed serially come to be processed efficiently in parallel (Shiffrin & Schneider, 1977). Rehearsal increases the efficiency of the mental operation, reflecting an elaboration of the related mental representation that can increasingly proceed efficiently and effortlessly (Logan, 1988). In other words, rehearsal is a mechanism for acquiring automaticity.

Automatic processing may also develop more indirectly by means of frequent and consistent environmental exposure to particular stimuli. Through this alternate route, people observe and store covariation information in their environment, and this process can be unintentional, effortless, unconscious, or uncontrollable (Devine, 1989; Smith & Decoster, 1999). In short, knowledge need not be consciously or effortfully learned and then routinized to be available for subsequent processing.

Practice and repetition concern the "quality" of the representation. Another factor for acquisition of automaticity is understanding the relevance of the social context. For example, the synapse model (Higgins & King, 1981) proposes that available knowledge will accumulate "charge" that rises above a variable threshold on the basis of contextual cues. Even so, there exists little research on this proposed process and on whether it is contextually cued or cued by internally generated thought, or both. In sum, this is a frontier for automaticity research that remains to be explored. How do cues to a construct actually trigger the construct in memory and raise its activation level? Does a basic process mediate between availability and activation?

More broadly, the availability of social knowledge is tied to socialization and development as they influence concept acquisition. Again, although the links between the automaticity literature and the developmental literature on concept acquisition are not often drawn and not well understood, their examination is likely to be important for a full understanding of automatic thought.

Automaticity on the Brain:
Process Meets Structure in Social Neuroscience

The rapid maturation of social neuroscience has benefited research on the automatic activation and use of social knowledge. For one, it provides converging evidence for fundamental findings from behavioral data, linking automatic processes to brain structures whose functions have been identified separately from years of neuroscience research. This contribution is not to be underestimated—as the successful convergence of disparate literatures provides mutual validation of constructs and principles derived from each. Identifying inconsistencies between literatures provides fertile ground for theoretical and methodological innovation.

Although some theorists suggest that this growing literature has yet to yield new principles of automaticity, it nonetheless holds promise for doing so, in the continuing challenge of bridging the gap between behavioral and neuroscience research on automatic thought. One of the best elaborated links between neuroscience and automaticity research is suggested by evidence that the amygdala, an area linked with emotional processing, is involved in automatic evaluation of social objects such as racial groups (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003; Cunningham, Nezlek, & Banaji, 2004; Phelps et al., 2000; see also Amodio, Harmon-Jones, & Devine, 2003). Indeed, evidence shows that the difference in amygdala activity in response to Black and White faces is correlated with behavioral indicators of automatic evaluation such as the Implicit Association Test and startle eye-blink (Phelps et al., 2000). Moreover, the dynamic interplay between automatic evaluation (e.g., via the amygdala) and self-regulatory processes via the prefrontal cortex (that modulates automatic responses) has also been shown (Amodio et al., 2004; Amodio, Kubota, Harmon-Jones, & Devine, in press; Cunningham et al., 2004; Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005). Demonstrating the societal relevance of this kind of research, arousal responses assessed by skin conductance based on fear conditioning to opposite-race (vs. same-race) faces, have recently been shown to extinguish significantly more slowly—among both African American and Caucasian participants, although the effect is moderated by whether or not one has had a cross-race dating relationship (Olsson, Ebert, Banaji, & Phelps, 2005). Taken together, these investigations show convergence across research domains examining automaticity and should thus spark further theoretical and empirical development.

Overall, converging evidence offers support for the five basic principles we have outlined here, that is, evidence that (1) automaticity and control are not discrete and opposing classes of processes; (2) contexts and situations are central to what mental representations are activated, applied, and how regulation is carried out; (3) regulation does not require the conscious will; (4) practice and repetition, habit formation, and developmental processes are central to the acquisition of concepts (to the availability and varying accessibility of social knowledge), and thus to the transition from conscious to automatic processing; and (5) automatic processes are reflected in recruitment of specific brain regions and functions. This evidence stems from a variety of research domains in social psychology, and we now turn to the evidence for these claims.

BASIC PROCESSES IN DETAIL: WHAT FIVE RESEARCH DOMAINS REVEAL ABOUT AUTOMATICITY

We organize our review the evidence for these basic principles of automaticity in terms of four basic processes—availability, accessibility, application, and self-regulation—and in turn address each process in terms of five different research domains in social psychology—that is, trait inferences, attitudes, stereotypes, goals, and relationships—by giving the most illustrative examples of the evidence.

Accessibility

Construct Accessibility Is Context Dependent: Contexts and Situations Shape Automatic Activation

SOCIAL CONSTRUCTS

Automatic accessibility accumulates and fades like a charge in an energy cell (Higgins, Bargh, & Lombardi, 1985; Higgins & King, 1981), and does so based on the concept's recency and frequency of use (Higgins et al., 1985; for reviews, see Bargh, 1996; Higgins, 1996b). Situational exposure to relevant cues influences recent and frequent concept use. Recent exposures temporarily increase activation potential. Frequent exposures influence the concept's chronic activation readiness (e.g., Bargh, Bond, Lombardi, & Tota, 1986; Bargh, Lombardi, & Higgins, 1988; Bargh & Pratto, 1986; Bargh & Thein, 1985; Higgins, King, & Mavin, 1982). Chronically accessible concepts have less impact on judgments initially compared with temporarily accessible concepts but more impact as time elapses (Higgins, et al., 1985). In particular, when the delay is brief (15 seconds), *recently* primed concepts guide judgments, but when the delay is extended (2 minutes), *frequently* primed concepts guide judgments. In short, the *decay function is not uniform* and is tied to how often a prime was seen and to how much time transpired since (Bargh et al., 1988). Chronicity and transient cuing appear to have an additive impact such that recent priming of a chronically accessible concept will result in high accessibility reflecting both its recency and its chronicity (Bargh et al., 1986).

The effects of recency and frequency on accessibility are not dependent on awareness of the exposures or knowledge of the relations between exposures and subsequent judgments. For example, subliminal exposure to cues increases their accessibility (Bargh & Pietromonaco, 1982). Individuals will apply such recently activated cues that were made accessible in one situation to another situation, without recognition that the prior exposure in-

fluenced subsequent judgments (indeed, without recognizing that the prior exposure occurred). Participants subliminally exposed to a list of words (one at a time) in which 80% reflected the trait "hostile" later judged a target person to be more hostile than controls (exposed to 20% or 0% hostility-relevant traits) (Bargh & Pietromonaco, 1982). Similar effects are observed among people known to have chronic levels of accessibility who judge a person (vis-à-vis the accessible dimension) in a way that assimilates the person to the accessible trait despite the perceiver lacking awareness that the trait is accessible or that the behavior judged is relevant to their chronic states (Higgins et al., 1985). In brief, both recent and frequent activation contribute to automaticity, when the construct is also relevant or applicable.

SPONTANEOUS TRAIT INFERENCES

People infer traits from behavior, even without realizing it (Moskowitz & Roman, 1992; Trope & Alfieri, 1997; Trope & Liberman, 1993; Winter & Uleman, 1984). This insight contradicts the initial assumption in social cognition that an explicit goal is needed to form an impression (Srull & Wyer, 1989). It also suggests a role for automaticity in the classic literature on correspondent inference (Jones & Davis, 1965; Ross, 1977).

Participants exposed to sentences describing behaviors that imply traits (e.g., "the librarian carries the old woman's groceries across the street") implicitly form trait inferences (e.g., "helpful"), as indicated in subsequent recall of the sentence when cued with the implied trait (Winter & Uleman, 1984). Activation of a trait category triggered by a behavior leads the trait to become associated with the behavior, facilitating joint retrieval in cued recall even though people are unaware of having made a trait inference. Implicit memory also arises in greater ease in learning lists of words including earlier inferred traits (Carlston & Skowronski, 1994). People also automatically attribute behavior to causes, whether physical events or previous behaviors (Hassin, Bargh, & Uleman, 2002). Automaticity in trait inference is revealed in reaction time paradigms where participants read a sentence and immediately after judge whether a word was present in the sentence. When the word was a trait implied in the sentence (vs. not implied) reaction times slowed and error rates were heightened, suggesting that the trait was spontaneously inferred during sentence comprehension thus leading to confusion regarding whether it was actually present (Uleman, Hon, Roman, & Moskowitz, 1996).

Implicit inferences enable rapid interpretation that probably sustains beliefs that the world is understandable and predictable (Anderson & Deuser, 1993; Bruner, 1957; Heider, 1958; Jones, 1979; Pittman & Heller, 1987). The automaticity of trait inference thus has functional value by making predictions of social behavior, something that the mind does effortlessly. If this were not an automatic process, then much of humans' deliberative resources would be spent categorizing, generalizing, and generating predictions about future social behavior, or else the social world would be continuously surprising and unpredictable.

AUTOMATIC EVALUATION

Object-evaluation associations, or attitudes, are readily activated automatically from memory (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Early work demonstrated that the brief presentation of a prime word (e.g., doctor) could facilitate (speed up) responses to a semantically related target word presented in sequence (e.g., nurse; Neely, 1977; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Automatic evaluation effects are measured using a similar procedure in which an image or word denoting an attitude object is presented immediately prior to the presentation of a target adjective that must be judged as *good* or *bad*. When the time interval between the prime and target is relatively brief (e.g., 300 msec), evaluative judgments to the target word are faster to evaluatively congruent (vs. incongruent) target words. This effect illustrates the unintentional and uncontrollable aspects of automaticity. Because of the very short time between presentation of the prime and target, respondents have no opportunity to prevent evaluative activation of the prime or its influence on judgment of the target stimulus.

Automatic evaluation effects are robust and replicable across a wide variety of stimulus modalities and procedures, including subliminal presentation (Greenwald, Draine, & Abrams, 1996; Wittenbrink, Judd, & Park, 1997), familiar attitude objects (Bargh et al., 1992; Fazio et al., 1986), images (Fazio, Jackson, Dunton, & Williams, 1995), odors (Hermans, Baeyens, & Eelen, 1998), newly learned words (De Houwer, Baeyens, Vansteenwegen, & Eelen, 2000), and novel objects (Duckworth et al., 2002). Moreover, they occur not only when people make explicitly evaluative judgments (good–bad; Fazio et al., 1986) but also when they make nonevaluative judgments, such as a lexical decision task (word/nonword; Kemp-Wheeler & Hill, 1992; Wittenbrink et al., 1997), or simply pronouncing a target word (Bargh, Chaiken, Raymond, & Hymes, 1996; Glaser & Banaji, 1999). Current mood state appears to influence automatic evaluation too. Recent evidence reveals that robust priming effects that occur in positive mood states are reduced or eliminated in negative mood states. A prominent interpretation is that positive moods serve as a "go" signal that promotes finding connections between concepts, while negative moods serve as a "stop" signal indicating that something is not right, leading stimuli to be scrutinized individually (Storbeck & Clore, 2005).

A central debate in this literature was spawned by the question: Is automatic evaluation universal and unconditional (Bargh et al., 1992; Chaiken & Bargh, 1993), or does it occur *only* for strongly held attitudes (Fazio et al., 1986)? In priming paradigms, strength of association is taken as an indicator of the *degree* of accessibility. The fact that even novel stimuli (whether abstract art or novel words) elicit automatic evaluation (Duckworth et al.,

2002) suggests that the phenomenon is not restricted to preexisting and strongly held attitudes and that it can be evoked online. At the same time, the strength of the object-evaluation association moderates automatic evaluation (e.g., Roskos-Ewoldsen & Fazio, 1992), and "evidence suggests that attitude strength does play a moderating role in the activation of the attitude to the extent the object is being given conscious scrutiny (i.e., when it is the current focus of the person's goals); otherwise, any and all attitudes become activated automatically" (Wegner & Bargh, 1998, p. 469).

THE TRIGGERING OF STEREOTYPES

The accessibility of stereotypes is temporarily heightened by exposure to a category member (e.g., a Black face) or attributes of a stereotype (Bargh, 1999; Devine, 1989; Fiske & Neuberg, 1990), with even subliminal cues activating the stereotype (Bargh, Chen, & Burrows, 1996; Chen & Bargh, 1997). Automatic stereotype activation is assessed through outcomes such as response facilitation (Banaji & Hardin, 1996; Blair & Banaji, 1996; Macrae, Milne, & Bodenhausen, 1994; Moskowitz, Salomon, & Taylor, 2000; Wittenbrink et al., 1997), inhibition (Kawakami et al., 2000; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999), word completion (Gilbert & Hixon, 1991; Spencer, Fein, Wolfe, Fong, & Dunn, 1998), and social judgment (Devine, 1989; Lepore & Brown, 1997). Automatic stereotype activation may be limited in time course, occurring as early as 500 msec (e.g., after seeing a Black face), but may not last long (not past 12 minutes, Kunda, Davies, Adams, & Spencer, 2002). It is also less likely among people low in prejudice (Kawakami et al., 2000; Wittenbrink et al., 1997) or those internally motivated to avoid it (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2001; Moskowitz, Gollwitzer, et al., 1999).

ACTIVATING SIGNIFICANT OTHERS EVOKES INFERENCES, EVALUATION, AND EXPECTANCIES

Knowledge of significant others can be activated automatically (Andersen & Cole, 1990; Andersen et al., 1995; Baldwin, 1992; Fitzsimons & Bargh, 2003; Shah, 2003a, 2003b). In one example, when cues presented "about" a new person implicitly activate the significant-other representation (Andersen & Cole, 1990), participants come to infer that the new person has additional, unlearned features in common with their significant other. The accessibility of the associated significant other leads to an increased likelihood of falsely remembering features of their significant other as having been presented as descriptors of the new person (Andersen et al., 1995; Andersen & Baum, 1994; Andersen, Reznik, & Manzella, 1996; Baum & Andersen, 1999; Berk & Andersen, 2000; Glassman & Andersen, 1999a, 1999b; Hinkley & Andersen, 1996).[4] This effect can persist and even increase after a 1-week delay (Glassman & Andersen, 1999a) and is observed when the cues are presented outside awareness (Glassman & Andersen, 1999b). Significant-other representations are chronically accessible without priming (Andersen et al., 1995, study 1) and

are more accessible with even minimal triggering (Chen et al., 1999), with transient cuing and chronicity combining additively (Andersen et al., 1995; Baldwin et al., 1996). In this paradigm, automatic evaluation also occurs—with the favorability of the significant other automatically evoked in terms of what is applied to a new person who shares some features with the significant other (Andersen & Baum, 1994; Andersen et al., 1996; Berk & Andersen, 2000).

When a significant other's name or face is presented subliminally, followed by supraliminal Chinese ideographs, people evaluate these unfamiliar stimuli more positively than stimuli primed with a nonsignificant other (Banse, 1999). More positive affect is also evoked in participants' facial expressions while reading descriptions of a new person who resembles a positive (vs. a negative) significant other (Andersen et al., 1996)—and this occurs within a couple of seconds, implying intentionality may be unnecessary.

Representations of significant others are also of interest in research based on attachment theory focused on the role of caretaker responsiveness in development (Bowlby, 1969; Thompson, 1998). Attachment style can be evoked by subliminal cues related to a significant other (Mikulincer, Hirschberger, Nachmias, & Gillath, 2001), or involving the person's first name (Baldwin, Keelan, Fehr, Enns, Koh-Rangarajoo, 1996; Gillath et al., 2006), or consciously visualizing him or her (Mikulincer & Arad, 1999; Zayas & Shoda, 2005). Also, priming specific significant others can evoke distinct if–then expectancies related to attachment styles (Baldwin, Fehr, Keedian, Seidel, & Thomson, 1993). For example, securely attached participants responded faster than insecurely attached participants to positive outcome phrases (e.g., come through), while the opposite occurs when the phrase indicates a negative outcome (e.g., leave me). Likewise, automatic expectancies for social rejection, particularly by significant others, can exist on a chronic basis and be activated by contextual cues (Downey & Feldman, 1996), leading to a hypervigilance to threat cues.

TRIGGERING MOTIVES AND GOALS

Like other knowledge concepts, motives and goals can be activated automatically by cues in the environment (Bargh, 1990), and unlike semantic constructs, such as trait constructs, they are uniquely equipped to energize and guide behavior (Bargh et al., 2001). When activated by contextual cues, goal states evoke behavior (Aarts & Dijksterhuis, 2000; Bargh & Barndollar, 1996; Bargh & Chartrand, 1999). Triggering cues consist of any goal-relevant cues in the environment or in one's thoughts that evoke the desired end states (goals), which then evoke behavioral strategies for goal attainment (Bargh, 1990; Hommel, Muesseler, Aschersleben, & Prinz, 2001; Kornblum, Hasbroucq, & Osman, 1990; Kruglanski, 1996). When a goal is associated with a context, the context will then activate the goal and goal-relevant behavior automatically (Aarts & Dijksterhuis, 2003; Bargh & Gollwitzer, 1994). For example, viewing images of a li-

brary leads to faster lexical decision judgments for library-relevant goals and behaviors (e.g., to be quiet) relative to control words or nonwords, but only when participants expected to visit a library. Holding a goal in mind thus increases the accessibility of relevant behavioral norms.

Moreover, the activation of goal states and goal pursuit leads to positive affect once the goals are attained, and to tension when they are not (Aarts & Dijksterhuis, 2000; Carver & Scheier, 1998; Geen, 1995; Gollwitzer & Moskowitz, 1996). Goals thus function as other mental representations with specific activation consequences, and also as processes that operate on other mental representations, such as by introducing motivated shifts in automatic processing. What distinguishes goal priming from trait priming is that goals are not truly primed unless the affect associated with them is triggered (its incentive value, e.g., Custers & Aarts, 2005) and a tension state exists while the goal is unattained (Custers & Aarts, 2005; Moskowitz et al., 2004; Strahan, Spencer, & Zanna, 2002). For example, priming people with the goal of quenching their thirst activates the goal of drinking something (measured by selection of drinks in a later task), but only if the person is thirsty (Strahan et al., 2002). When initial tension is present, the prime triggers the goal. This does not imply that the goal must already be activated for automatic priming to occur. Rather, desiring an end state involves experiencing associated tension prior to goal satisfaction and is a precondition for nonconscious goal priming.

Goals are more pronounced among people who regard them as highly desirable (Aarts, Dijksterhuis, & DeVries, 2001; Aarts et al., 2005; Strahan et al., 2002), and the habitual pursuit of some goals over others leads them to become chronically accessible. Chronic goals have been widely shown to impact the way people process information. Some goals, such as to have control (Lefcourt, 1976; Pittman & D'Agostino, 1989; Pittman & Pittman, 1980; Rotter, 1966), to engage in cognitive activity (Cacioppo & Petty, 1982), to avoid uncertainty (Sorrentino & Short, 1986), and to have closure (Neuberg & Newsom, 1993), thus influence how information is perceived and categorized. In addition, a goal can be activated by contextually cuing a means to a goal. After subliminal priming with the means to achieving a goal, people respond more quickly to words reflecting behaviors that can achieve the goal than to control words—even when the task is irrelevant, such as deciding if a word is a trait or not (Shah & Kruglanski, 2003).

Complexity in Contextual Cues

In most circumstances, stimulus cues are complex, introducing a variety of possibilities for automatic categorization. The selection of cues for automatic activation is influenced by situational factors. For example, recent research made use of stimulus items that varied in evaluation (good or bad) and in semantic meaning (animal or religion) in a paradigm that pitted semantic priming against evaluative priming (Do animals prime other animals regardless of valence or do good words prime other

good words regardless of semantic category?) (Storbeck & Robinson, 2004). With four categories of primes and targets (good animals, bad animals, good religious terms, bad religious terms), the data showed little to no evaluative priming and substantial semantic priming. In other words, animals, good or bad, primed animals relative to religions; moreover, good animals did not prime good religions. At the same time, when the priming task involves stimuli from a single semantic category (animals) evaluative priming is observed, suggesting automatic evaluation may be conditional on the availability of obvious, alternative semantic features for categorization.

THE MULTIPLE CATEGORY PROBLEM AND CATEGORY SPECIFICITY

Group memberships are often activated automatically, influencing perceptions without intention or control, but if a stimulus belongs to multiple categories, what is automatically activated? The presence of multiple cues can be resolved by individuation, that is, by considering an individual's personal qualities, if one has the time and motivation (Fiske & Neuberg, 1990). However, one category is likely to become dominant in activation depending on the social context. In the competition to see which will be activated, some categories are automatically inhibited (while others remain active) due to inconsistencies among the categories and their fit with contextual cues (Bodenhausen & Macrae, 1998; Moskowitz et al., 2004; Sinclair & Kunda, 1999). An Asian female activates a different response after subliminal exposure to the prime "Asian" (vs. "female") (Macrae et al., 1995). In a lexical decision task following the "Asian" prime, response latencies show that Asian stereotypes are automatically activated while female stereotypes are inhibited, and vice versa following the "female" prime.

Exposure to a category label also activates multiple associated exemplars that have stereotype-relevant attributes (Rothbart, Sriram, & Davis-Stitt, 1996). Typical members of rival fraternities were more likely to be retrieved than atypical ones, which are inhibited. Automatic exemplar activation is thus relevant in stereotyping (Smith & Zarate, 1992; see also Andersen & Cole, 1990; Andersen & Glassman, 1996).

CONTEXTUAL SPECIFICITY AND THE MALLEABILITY OF AUTOMATIC EVALUATION, STEREOTYPES, AND GOALS

Accessible constructs are often assumed to be rooted in long-term experience and to be relatively insensitive to the situation (Bargh, 1997; Devine, 1989; Dovidio & Fazio, 1992). Over the last 10 years, a major shift has occurred with evidence indicating that activation depends on the social context. Automatic processes are responsive to immediate motives, goals, and contexts (Blair, 2002; Kunda & Sinclair, 1999; Kunda & Spencer, 2003; Kunda et al., 1997; Mitchell, Nosek, & Banaji, 2003; Moskowitz et al., 1999a).

Greater automatic negative evaluation of Black targets is evoked in a ghetto context (vs. in a religious context; Wittenbrink, Judd, & Park, 2001), for example, and a

weaker automatic negativity toward Blacks relative to Whites is observed immediately after exposure to admired African Americans and disliked Caucasians compared to a control condition (Dasgupta & Greenwald, 2001). Other research has shown that Black targets elicit more automatic negative evaluation than do Asian targets in a classroom but more automatic positive evaluation on a basketball court (Barden, Maddux, Petty, & Brewer, 2004). These effects suggest an interaction between group membership, social roles, and situations in determining automatic evaluation.

Simple manipulations such as having participants think about strong women (Blair, Ma, & Lenton, 2001) or provide information about well-known female leaders (Dasgupta & Asgari, 2004) can reduce automatic activation of gender stereotypes of women as weak rather than strong, as supporters rather than leaders. Moreover, evidence suggests that participants' belief that the world is dangerous is related to stronger automatic stereotypes associating African Americans with danger, but only when seated in a dark room (Schaller, Park, & Mueller, 2003).

Another component of social context is in the variety of cues that emanate from a single target. For example, a target's gaze direction influences the relevance of the target for the self. When a target person looks directly at the perceiver (vs. not) the target is more likely to be categorized (Macrae, Hood, Milne, Rowe, & Mason, 2002). Finally, context can present challenges that overburden the person, or processing tasks that usurp attentional or cognitive resources that would otherwise be required for stereotype activation (Gilbert & Hixon, 1991; Li & Moskowitz, 2005; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999).

Instead of emphasizing which cue (of multiple cues) is activated based on the context, research now emphasizes how multiple stereotypes and attributes may be activated simultaneously and automatically with each constraining the other's meaning and constructing a unique, emergent impression (Kunda & Thagard, 1996). This connectionist perspective is supported by computer simulations (Kunda & Thagard, 1996) and provides grounding for taking seriously the notion that automaticity is *truly* contextualized and does not just reflect which of a defined set of mental representations is activated (Mitchell et al., 2003).

GOAL ACTIVATION IS ALSO CONTEXTUALIZED IN SPECIFIC TERMS

Benign situations in which searching and exploring are permissible are known to trigger approach goals, whereas avoidance goals are associated with situations that are problematic and require rectifying action. Situations that call for specific types of motor actions can also determine whether approach or avoidance goals are activated. Contexts that involve tasks where one must contract flexor muscles trigger approach goals, whereas contexts that involve acting through the use of extensor muscles trigger avoidance goals (Friedman & Förster, 2000). Automaticity is shown by an inability to detect any

connection between the muscular activity and goal activation, as well as through the use of dependent variables that reflect automatic processing.

Observing another person in goal pursuit can also prime the goal in the perceiver if he or she values the goal (Aarts, Gollwitzer, & Hassin, 2004) in a kind of goal contagion where the other's goal is activated in one's own processing. This offers a naturalistic explanation for how goals are triggered automatically across various interpersonal situations—they are inferred from the behavior of others (see also discussion of such activation processes in trait inference by Moskowitz & Roman, 1992).

Furthermore, people with the goal of being egalitarian (and who strongly believe in this) reveal automatic activation of this construct when exposed to members of disadvantaged groups (Moskowitz et al., 2000). When such people are exposed to faces of Black men, the stereotype of African American men is inhibited and the goal of being egalitarian becomes more accessible (as measured in lexical decision judgments to goal-relevant vs. stereotype-relevant words; Moskowitz et al., 2000, 2004). Even strangers, then, can trigger goal concepts linked to specific actions. And if people have the goal of seeing the world as just (Lerner, 1980), this belief can be triggered by environmental cues that suggest injustice (Hafer, 2000).

INTENTIONALITY IN ACTIVATION AND LACK OF AWARENESS IN PROCESSING

Even when there are in fact intentions to judge or to evaluate, automatic effects of social knowledge are still observed. For example, despite the fact that respondents are intentionally evaluating novel targets, the unawareness component of automatic attitudes is shown in affect misattribution (Payne, Cheng, Govorun, & Stewart, 2005). That is, participants presented with a to-be-ignored prime (e.g., faces of Blacks or Whites) quickly followed by a novel pictograph to be rated for pleasantness exhibit an influence of these primes on their evaluations of novel pictographs. This suggests that people misattribute their affective reaction evoked by a prime to their evaluation of the novel object (Payne et al., 2005). Little awareness or control is involved, and warnings to discount the affective implications of the primes do not reduce the effect. In this case, intentions are directly opposite to the observed effects, which rely on lack of awareness and control.

By contrast, automaticity in goal activation may at times depend on having compatible intentions. Automatic goal triggering by a contextual cue is facilitated by the explicit intention to pursue the goal when the appropriate context occurs (Gollwitzer, 1999). Automatic goals enable behavior to be regulated automatically, and yet, it is the conscious valuing of the goal (of this desired end state) and having the intention to pursue it that enables automatic goal-directed behavior when relevant cues are encountered.

In other circumstances, goal activation parallels other social knowledge activation, showing that intentions are not necessary for goal activation. Once goal pursuit has

become routinized by being paired with specific contexts, the *explicit* intent to pursue the goal is no longer needed. The contextual cues alone can trigger the *implicit* intent to pursue the goal (Aarts & Dijksterhuis, 2000; Aarts et al., 2004; Bargh et al., 2001; Moskowitz et al., 2004; Shah & Kruglanski, 2002). Though needed initially, *explicit* intent recedes to the background as the affordance it established with the context takes hold.

Automatic Activation in Social Neuroscience as Triggered by Contextual Cues

Relevant neuroscience evidence on automatic evaluation concerns the amygdala—a small structure in the medial temporal lobe—which has been shown to be important for emotional learning (see LeDoux, 2000, for a review) and has been linked directly to fear conditioning in humans (Phelps et al., 2001). Amygdala activation is particularly pronounced in response to aversive cues. This work is informative about possible neurocortical underpinnings of automatic evaluation. Research shows that the amygdala is recruited in processing evaluative information, and that this occurs automatically and unconsciously (Morris, Ohman, & Dolan, 1998; Whalen et al., 1998). Research using fMRI reinforces the conclusion drawn from behavioral research that automatic evaluation occurs whether or not participants are intentionally evaluating stimuli (Bargh, Chaiken, et al., 1996). For example, participants rating famous names as *good* or *bad* (evaluative) or as *past* or *present* (nonevaluative) showed greater amygdala activity regardless of the task for the negative famous names (vs. positive famous names), suggesting immediate evaluation without intention (Cunningham et al., 2003).

To contrast automatic and controlled evaluative processing, the amygdala can be compared with the prefrontal cortex (PFC), a region thought to be involved in higher-order processes involving executive control and deliberative processing (Duncan & Owen, 2000; Stuss & Benson, 1984). In the study by Cunningham and colleagues (2003), activation of the ventrolateral PFC (but not the amygdala) was shown to be influenced by the evaluative ambivalence of target names, but only in the evaluation condition. This effect suggests that controlled, reflective processes via the PFC may be recruited to deal with attitudinally ambivalent cues when relevant to the response task. By contrast, automatic amygdala responsivity is relatively insensitive to the task demands. This neurological difference between automatic and controlled evaluative processes is thus consistent with social cognitive theories drawing this distinction (e.g., Cacioppo & Berntson, 2001; Greenwald & Banaji, 1995; Smith & DeCoster, 1999; Wilson et al., 2000).

The notion that amygdala activation may mediate automatic evaluation effects (Phelps et al., 2000) is also shown in greater amygdala activation in response to Black (vs. White) faces—as suggested by evidence that fMRI measures of amygdala activity to Black (vs. White) faces—is correlated with automatic racial attitudes tapped by the Implicit Association Test (IAT) and also by startle eye-blink to Black (vs. White) faces (Phelps et al.,

2000). This latter physiological measure is associated with threat, and thus is also correlated with the IAT assessing the automatic association of *good* and *bad* with *Blacks* and *Whites* (see Nosek, Greenwald, & Banaji, in press, for a review). Likewise, significantly greater amygdala activation was observed in response to subliminally presented Black (vs. White) faces (Cunningham et al., 2004) and startle–eyeblink responses revealed greater bias toward Black (vs. White) faces among individuals with less personal motivation to respond without prejudice (just 400 ms after a face was presented; Amodio et al., 2003). Together, this evidence supports the interpretation that evaluative processes are activated automatically (unconsciously and effortlessly) and consequently guide immediate responses to affectively charged information.

Recent evidence using ERPs of the brain has focused on the amplitude of the late positive potential (LPP), which is thought to measure categorization and which has shown differences when the target word and the context match in evaluation (Ito & Cacioppo, 2000; see also Crites, Cacioppo, Gardner, & Berntson, 1995). That is, even when participants are instructed to provide an overt response that is opposite to the evaluation of the stimulus (e.g., to say "positive" to a negative stimulus), or to perform a nonevaluative task (to make the choice "people" or "no people"; Ito & Cacioppo, 2000), there is a LPP difference when the evaluation of prime and target words match (vs. mismatch). This is the same context in which automatic evaluation without *awareness* or *intention* occurs (Bargh, Chaiken, et al., 1996).

Research using ERP and fMRI has also begun to chart the neural correlates of stereotype activation. ERPs have been examined in response to sentences that were stereotype congruent versus -incongruent, with regard to an antecedent noun and a reflexive pronoun (e.g., doctor, he, vs. doctor, she). Sentences that violate gender stereotypes produce a larger P600 response, compared to sentences that are stereotype congruent (Osterhout, Bersick, & McLaughlin, 1997). Similarly, both a later and larger P300 response to race stereotype-incongruent sentences emerges relative to congruent face-trait priming (Bartholow, Dickter, & Sestir, 2006). Importantly, this difference in the P300 is consistent with a response latency measure of stereotype activation from priming. Moreover, none of these purported indicators of stereotype activation (P600, P300, RT) is affected by alcohol intoxication, supporting their automaticity, whereas the ability to inhibit responses on stereotype-congruent trials was affected by alcohol consumption, and inhibition (vs. inability to inhibit) involved a stronger negative slow wave ERP response.

SIGNIFICANT-OTHER ACTIVATION

Neuroscience has also examined brain activity while people are exposed to pictures of human faces that are personally familiar. Exposure to a significant other's face versus one's own face leads to a markedly different pattern of brain activation as assessed by fMRI (Kircher et al., 2001) and distinct from that evoked by a

nonsignificant other face (Gobbini, Leibenluft, Santiago, & Haxby, 2004). A photograph of a close friend or family member leads to diminished right amygdala activation compared with one of a famous person or stranger, although a photo of one's own child evokes *stronger,* not weaker, amygdala activation, perhaps due to a vigilant protectiveness with one's own offspring (Leibenluft, Gobbini, Harrison, & Haxby, 2004). And, although familiarity (alone) also diminishes amygdala activation, this is less true for merely familiar than for significant-other faces (Gobbini et al., 2004); hence, the effect is not reducible to familiarity, although it varies depending on the exact nature of the relationship.

In sum, the evidence suggests that social neuroscience is contributing to our understanding of how automatic processes transpire.

Application

Contextual Triggering of Application

APPLICATION OF TRAIT CONCEPTS TO A TARGET PERSON

The more accessible a concept, the more likely that it will be applied to a potentially relevant, but ambiguous behavior (Bargh et al., 1986, 1988; Higgins et al., 1977, 1985; Srull & Wyer, 1979, 1980; for a review see Stapel & Koomen, 2001b). For example, subliminal priming of trait concepts increases temporary accessibility of the trait and, as a consequence, the trait shows an increased likelihood of being applied in making a trait inference about an ambiguous behavior (Bargh & Pietromonaco, 1982). Spontaneous trait inferences are also made based on observed behaviors and are applied to whoever is engaging in the behavior as if the trait is "about" the person (Carlston & Slowronski, 1994; Moskowitz, 1993; Todorov & Uleman, 2002). When a behavior implying a trait is paired with a face and with the instruction to remember it (Todorov & Uleman, 2002), exposure to the same face later paired with the implied trait that was *not* seen yields more false recognition of that trait, and a slower response to correctly say "no" to it having been presented. This shows that the trait becomes associated with the person (the face) in memory and is consistent with the two-stage inference and application model in which a target's behavior is first identified and then applied as an attribution to the target's stable qualities (Trope, 1986). It is equally possible, however, that there is simply one act of spontaneous trait inference activation that is always "about" the person, in which case activated trait concepts are simply applied to persons implicitly in judgment.

APPLICATION OF AUTOMATIC EVALUATION

The applicability of automatically activated attitudes has been defined by the *m*otivation and *o*pportunity as *de*terminants of attitude-to-behavior processes (MODE) model (Fazio, 1986), which predicts that accessible (stronger) attitudes are likely to be applied to judgments when perceivers have no motivation or opportunity to override it. Motivation and opportunity to override depend on numerous factors such as awareness of the evaluation, availability of cognitive resources to self-regulate, the intention to avoid the automatic response, and the controllability of the action. Most research on the application of automatic evaluation involves attitudes toward social groups and group members, especially in racial attitudes (i.e., prejudice).

Higher levels of automatic racial prejudice lead to lower-rated friendliness when interacting with a Black than a White person, as judged by both the target and by observers (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Dovidio et al., 2002; Fazio et al., 1995; McConnell & Leibold, 2001; for similar effects for antifat attitudes, see Bessenhoff & Sherman, 2000). Higher levels of automatic preferences for Whites compared to Blacks are also associated with more blinking, less eye contact, less speaking time, less smiling, and more speech hesitations and speech errors (with a Black vs. a White interaction partner; Dovidio et al., 1997; McConnell & Leibold, 2001). The same behaviors are not predicted by explicit racial prejudice and, moreover, automatically activated attitudes do not predict behaviors that are more carefully controlled (Dovidio et al., 1997, 2002). A meta-analysis of predictive validity studies with the IAT suggests that automatic attitudes are better predictors, relative to related self-report measures, of behavior in socially sensitive attitude and stereotyping domains (Poehlman, Uhlmann, Greenwald, & Banaji, 2005). Finally, in a rare study involving minorities as participants, African American participants' automatic prejudice toward their ingroup predicted whether they explicitly preferred a White versus Black partner in an intellectually challenging task (Ashburn-Nardo, Knowles, & Monteith, 2003).

APPLICATION OF STEREOTYPES

One factor influencing the automatic application of stereotypes is the degree to which a target possesses features typical of a stereotyped group (Blair, Chapleau, & Judd, 2005; Blair, Judd, & Chapleau, 2004b; Blair, Judd, & Fallman, 2004a; Blair, Judd, Sadler, & Jenkins, 2002; Deaux & Lewis, 1984; Livingston & Brewer, 2002; Maddox & Gray, 2002; Sczesny & Kuhnen, 2004; Uhlmann, Dasgupta, Elgueta, Greenwald, & Swanson, 2002). People with more Afrocentric facial features (even if they are White) are judged as more likely to possess stereotypically African American traits than those with less Afrocentric facial features (Blair et al., 2002; Blair, Judd, & Chapleau, 2004; Blair, Judd, & Fallman, 2004). This differential application of stereotypes occurs without awareness and despite attempts to avoid it (Blair et al., 2002; Blair, Judd, & Fallman, 2004). Moreover, when in a categorization mindset, stimulus objects pertinent to categories will stand out and be given attention (Higgins, 1996b). A recent fMRI study (Wheeler & Fiske, 2005) showed enhanced amygdala activity among White subjects in response to Black (vs. White) faces when the task was to judge age, but not when it was to look for a dot on each face (or to judge whether the person liked a particular vegetable). The first task ostensibly encouraged social

categorization by age, enhancing the salience of the target's social category memberships (including race). The goals of a task thus shape which features are attended to and the apparent applicability of an activated category.

APPLICATION OF
SIGNIFICANT-OTHER REPRESENTATIONS

Significant-other representations are clearly activated and applied automatically, perhaps even more so than other forms of social knowledge, with effects on inferences and memory (Andersen & Cole, 1990; Andersen et al., 1995; Chen et al., 1999; Glassman & Andersen, 1999a). Also, facial affect arises relatively immediately that matches the overall valence of the significant other in response to the relevant stimulus cues (e.g., Andersen et al., 1996). Although it is not clear that this facial affect is "about" this new person, each stimulus cue that evokes this affect is presumed to be "about" the new person. Automatic expectancies of acceptance or rejection, and also relevant overt behaviors, are applied in the encounter as well (Berk & Andersen, 2000). Advanced priming can trigger application of significant-other representations, just as applicability-based cues in a person may do so (Andersen et al., 1995; see Chen & Andersen, 1999). Both forms of transient contextual cuing (incidental stimuli presented in advance, i.e., priming, and cues "about" the target presented when learning about the target) can provoke automatic application, as shown in relevant inferences and memory. When primed with a significant-other representation, people also respond faster in a lexical decision in which they first read an "If . . . " statement (e.g., "If I fail . . .) and then a "then . . . " statement (e.g., "then I will be rejected") if it is consistent with their significant-other relationship (Baldwin & Sinclair, 1996).

APPLICATION OF GOALS IN BEHAVIOR

Several goals have been shown to silently direct the extent to which stereotypes are applied, such as the need for structure (e.g., Kruglanski & Freund, 1983; Neuberg & Newsom, 1993), the need to repair self-esteem (e.g., Fein & Spencer, 1997; Spencer et al., 1998), and accuracy goals (e.g., Weary, Jacobson, & Edward, 2001). For example, when uncertainty beliefs are primed, people show no evidence of stereotyping, instead rating a member of a stereotyped group equally likely to be guilty of a stereotype-consistent crime as a member of a non-stereotyped group (Weary et al., 2001). Once any goal is activated, it may presumably provoke consistent behavior. As long as the new stimulus context or person presents an opportunity to act in a goal-directed fashion, the behavior will be applied regardless of the source of goal activation (Bargh & Chartrand, 1999; Lewin, 1936).

Implicit goal application is readily assessed through task interruption and spontaneous resumption (e.g., Wicklund & Gollwitzer, 1982; Zeigarnik, 1927). People primed with a high-performance goal (vs. not) often choose to continue a challenging word-formation task (e.g., constructing words using lettered tiles) when asked to choose whether or not to move on to a task involving humorous cartoons (Bargh et al., 2001). Goals can also be applied in one's own behavior by observing another person perform a similar behavior (goal contagion), as noted (Aarts et al., 2004). Participants who read a script describing another person acting in a self-serving way (with the goal to earn money) or in an egalitarian way (with the goal to help others) later adopted the same goal in their own behavior.

Intentionality and Automatic Goal Application

Can goal application be automatic if it is intended? Consciously deciding to pursue a goal in advance of doing so enables goal pursuit to occur outside awareness. Relevant cues in the environment will then activate the goal, triggering preexisting behavioral intentions that predict goal pursuit. Participants instructed to formulate a plan about how and when to collect a coupon from an on-campus office (an implementation intention vs. simply being given the goal), while pursuing a second, distracting goal of going elsewhere in the same building, picked up the coupon 80% of the time versus 50% if they had no implementation intention (controls) (Aarts, Dijksterhuis, & Midden, 1999) (i.e., they applied the goal in their behavior).

Implementation intentions heighten *efficiency* in initiating and continuing actions in both experimental tasks showing the efficiency under cognitive load and in naturally occurring cognitive load conditions (Brandstätter et al., 2001). When heroin addicts in withdrawal and recovered heroin addicts were asked to write a resume, those for whom the task specified where and when to do this were more likely to get it done—regardless of withdrawal/recovery status (Brandstätter et al., 2001). Goals are thus automatically applied to behavior efficiently, regardless of load, when behavioral intentions are activated.

GOAL-RELEVANT COGNITION

When goals are activated in a context, they can be applied by directing goal-relevant cognition (Chartrand & Bargh, 1996; Moskowitz, 2001). When explicitly asked to form an impression of a target person, people recall more about the person than when asked simply to remember the same information (Hamilton, Katz, & Leirer, 1980). For years, the explicitness of the instruction was assumed to be critical for engaging the impression formation goal and its impact on memory. But, indirect priming of impression formation (vs. memorization) goals elicits the same effect, and participants report no awareness of having had the goal in the first place, let alone of its impact on memory (Chartrand & Bargh, 1996). As another example, people asked to explicitly form an impression (vs. memorize) show recall that correlates more with schema-incongruent information (presented online) but judgments that correlate more with schema-congruent information (Hastie & Kumar, 1979). Subliminal priming of these goals has similar effects (Chartrand & Bargh, 1996), suggesting that goals can be

activated and applied automatically, without the person even being aware of pursuing them. As a final example, memory is often organized around broader life tasks relevant at the time information is acquired (e.g., Nidenthal, Cantor, & Kihlstrom, 1985; Woike, Lavezzary, & Barsky, 2001). Priming a goal can thus promote differences in memory organization and encoding (Woike et al., 2001) such that when one has the goal of agency (vs. communion) primed, there is increased memory for differentiation *when* the information is agentic but not when it is communal.

APPLYING GOALS TO AVOID UNDESIRED STATES

Cues in the environment can also suggest to-be-avoided states and evoke goals relevant to avoidance of such states. For example, the goal of avoiding death may lead to processes that buffer these concerns, even when death-related concepts are primed outside awareness (Arndt, Greenberg, Pyszczynski, & Solomon, 1997). Preexisting goals and behavioral intentions are thus applied and may even be provoked by cues that are antagonistic to the goal.

This effect is not limited to concerns about death. Both a temptation (delicious but fattening foods) and a focal goal (dieting) trigger the goal of being health conscious in eating behavior (Fishbach et al., 2003). That is, participants implicitly primed with food temptations by having been placed in a room with magazines and photographs depicting fattening food or implicitly primed with the goal of dieting both chose an apple as a gift versus a chocolate bar more than control participants. This reflects application of the goal regardless of whether it was the goal or the temptation that was primed. Similarly, when the goal of looking well groomed is primed, people apply the goal in action even if new contextual cues are inconsistent with this goal (dirty shoes). When the goal is accessible (vs. not), it facilitates response times to instrumental actions (e.g., polishing) (Custers, 2006).

Application by Means of Concurrent Activation of the Self or of One's Goals

Automatic stereotypes are also concurrently activated and applied to the self-concept in terms of self-evaluation. The strength of women's automatic gender–math stereotypes (math = male) and gender identities (female = me) predicts the strength of self-stereotyping or identification with math (math = me) (Nosek, Banaji, & Greenwald, 2002a) when these links exist in memory (Greenwald et al., 2002). Women who strongly and automatically identify with their gender and strongly and automatically stereotype math as male automatically apply math to the self less, that is, they readily process this association. Stereotypes and identity are thus interdependent in automatic activation and this occurs by means of concurrent activation.

Concurrent activation is also shown when a significant-other representation is activated and this activation spreads to the self, leading what is applied to the self to shift as well (Andersen & Chen, 2002). One's sense of self

at the moment becomes infused with the relevant relational self with this other (Hinkley & Andersen, 1996), including relevant shifts in self-evaluation.[5] When a new person resembles a significant other versus does not, people describe themselves in more self-with-significant-other terms (covarying out pretest scores), showing automatic application to the self (Andersen, Reznik, & Glassman, 2005).

Likewise, exposing graduate students to the disapproving face of their advisor outside awareness or, for Catholics, the face of the Pope (vs. an unfamiliar frowning face), leads to shifts in self-evaluation (Baldwin et al., 1990). Visualizing a significant other intentionally can also have a negative impact on self-evaluation if the relationship involves contingent acceptance based on success or failure (Baldwin & Holmes, 1987). Activating a significant-other representation leads to self-application of the *standards* that the significant other holds for the self (Reznik & Andersen, 2006; see Andersen & Chen, 2002), and to application of the *interpersonal role* with the significant other to new encounters (Baum & Andersen, 1999).

The latter effects both involve concurrent goal activation as well. Activating a significant-other representation, such as one's mother, activates the *goals* with this parent (such as the goal to be close, to help, or to achieve), which are then applied to immediate ratings, social interactions, or tasks (Andersen et al., 1996; Berk & Andersen, 2000; Fitzsimons & Bargh, 2003; Shah, 2003a, 2003b). For example, when primed with the name of their mother, participants show enhanced performance on an anagram task when their mother values achievement (Fitzsimons & Bargh, 2003).

Beyond goals per se, automatic evaluation is also applied to the self when attachment style with significant others is activated, as shown in a Stroop task—with anxious–ambivalent individuals responding most slowly for negative, self-relevant traits, while dismissive individuals were slowest for positive traits, and secure individuals slowest for all self-relevant traits (Mikulincer & Shaver, 2004). Moreover, individuals who anxiously expect rejection by others overreact to it with hostility and rage, applying such terms to the self (Downey & Feldman, 1996).

Inhibition of Application

INHIBITION BASED ON CONTRADICTORY INFORMATION

The reigning assumption in the early automaticity literature was that activation virtually always produces application, and application was thus a good proxy for activation. The current literature now draws a distinction between activation and application because of evidence suggesting that activated traits are not always applied (Bargh, 1989; Bargh, Chen, et al., 1996; Higgins, 1996b). A primed trait concept may not be applied to a behavior when it matches on evaluation but mismatches on applicability. Priming a "hostile" animal name, such as tiger or shark, activates the construct "hostile," but it is not applied to a target person (Stapel et al., 1997). Likewise,

when there is dissimilarity between two significant others and one is activated, the other may be inhibited, as not all significant-other representations are used simultaneously (Baldwin, 1992; Hinkley & Andersen, 1996) and attachment style varies by relationship (Baldwin et al., 1996; Pierce & Lydon, 2001). Sometimes, however, contradictions do not prevent application. When a new person with features of a significant other occupies a contradictory role, inferences may still derive from the significant-other representation (Baum & Andersen, 1999).

INHIBITION BASED ON INCOMPATIBLE GOALS

Inhibition can also occur in goal application, often assessed through task persistence. When a goal that is incompatible with an active goal is made accessible, some inhibition of the primary goal is observed. For example, subliminally priming an incompatible goal undermines performance and persistence on the focal goal, and this is moderated by the degree to which people perceive the goals as incompatible (Shah & Kruglanski, 2002). Perceptions of compatibility promote facilitation over inhibition. People effectively pursue an active focal goal by inhibiting incompatible concepts and goals (Moskowitz et al., 2004; Shah, Friedman, & Kruglanski, 2002). An activated stereotype is less likely to be applied when the focal goal to be egalitarian is also activated. Men with chronic egalitarian goals thus do not show facilitated lexical decision judgments to stereotypically female words (vs. control words) when primed with female faces, while men without this accessible focal goal do (Moskowitz, Gollwitzer, et al., 1999a).

People also shield goals from incompatible thoughts, even though adopting the goal of avoiding something (such as a temptation) through conscious suppression can lead to hyperactivation of to-be-avoided associations rather than to their inhibition (Wegner, 1994). That is, if the goal to suppress a particular response becomes associated with the to-be-suppressed item, the presence of the temptation can automatically trigger the focal goal of avoiding it (Shah et al., 2002). When goals are incompatible, they will have an inhibitory relation, so that priming one inhibits the other, and the greater commitment one has to the goal, the stronger the inhibition of alternative goals (Shah et al., 2002). Participants primed subliminally with a focal goal (from a list of three self-generated goals) responded more slowly in a task in which they were to decide whether or not a target word was an attribute of a nonfocal goal, with more inhibition of the alternative goal shown by those who were more committed to the goal (Shah et al., 2002).

Misattribution Based on Ease of Processing

In some instances, the ease with which information is processed also guides application. Some of the best evidence for the power of automatic stereotypes is the fact that stereotypes can work against attempts to be accurate and to not use stereotypes, as reflected in stereotype-consistent memory errors. Prior exposure to a name in-

creases its perceptual fluency, leading to misattributions of fame (Jacoby et al., 1989). Given the greater association of fame with males than with females, this "becoming-famous-overnight" effect occurs more for male names (vs. female names) (Banaji & Greenwald, 1995). In another paradigm, participants presented with several lists of words, including stereotypically male or female roles, showed more false recognition of other male roles in a later memory test—without awareness that the stereotypes influenced memory (Lenton, Blair, & Hastie, 2001; see also Macrae, Schloerscheidt, Bodenhausen, & Milne, 2002). Activated stereotypes also affect memory in that people tend to remember stereotype-consistent information better than inconsistent or neutral information (Fyock & Stangor, 1994). Indeed, practice in making a trait inference from a behavior leads this and other trait inferences to be made more efficiently, and this efficiency is misattributed to a positive evaluation of the behaviors (Smith, 1989). People are in fact more likely to infer that they favor a position when asked to list a few thoughts in support of it (vs. many) because it is easier to do so, and this ease is misattributed to a positive evaluation (Schwarz & Clore, 1988). This misattribution based on ease of processing reflects yet another means by which construct activation can lead to application.

Self-Regulation

Goals Shape Self-Regulatory Processes

Some goals have been suggested as core human motives, such as the goal to understand or comprehend, the goal to get along with others, to belong and affiliate with others, the goal to enhance self-esteem and feel positively about the self, and the goal to be accurate and have valid beliefs (Fiske, 2005; Moskowitz, 2005; Smith & Mackie, 2005; see also Andersen & Chen, 2002; Baumeister & Leary, 1995). These and other goals are central elements of self-regulation. For example, the goal to comprehend and also the goal to avoid bias shape the way stereotypes and traits are activated and applied in both judgment and behavior (for reviews, see Kunda & Spencer, 2003; Sinclair & Kunda, 1999; Spencer et al., 1998). These goals may contribute to the inhibition of knowledge activation or disrupt its application (e.g., Moskowitz, Gollwitzer, et al., 1999; Moskowitz et al., 2004). Emergent evidence suggests that self-regulatory processes can take place relatively unconsciously and effortlessly, and this tell us something about the complexity of "executive" operations.

As noted before, goals define end states and are associated with behaviors geared toward attaining them. Goal application can thus be used as a measure of self-regulation. For example, the degree to which one can complete a goal-relevant task while putting effort into a distracter task standing in the way of the goal reflects the automatic application of the goal in behavior (Aarts et al., 2004). Self-regulation occurs after goal priming as the incentive value of the goal becomes greater (Förster, 2005) and approach behavior ensues (Bargh et al., 2001; Förster, Higgins, & Idson, 1998).

THE GOAL OF COMPREHENSION

A fundamental motive in human psychological functioning is to seek comprehension, a degree of certainty in one's beliefs—that is, clear meaning and understanding (Heider, 1944). Although people do not always have an explicit intention (a conscious goal) to infer meaning, they make meaning all the time. Seeking comprehension is a distal goal that is fulfilled proximally with actions, such as making trait inferences about people, situations, and social objects (Uleman, Newman, et al., 1996). The automatic operation of motives or goals in trait inference involves self-regulation of trait activation and use. Once a trait is inferred based on an observed behavior in another person, the observation is meaningful.

The literatures on both attribution and accessibility have addressed such questions for decades. Our focus here is the automatic components of regulatory processes. For example, spontaneous trait inference, an automatic process of trait activation and application from observing another person's behavior, is moderated by a chronic need for structure. People with this activated goal show more spontaneous trait inferences, despite being neither aware of having the goal nor aware of forming the inference (Moskowitz, 1993). It is also moderated by chronic authoritarian goals (Uleman, Winborne, Winter, & Shechter, 1986).

Comprehension goals (the desire to understand) can inhibit stereotyping when a person feels that stereotyping would impede understanding. Priming these goals can thus lead people to be more attentive and careful in their judgments (Leyens, Yzerbyt, & Schadron, 1992; Macrae et al., 1995; Yzerbyt, Schadron, Leyens, & Rocher, 1994). On the other hand, comprehension goals can instead exacerbate stereotyping when it seems to the person that stereotyping might facilitate understanding (Kunda & Spencer, 2003; Pendry & Macrae, 1996; Stapel & Koomen, 2001a).

THE GOAL OF BEING ACCURATE AND AVOIDING BIAS

Research on the goal to avoid bias is most clearly seen in studies of stereotype application. Old assumptions about the nature of automatic stereotyping suggest that one cannot alter a stereotype's automatic activation, which led to early research on how one can regulate outputs from already activated stereotypes (Devine, 1989; Fazio, 1990; Fiske & Neuberg, 1990). Self-regulation can help people to avoid the effects of automatic stereotypes by suppressing biased impulses, attending to alternative sources of information (e.g., focusing on egalitarian responses), and attempting to correct for biased actions (see Devine & Monteith, 1999). These are output-control strategies and they have preconditions, such as awareness of the potential influence of the automatic stereotype, the knowledge that one was influenced and by how much, the motivation to avoid the influence, and the capability to implement an effective strategy to change the output (Moskowitz, Skurnik, & Galinsky, 1999; Wegener & Petty, 1997; Wilson & Brekke, 1994). This is asking a

lot, which explains one reason why stereotyping, prejudice, and discrimination are hard to eliminate, even among committed egalitarians.

Conscious attempts to suppress stereotypes can have unintended consequences, such as leading people who successfully suppress a stereotype to have a more accessible (automatic) stereotype later (Macrae, Bodenhausen, et al., 1994; see also Galinsky & Moskowitz, in press; Wegner & Erber, 1992). This backfiring of suppression occurs only for people highly motivated to avoid prejudice, showing that this paradoxical effect is moderated by motivation. In sum, automatic stereotyping involves failure to control biases that stereotypes introduce, despite one's best intentions (Monteith, 1993). While control of stereotype application can sometimes be demanding, conscious, and effortful, it can also occur automatically. Stereotypes may be automatically evoked, but motivation to avoid them may be as well, introducing the potential for automatic self-regulation.

Chronic egalitarians have a persistent goal to respond without stereotyping and prejudice (e.g., Plant & Devine, 1998). They readily detect information relevant to the goal of being egalitarian and inhibit incompatible responses (Moskowitz et al., 2004). Repeated attempts to suppress a stereotype are a form of practice that can lead to automatic inhibition with exposure to relevant stimuli (Bargh, 1990). Ultimately, the cues that activate the stereotype also activate the regulatory response by means of "spreading inhibition." Focusing attention on ignoring a stimulus (and its semantic associations) can lead to inhibition (Fox, 1995), termed "negative priming" (Tipper, 1985). In negative priming, spreading inhibition leads responses to "slow down" (vs. facilitation) when the relevant words are ones that are supposed to be ignored (Neill, Valdes, & Terry, 1995). Moreover, men with chronic egalitarian goals presented with pictures of women inhibit stereotypical thoughts automatically, whereas men with no such goals have more accessible stereotypical thoughts (Moskowitz, Gollwitzer, et al., 1999a). Chronic egalitarians also pronounce stereotypical words more slowly in response to female versus male pictures that were to be ignored, and thus seem to ignore stereotypical associations habitually. A female face triggers the goal and self-regulation, which then wards off the incompatible response, and this appears automatized. This and other research on racial attitudes and stereotypes shows that people with chronic goals to avoid bias develop automatic responses to stereotype-relevant stimuli that facilitate the avoidance of bias, such as the inhibition of stereotypes and the activation of goals, beliefs, and actions that oppose stereotyping (Amodio et al., 2004; Devine et al., 2001; Glaser, 2005; Moskowitz, Gollwitzer, et al., 1999; Moskowitz et al., 2000). This likely tells us something about the nature of control processes writ large. Goal activation can inhibit other relevant knowledge and goal activation and application and can do so automatically when rehearsed in advance.

A goal (such as to avoid prejudice) need not be chronic in order for implicit control (such as inhibition of stereotypes) to be produced. Indeed, even people without

chronic tendencies to be egalitarian can inhibit stereotype activation if sufficiently committed to the goal. This goal can be produced in the moment by giving people implementation intentions, say, to be egalitarian (to formulate specific plans regarding when and how they will be nonbiased) or by having people contemplate an instance in which they were prejudiced (failed to be egalitarian). Failure to attain the goal of being nonbiased is thought to lead to a state of "incompleteness" that accompanies failure in committed goal pursuit. This state triggers compensatory responses, including automatic processes that prepare the individual for goal-relevant behavior (Moskowitz, 2001, 2002). These compensatory processes include control over stereotype activation, inhibition of the stereotype (Moskowitz et al., 2004), and readiness to respond to goal-relevant cues.

One example of such readiness is the automatic process of selective attention whereby stimuli relevant to the goal of egalitarianism are attended to more quickly (relative to a condition involving no such incompleteness). This type of automatic attention effect has been shown in a computer task in which words relevant to egalitarianism are presented as distracters vis-à-vis a focal task (Moskowitz, 2002). When people are asked, as a focal task, to press a button indicating the direction that stimuli have moved on a computer screen (which move so quickly that they cannot be detected as words nor can their relevance be determined), their reaction times are slower if these distracter words are in fact relevant to egalitarianism. Only "incomplete" participants (those failing to reach the goal of being accurate and nonbiased) show this kind of distracted attention, thus illustrating attentional selectivity as a compensatory response to the goal state.

Beyond the control that individuals themselves exert over automatic stereotypes (or their outcomes), *social* regulation may also influence stereotypes and prejudice. Some people likely suppress stereotypes not because they do not believe in them but, rather, because of pressure from societal norms about what is considered appropriate (on internal and external motives, see Plant & Devine, 1998). Automatic stereotypes and prejudice are responsive both to motivation and to social context. Changes in context, whether instituted by the individuals, by others, or by settings, can alter both processes and outcomes (Blair, 2002). Exposure to minority leaders, and also to female scientists, can reduce automatic gender stereotypes and prejudice (Dasgupta & Asgari, 2004; Dasgupta & Greenwald, 2001; Richeson & Ambady, 2003). Increases in the numbers and visibility of such individuals may thus be able to reduce stereotyping and prejudice over time in society. Encouraging individuals to confront prejudices (e.g., diversity seminars) also appears to be associated with decreases in automatic stereotyping and prejudice (Rudman, Ashmore, & Gary, 2001), altering students' automatic associations and diminishing negative associations toward African Americans.

Indeed, simply making egalitarian norms salient can also impact automatic stereotypes (Sechrist & Stangor, 2001; Sinclair, Lowery, Hardin, Colangelo, 2005; Spencer, Fein, Zanna, & Olson, 2003). Lower levels of automatic race stereotypes arise when people are led to believe their stereotypes are "out of step" with those of their peers (Sechrist & Stangor, 2001). The expression of automatic stereotypes can even be completely eliminated (or reversed) when people are told that relative to their peers they score high on "contemporary racism" and low on "racial tolerance" as compared with scoring that way on nonracial dimensions (Spencer et al., 2003).

Automatic processes are thus more malleable than previously believed, and there are multiple influences on stereotypes that can succeed in changing their expression. For example, inducing people to focus attention on counterstereotypical associations can combat automatic stereotypes (Blair & Banaji, 1996). People led to expect counterstereotypical target associations (in stimuli to be encountered), rather than to expect stereotypical associations, produce significantly weaker automatic gender stereotypes. Counterstereotypical associations can also be formed through mental imagery (Blair et al., 2001). Spending 5 minutes creating a mental image of a strong (counterstereotypical) woman before completing a measure of automatic gender stereotypes leads to weaker automatic stereotypes (relative to creating a neutral image or no image or other controls, such as an image of a weak woman or a strong man, or while simply trying to suppress stereotypes). Counterstereotypical mental imagery also moderates automatic stereotyping on the IAT as well as on word detection sensitivity (d') on the go/no-go association task (Nosek & Banaji, 2001), and indeed on false alarms in a memory task based on a false memory induction.

In short, when people want to avoid bias, they may *try* to recognize when they have acted (or are about to act) in a way that falls short of their personal standards, in part because failing to do so would stimulate negative affect (guilt and disappointment; Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, 1993). Hence, cues for stereotyping come to be associated with aversive feelings anticipated based on regulatory failures (e.g., Monteith & Voils, 1998; Monteith, Voils, & Ashburn-Nardo, 2001; Monteith et al., 2002; Moskowitz, 2002; Moskowitz et al., 2004), such that cues of stereotyping are responded to as if they are *cues for control*, and they then provoke automatic self-regulation. This is unlikely to be something simply about stereotyping per se but, rather, is likely to tell us something about the nature of regulatory processes generally.

THE GOAL OF SELF-ENHANCING

People possess automatic positive self-views (Greenwald & Banaji; 1995; Hetts & Pelham, 2001), as shown in their preference for the numbers in their own birthday (Kitayama & Karasawa, 1997), for people who share their name or birthday (Finch & Cialdini, 1989; Jones, Pelham, Carvallo, & Mirenberg, 2004; Miller, Downs, & Prentice, 1998), and for the letters in their own name (DeHart, Pelham, & Murray, 2004; Koole, Dijksterhuis, & van

Knippenberg, 2001; Nuttin, 1985; Pelham, Mirenberg, & Jones, 2002).

One route to evoking self-enhancement goals is to threaten the self. Triggering negative stereotypes relevant to the self can thus provoke heightened ingroup favoritism and outgroup denigration, as a response to the threat to self-enhancement goals (Fein & Spencer, 1997; Sinclair & Kunda, 2000; Spencer et al., 1998). That is, people implicitly seek to restore feelings of self-worth. Automatic self-protective biases in response to threat are evoked in stereotyping because the threat provokes outgroup derogation in a compensatory fashion (Spencer et al., 1998). However, people who are motivated to see another person positively will also tend be more inclined to see the person as similar to the self (Sinclair & Kunda, 1999).

Self-enhancement emerges in compensatory responses evoked by failures or other threats to the self (Greenberg & Pyszczynski, 1985; Steele, 1988). For example, activation of a negative significant-other representation brings an influx of negative self-qualities into the working self-concept that activates a flood of self-enhancing qualities—that react against the negative relational self (Hinkley & Andersen, 1996; see Reznik & Andersen, 2004, for a related effect). Assessed by means of freely listed self-concept features, this effect is undergirded by automatic activation of the significant-other representation (Andersen et al., 2005).

Automatic self-regulation vis-à-vis significant others depends on the quality of the relationship (Chen et al., in press). People may respond to threat by automatically activating significant-other representations—an attachment figure—and may do so without awareness and in a compensatory fashion (Mikulincer, Gillath, & Shaver, 2002). When threat cues are presented subliminally, they may provoke thoughts of attachment figures, perhaps because thoughts of these others are comforting among securely attached individuals. People high in attachment anxiety are also prone to activation of attachment figures under threat, while those high in attachment avoidance show diminished accessibility (Mikulincer et al., 2002). Indeed, individuals who are anxiously attached respond to threat cues by pursuing closeness, and accentuating agreement with the other, while those who are avoidantly attached pursue self-reliance and emotional distance (Mikulincer, 1998; Mikulincer, Orbach, & Iavneli, 1998). To the extent that people are "good self-regulators" (i.e., good at delay of gratification), they also tend to have fewer negative interpersonal experiences, even when highly sensitive to rejection, and this may allow such people to function even more effectively than those low in rejection sensitivity (Ayduk et al., 2000). Such a process may begin effortfully (see Yovetich & Rusbult, 1994) but can presumably become automatic with practice.

Parents who perceive they have little power relative to their children respond to their child's behavior as if it is a threat—showing the activation of dominance ideation when under a cognitive load (Bugental, Lyon, Krantz, & Cortez, 1997). With sufficient cognitive capacity, such parents effortfully regulate their responses and can respond adaptively, but when distracted, they cannot, suggesting that their dominance ideation is automatic and their self-regulation is not.

THE GOAL OF GETTING ALONG

The goal of getting along with others, to make a good impression, and to fit in moderates stereotyping (Lowery, Hardin, & Sinclair, 2001; Richeson & Ambady, 2003; Sechrist & Stangor, 2001; Sinclair et al., 2005; see also Chen, Shechter, & Chaiken, 1996). People who believe they will evaluate a Black target versus interact with a Black partner of equal status show more automatic prejudice (Richeson & Ambady, 2003). Moreover, those believing they are in a subordinate role (vs. a superordinate one) produce the lowest levels of automatic prejudice. The other's status (i.e., authority role) outweighs race in the prejudice experienced. White individuals also show less automatic prejudice in the presence of a Black individual (rather than White) when this person is the experimenter (Lowery et al., 2001). This appears to arise due to the motivation to get along with and "tune in" to the experimenter's presumed beliefs. Social tuning can also reduce the expression of automatic prejudice even when an experimenter is White, is liked, and is presumed to hold egalitarian attitudes (Sinclair et al., 2005). Hence, social roles and perceptions of socially significant others' beliefs can also moderate automatic prejudice.

Protection of the other can also be evoked as self-regulation in the context of significant-other relationships. A positive relationship will lead people to treat negative aspects of their partner as virtues, with these flaws being seen as endearing, and this may contribute to relationship maintenance (Murray & Holmes, 1993). Negative cues about a new person who resembles a positive significant other (in transference) evoke a compensatory process in immediate facial affect, which becomes much more positive (but not in a control condition) (Andersen et al., 1996). In such a positive transference, more positive affect is expressed in momentary facial expressions in response to *negative features* than to comparable positive features "about" this new person, but this does not occur in the absence of transference, nor does it occur in negative transference. This highly positive response to negative cues also arises when the significant other is a parent, even if the parent was physically abusive (Berenson & Andersen, 2004), presumably protecting the valued relationship. More generally, people also show the same kind of name–letter effect for a significant other as they show for the self (DeHart et al., 2004). That is, people with high self-esteem show the effect both for romantic partners and for best friends, even when the relationship is troubled, while those low in self-esteem show it only when the relationship is going well.

Activation of a significant-other representation also indirectly activates goals that the significant other holds (Moretti & Higgins, 1999; Shah, 2003a), and these have a self-regulatory function (Baldwin et al., 1990). The goals *they* hold influence the individual's behavior (Andersen & Chen, 2002; Baldwin & Holmes, 1987; Shah, 2003a, 2003b), making him or her more inclined to achieve or to be helpful as the other prefers (Fitzsimons & Bargh,

2003; Shah, 2003a, 2003b). Evidence for such automatic self-regulation based on significant-other representations comes largely from priming studies presenting the name of a significant other outside conscious awareness (Shah, 2003a), which results in reports of greater commitment to the significant other's goals and more likeliness to behave in ways consistent with these goals.

Contrast Effects and Self-Regulation

DUAL-PROCESS MODELS

Most social psychological theory relevant to automatic processes postulates a distinction between automatic and controlled processes in which the latter modifies, inhibits, replaces, or enhances the former (Devine, 1989; Fazio, 1990; Fiske & Neuberg, 1990; Gilbert, 1999; Greenwald & Banaji, 1995; Smith & DeCoster, 1999; Strack & Deutsch, 2004; Thompson et al., 1994; Wegner & Bargh, 1998; Wilson et al., 2000).

Dual-process models (Chaiken & Trope 1999) assume, for example, that when a trait is automatically primed while one is inferring a trait from a behavior, correction may occur in an effortful manner. Correction was long considered essential in contrast effects (Martin, 1986), a process in which one adjusts an initial inference or judgment to satisfy a motive (e.g., to be unbiased). Correction models tend to assume that correction will not occur unless one is aware of a potentially biasing information source and has a theory about the direction of the bias, is also motivated to correct for it, and has the cognitive capacity to do so (Moskowitz, Skurnik, & Galinsky, 1999; Wegener, Dunn, & Tokusato, 2001). The response is then adjusted in a direction *opposite to* the perceived bias and in a magnitude commensurate with the bias's perceived magnitude. Contrast emerges from these regulatory steps when efforts to correct overshoot the mark—resulting in overcorrection (Martin, 1986; Martin et al., 1990; Moskowitz & Skurnik, 1999; Schwarz & Bless, 1992).

Using blatant priming, participants are more likely to see the connection between primes and the target judgments and this leads contrast effects to be shown (Lombardi, Higgins, & Bargh, 1987; Martin, 1986; Martin et al., 1990; Moskowitz & Roman, 1992; Moskowitz & Skurnik, 1999; Newman & Uleman, 1990; Strack, Schwarz, Bless, Kübler, & Wänke, 1993). Regulation in this process often involves subtracting the estimated bias contributed by the detected prime from judgments, which requires awareness of the prime. In such cases, initiation of correction is thus likely to be deliberate, even though the process may involve some automatic components. Often, when a contrast effect emerges in judgment following a blatant prime, people fail to report that the eliciting prime had any impact on the corrective judgment (e.g., Moskowitz & Skurnik, 1999), or that they had any explicit intention to correct (e.g., Wegener & Petty, 1995).

Deliberative corrective strategies depend on naive theories about what was influential and how it should be corrected. Even so, such theories can influence judgment without awareness (Anderson & Lindsay, 1998; Wegener, Petty, & Dunn, 1998). When people lack awareness that they implicitly used a theory in correction, they still show correction effects, such as by judging attractive vacation spots like the Bahamas against a high, abstract standard and Kansas City with a lower standard unless these places are compared directly (Wegener & Petty, 1995). When people are explicitly instructed not to be biased, they will rate Kansas City as much more desirable (compared with control participants), showing a different kind of correction and overcorrection with awareness (Wegener & Petty, 1995).

The latter is metacognitive in that it involves "thinking about thinking"—and is thus not fully implicit but involves judging whether one's own reaction is to be trusted, how one might adjust it, and how ongoing cognitive activity might have been influenced (Moskowitz, 2005; Skurnik, Schwarz, & Winkielman, 2000; Wegener et al., 1998). Such conscious processes are essential to our understanding of self-regulation and are not necessarily automatic. But although self-regulation is dictated by theories that implicitly influence judgments (and perhaps some consciousness), implicit processes are still integral. Indeed, explicit attitudes are intriguingly different from automatic ones in that people can alter explicit attitudes about groups when the attitude's basis is shown to be faulty but automatic prejudices remain largely unaffected (Gregg, Seibt, & Banaji, in press). Distinguishing automatic from nonautomatic self-regulation is thus critical.

STANDARD OF COMPARISON MODEL

Early work on contrast effects arose in the priming literature and showed that primed social knowledge that is moderate (i.e., not extreme) tends to produce judgments that assimilate a new stimulus to the prime, whereas primed categories that are extreme produce contrast (Herr, 1986). A standard-of-comparison model accounts for this by arguing that the triggered (primed) construct is used as a standard against which subsequent information is implicitly compared (Stapel & Koomen, 2001b). Ambiguous/vague behavior that is relevant to the prime and subsequently encountered is then assimilated to the primed construct given that the construct is moderate (Herr, 1986; Herr, Sherman, & Fazio, 1983; see Stapel, Koomen, & van der Pligt, 1996). The behavior is implicitly judged to be similar enough to the activated knowledge to fall within the range of acceptance. An extreme prime, by contrast, is dissimilar to whatever is judged next and thus this information/behavior is contrasted away from the primed construct to which it is implicitly compared. In this way, a contrast effect can arise as a form of automatic self-regulation as people pursue the goal of understanding new behavior. The comparison process is initiated and carried out without need of conscious monitoring or control (Moskowitz & Skurnik, 1999; Stapel & Koomen., 2001b), and even under cognitive load. As people pursue the distal goal of making meaning, using social comparison, and activated knowledge, they self-regulate based on standards or reference

points. In short, they appear to assume that moderate constructs (e.g., Bart Simpson), when primed, share features with a to-be-judged stimulus while extreme constructs (e.g., Hitler) do not. Donald may be hostile, but he is unlike Hitler.

INTERPRETATION VERSUS COMPARISON GOALS

The use of exemplars (Hitler and Bart Simpson) in describing the standard-of-comparison model is no accident. The standard-of-comparison model and correction models make opposite predictions about when contrast will occur. The former claims it occurs following extreme primes, while the latter claims it occurs following moderate primes (Moskowitz & Skurnik, 1999). However, research on the former tends to prime exemplars (nouns, often "proper" nouns) while the latter primes traits (adjectives). The two models describe different basic processes that produce contrast, each with a separate role for automaticity. Exemplar primes are used as standards of comparison that produce contrast through an automatic process of *comparison*, whereas trait primes serve as interpretive frames that produce contrast through conscious correction for perceived bias in judgment (with lack of awareness of the assumptions guiding the correction and perhaps of the fact (or degree) of the correction.

Given these distinct regulation processes, goal activation may play a central role in determining how activated constructs are used (Stapel & Koomen, 2001a). The goal to compare is likely to trigger standard-of-comparison-type processing (even when a trait is primed), while the goal to interpret is likely to trigger use of an interpretive frame (even though implicit, such as when a feeling of fluency is interpreted as being caused by a quality of the target). Research manipulating interpretive (vs. comparative) goals by means of instructions, followed by the activation of a relevant trait concept, results in different effects as a function of the goal (Stapel & Koomen, 2001a). With interpretation goals, activated concepts are used as interpretive frames resulting in assimilation. With comparison goals, contrast occurs. The primed concept thus serves as a standard. Interpretation and comparison are only a small subset of the goals people hold in social situations (e.g., see the discussion of cognitive dissonance below), but they are nonetheless quite important.

In summary, although some ways of conceptualizing self-regulation and self-control invoke explicit, conscious, or intentional processes, regulatory processes are not isomorphic with "conscious" control, because they can occur automatically (Custers, 2006; Fishbach et al., 2003; Förster & Higgins, 2005; Moskowitz, Skurnik, & Galinsky, 1999; Moskowitz et al., 2004; Shah & Kruglanski, 2002). In the attitude domain, even though automatic evaluation effects are facilitated by exposure to evaluatively congruent primes (Bargh et al., 1992; Fazio et al., 1986), this can be reversed when the primes are evaluatively extreme, and even when there is no intention to evaluate (Glaser & Banaji, 1999). One suggestion is that this contrast effect is evidence for regulatory processes that debias the effects of evaluatively extreme

primes, especially when subjects are motivated to respond accurately (Glaser & Banaji, 1999).

Most of these theories have been built on behavioral methods, and neurological investigations are also beginning to provide some converging evidence for the automatic- and controlled-processes distinction in evaluation. For example, a recent study presented Caucasian subjects with Black and White faces subliminally or supraliminally during event-related fMRI (Cunningham et al., 2004). For subliminal primes, the amygdala—a brain region associated with emotional processing of threat—was more active for Black versus White faces (see also Amodio et al., 2003). This supports prior research suggesting that automatic evaluation occurs and can do so without awareness of the eliciting stimulus. Furthermore, this difference in amygdala activation to subliminally primed faces was positively correlated with automatic evaluation on the IAT, $r (11) = .79$, $p < .01$. By comparison, when the faces were presented within conscious awareness (supraliminally), the difference in amygdala activation to Black versus White faces was significantly reduced, implying, perhaps, that other processes evoked by supraliminal presentation intervened in amygdala activation. Moreover, the difference in amygdala activation between the subliminal and supraliminal priming conditions was in fact correlated with activation of the dorsolateral PFC and the anterior cingulate, two areas of the brain associated with executive function and regulation (MacDonald, Cohen, Stenger, & Carter, 2000; Ochsner, Bunge, Gross, & Gabrieli, 2002). The latter finding thus lends some neurological support to the notion that automatic evaluation can be inhibited or suppressed by the introduction of controlled processing.

This hypothesis has also been tested directly by measuring ERPs among participants with strong automatic race-biased associations doing a task requiring the inhibition of racial biases (Amodio et al., 2004, in press). Higher levels of ACC activity, as indicated by the error-related negativity ERP component, predicted more controlled (i.e., less race-based) behavioral responses. Similarly, when individuals showing automatic racial bias engage in interracial interactions (Whites with Blacks), they self-regulate—and thus are slower afterward on a relevant Stroop color-naming task, suggesting resource depletion from that regulation (Richeson & Shelton, 2003). More to the point, those most impaired on the Stroop (who are also more implicitly biased) show more prefrontal activity (right dorsolateral PFC), as assessed by fMRI, when presented with Black faces, suggesting that such stimuli really evoke self-regulation, presumably due to social norms not to appear biased (Richeson et al., 2003).

Cognitive Dissonance and Self-Regulation

How do standards and expectancies salient at the time influence self-regulation? Cognitive dissonance is considered to be a negative drive state aroused when an individual perceives that he or she has more dissonant than

consonant cognitions about a particular entity. This aversive state motivates arousal reduction (Festinger, 1957), which can be accomplished through compensatory action (bringing behavior in line with cognition) or compensatory cognition (changing one's cognitions). The compensatory act alleviates the negative motivational state (Stone & Cooper, 2001). However, cognitions seen as dissonant in the context of one goal or standard may not seem dissonant in the context of another. Different types of behavioral discrepancies can trigger distinct motivational states within the rubric of dissonance (Stone & Cooper, 2001). Under some conditions, behavioral discrepancies trigger consistency-restoring responses—in upholding important attitudes (e.g., Harmon-Jones, Brehm, Greenberg, Simon, & Nelson, 1996). In other circumstances, dissonance motivates the reduction of aversive consequences of unwanted behavior, or the desire to affirm positive self-aspects, even those unrelated to the inconsistencies (e.g., Steele & Liu, 1983).

Evidence for automaticity in dissonance has emerged in that the various motivations that create dissonance—needs for psychological consistency, self-consistency, self-affirmation, reduction of aversive consequences—can be primed (Stone & Cooper, 2001). As an example, people high in self-esteem (vs. low) show more attitude change when implicitly primed with their *personal standards*. That is, they show more postdecision justification of a discrepant act. This dissonance reduction is evidence that dissonance motivation had been primed. By contrast, when implicitly primed with *normative standards* (e.g., "act as I should"), they show about the same level as do people low in self-esteem (Stone, 2001). Implicit priming in dissonance is also seen in priming dissonance reduction strategies as well as in priming arousal of dissonance motives. For example, dissonance reduction through use of self-resources for self-affirmation follows implicit priming of positive self-traits (e.g., I am creative) with behavior justification and counterattitudinal advocacy as the dependent variables (Stone & Cooper, 2003).

Dissonant cognitions can create tension, as can dissonance between a desired goal state and an actual goal state. Failure to attain a goal is thought to lead to a state of "incompleteness" that triggers compensatory responses, including automatic processes that prepare the individual for goal-relevant behavior. One such response to an incomplete goal is the automatic process of attentional selectivity to goal-relevant stimuli induced by failure in goal pursuit (Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). Having done poorly on a putative intelligence test involving unsolvable problems and thus failing at the self-image goal of being intelligent, participants first allowed to affirm an aspect of their self-image that was irrelevant to intelligence should presumably experience a buffer against the failure feedback. Indeed, a lexical decision task assessing implicit processes of facilitated attention to goal-relevant words showed that people who had not self-affirmed were faster to recognize words related to intelligence than those who had (or who received no failure feedback). Self-affirmation blocked selective attention, suggesting that attentional selectivity

repaired the tension associated with the failed goal pursuit. The mechanism underlying these effects is thought to be tension reduction. Unfulfilled goals evoke tension, which then evokes a drive to reduce. Such a reduction can be achieved by preparing to detect and pursue goal-relevant objects to compensate for the shortcoming by. Once the tension is repaired (as through affirmation), the process is shut down (e.g., Moskowitz et al., 2004).

Regulatory Focus and Self-Regulation

Once self-regulatory focus is activated, there is an automatic inclination to pursue goals, solve problems, and make decisions with this specific and distinct strategy. Regulation through approaching desired (ideal) states is *promotion focus*, while regulation through avoiding undesired states (which would result from behaviors one ought not to do) is *prevention focus* (Higgins, Roney, Crowe, & Hymes, 1994). The regulatory focus that becomes accessible during goal pursuit affects how goals are pursued. The same goal can be pursued through prevention or promotion, altering cognitions and behaviors (Förster et al., 1998; Förster, Higgins, & Strack, 2000; Friedman & Förster, 2001; Grant & Higgins, 2003; Higgins & Spiegel, 2004; Idson, Chen, Liberman, & Higgins, 2000; Shah & Higgins, 2001; Shah, Higgins, & Friedman, 1998).

A regulatory focus can be implicitly primed by leading people to think about achieving ideals or about preventing negative outcomes, while unaware of the influence that this framing has on their performance. People who are told they can earn extra money by performing at a certain level (inducing promotion focus) or can avoid losing money by such performance (inducing prevention) show regulatory focus effects (Förster et al., 1998). Asking people to solve the puzzle of finding a way for a rat to get out of a maze can also prime regulatory focus. The task is framed as a helping task implicitly when the route is to a piece of cheese versus the route to escape a hovering predator (Friedman & Förster, 2001). Behavioral measures of implicit promotion and prevention focus have also been associated with approach- versus avoidance-related patterns of neural activity, respectively (Amodio et al., 2004).

In this process, promotion focus induces a cognitive style characterized by "strategic eagerness" and prevention focus induces an inclination toward "strategic vigilance." People primed with a promotion focus generate more alternative hypotheses than those primed with prevention (Liberman, Molden, Idson, & Higgins, 2001)—when viewing ambiguous objects and asked to list as many identities for the objects as possible. Under promotion (vs. prevention), they are also faster in generating each hypothesis. After reading a scenario about someone performing helpful acts that could be situational or dispositional in cause, those in promotion focus show a positive association between these explanations, whereas those in prevention focus show a negative association, suggesting that promotion-primed participants are more open to alternative hypotheses about the causes of behavior.

The evidence reveals enhanced memory for stereotype-inconsistent information in promotion focus, especially when people are prejudiced because stereotype-inconsistent behavior then represents an even greater mismatch from expectations (Förster et al., 2000). Promotion focus should also be a riskier cognitive style and should lead to particularly speedy responses, whereas prevention focus should foster caution and accuracy. In promotion focus, people are in fact faster, especially when goal attainment is imminent, while in prevention focus there is more accuracy than speed (Förster, Higgins, & Bianco, 2003). Regulatory focus is also evoked when a new person with features of one's parent activates this parental representation and the parent's standards, which then guides self-regulation (Andersen & Chen, 2002; see also Shah, 2003b). The general point is that regulatory focus, when evoked, changes how contextual cues are processed and pursued.

In short, the influence of promotion and prevention orientations is automatic, as is the influence of the strategic inclinations with which they are linked, and each can activate the other. The orientations are broad and encompass many types of goals, including new, situated goals, and the inclinations general in that they can be realized with a variety of means, including new, situated means. The model thus extends beyond those targeting goal–means associations already established in memory (e.g., Bargh, 1997; Gollwitzer, 1999; Kruglanski et al., 2002), and once activated, these orientations implicitly influence processing.

Availability: Acquisition of Automaticity

Mechanisms of Acquisition in Automaticity

PRACTICE

The question of *how* automaticity is acquired is addressed by evidence showing that repetition of trait inferences leads to their automatization. For example, practice in inferring intelligence based on behaviors leads people to make such judgments more quickly and efficiently, and also to rate intelligent but unfriendly behaviors more positively (relative to practice in inferring friendliness) (Smith, 1989). This suggests that procedural efficiency may be misattributed to desirability (Smith, 1989; see also Monahan, Murphy, & Zajonc, 2000; Winkielman & Cacioppo, 2001). Research on spontaneous trait inferences has also shown that practice influences acquisition (Bassili, 1993; see Winter & Uleman, 1984).

In the literature on automatic evaluation, repetition is also influential. A prominent example is shown in the mere exposure effect (Zajonc, 1968). Repeated exposure to a stimulus object increases its ease of processing on subsequent presentations (Gordon & Holyoak, 1983). Such enhanced fluency is thought to produce a positive feeling of familiarity associated with the object, and the evaluation of novel stimulus objects can be acquired on this basis, even when the repetition is subliminal (Bornstein, 1989; Kunst-Wilson & Zajonc, 1980). Practice thus offers a means of automatic attitude acquisition. Practice in expressing attitudes can also lead to height-

ened attitude accessibility in terms of response latency and to greater attitudinal certainty (Holland, Verplanken, & van Knippenberg, 2003). Significant-other representations are also likely to be acquired in part based on repetition as people regularly interact with significant others, as the classic practice effect of consolidating memory and increasing positive evaluation (Chen & Andersen, 1999).

Acquisition of automatic self-regulation has also been shown based on practice in *not* resorting to the use of stereotypes. This can alter automatic responses. When people repeatedly practice negating stereotypes that they hold (e.g., about skinheads or racial groups), they show reduced stereotype activation. Using a priming-based Stroop task or a person categorization task shows diminished stereotype activation among those who practiced stereotype inhibition versus those who did not. Whereas this speaks more to diminished accessibility than to the altered availability of stereotypes in memory, it suggests that practice in negating stereotypes increases participants' proficiency in doing so, and this may occur by means of restructured associations in memory (i.e., the existence of incompatible associations) (Kawakami et al., 2000).

As with other social knowledge, research on goals suggests that routine pairings of goals with specific environments, and practice at pursuing a goal in a specific context, enable links to be forged in memory between goals and the environments in which they are pursued (Bargh, 1990; Bargh & Gollwitzer, 1994; Bargh et al., 2001). The goal is thus ultimately responsible for the repetition and for the habitual response (Aarts & Dijksterhuis, 2000). "Learned sequences of acts . . . become automatic responses to specific cues, and are functional in obtaining certain goals or end states" (Verplanken & Aarts, 1999, p. 104). Repetition increases efficiency in goal-related cognition until consciousness is unnecessary for goal pursuit (Bargh, 1990). The link between the relevant situation or cue and the goal becomes available in memory and can thus be activated and manifested in behavior (e.g., Aarts & Dijksterhuis, 2000, 2003).

CLASSICAL CONDITIONING

Attitudes can, of course, be classically conditioned as when unconditioned stimuli (e.g., positive or negative facial expression) are presented subliminally as primes (Krosnick, Betz, Jussim, Lynn, & Stephens, 1992) or as simultaneous stimuli (Olson & Fazio, 2001). Positive classical conditioning toward the self can be produced as well by repeatedly pairing smiling faces with self-relevant information in a computer game (Baccus, Baldwin, & Packer, 2004; *www.selfesteemgames.mcgill.ca*).

Acquisition of automatic attitudes can also involve physiological feedback. For example, exposure to novel stimuli while one's arm is flexed—as if pulling something toward the body—results in positive evaluations of those stimuli. Negative evaluations arise after exposure to items encountered when the arm is extended—as if pushing something away from the body (Cacioppo, Priester, & Berntson, 1993). These acts appear to operate as physi-

cal indicators of approach and avoidance and thus serve as information in acquiring evaluations of the stimuli. Moreover, recent research suggests that whether approach-and-avoidance motions lead to positive or negative evaluations is itself contextually situated, with approach enhancing positivity for positive things but not for negative, and vice versa for avoidance (Centerbar & Clore, 2006). Although not challenging the overall effect, this focuses again on context-based variability.

CONSTRUCTING LINKS TO THE SELF

Another mechanism of acquisition for automatic evaluation is based on group membership, such as that found in the minimal group paradigm (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971). That is, liking for one's own group occurs even when the group and one's membership in it are random or fleeting (see also Ashburn-Nardo, Voils, & Monteith, 2001; see also Otten & Moskowitz, 2000; Otten & Wentura, 1999). Participants placed in a rather meaningless group that preferred the art of one painter (Quan) versus another (Xanthie) later showed an automatic preference for their own group compared with the other group, especially among those who were feeling angry (DeSteno, Dasgupta, Bartlett, & Cajdric, 2004). Automatic preferences can also develop even when people are not assigned to be a member of a particular group (Greenwald, Pickrell, & Farnham, 2002). People who are simply exposed to group members tend to like the group and identify with its members immediately and implicitly, as compared with an unknown group. Self-categorizing an object or group thus leads to some modicum of positive evaluation.

Early Acquisition of Concepts in Development

Humans expend considerable energy making sense of their environments, determining what particular social and nonsocial objects are, and ascertaining what qualities and functions those objects may have. Developmental research suggests that children infer stable traits in others by ages 9–10, whereas prior to that (e.g., ages 5–6) tend to rely on more basic evaluation (Alvarez, Ruble, & Bolger, 2001). At ages 9–10, trait concepts also mediate behavioral predictions in that children predict behavior in new situations to be consistent with prior behavior and relevant trait inferences, suggesting that the use of dispositional trait concepts emerges at this later moment in childhood (Rholes & Ruble, 1984). Children ages 9–10 even make spontaneous trait inferences to a greater extent than do adults (college students) (Bassili, 1993), suggesting the tendency to make automatic trait inferences may be acquired developmentally.

Developmental research also shows object-evaluation associations early in childhood, at least by ages 5–6 when they are clearly operative in judgments about others (Alvarez, Ruble, & Bolger, 2001; Ruble & Dweck, 1995). The evaluative component of categorization is well in place by ages 4–6 (Aboud, 1988), although children may not yet know stereotypes of various ethnic groups and may nonetheless show effects in evaluation. This is of in-

terest in part because stereotypes and attitudes can be automatized through direct and indirect socialization (Devine, 1989; Greenwald & Banaji, 1995; Smith & Decoster, 1999). Societal or cultural effects are suggested most clearly in the fact that societally disadvantaged groups show lower levels of automatic ingroup bias (and may even favor outgroups) on implicit bias measures and yet show high levels of explicit ingroup bias (Hummert, Garstka, OBrien, Greenwald, & Mellott, 2002; Lane, Mitchell, & Banaji, 2005; Livingston, 2002; Nosek, Banaji, & Greenwald, 2002a; Rudman, Feinberg, & Fairchild, 2002). Automatic stereotypes may thus be ingrained through socialization, even if the stereotypes contradict endorsed beliefs. Recent research shows that implicit and explicit attitudes have different developmental trajectories. Even young children (6-year-olds) acquire implicit racial attitudes, which remain strong in older children and adults (Baron & Banaji, 2005; Dunham, Baron, & Banaji, 2006), while explicit racial bias is less apparent in older children and often absent in adults (Baron & Banaji, 2006).

Group stereotypes and intergroup evaluations also arise in development as linked with the functional significance of the category (Bigler, 1995). Children 6–11 years of age randomly assigned to a condition in which gender has functional uses in their classroom (for 4 weeks; e.g., sex-segregated seating and bulletin boards) acquire pronounced occupational stereotypes defined by gender. Minimal group membership effects can also be produced at the same age if the classification is functional (Bigler, Jones, & Lobliner, 1997). Randomly assigning children of this age to a "red" or "blue" group used in class seating arrangements (vs. to no category or no functional category) leads to more intergroup bias after 4 weeks, even in the absence of any explicit evaluative differences. In addition, when one of the assigned groups is implicitly depicted as relatively low in status, the ingroup favoritism effect occurs for children in the high-status group but not in the low-status group (Bigler, Brown, & Markell, 2001). Categories with functional uses in one's environment become salient and are used repeatedly, suggesting a mechanism by which social categories become associated with beliefs and come to function automatically.

Given that stereotypes tend to be noun categories and traits attribute categories, it is of interest that noun categories are learned earlier in development than are adjective categories (Gelman & Markman, 1985; Gelman & Raman, 2003). Children as young as 2 label abstract objects with noun phrases and, by age 3, refer to attributes in a stimulus or context, such as an object's color. Certainly, children 3–4 years of age use category membership (noun labels) more than perceptual features (adjectives) in their inferences, suggesting that the noun–adjective distinction arises rather early as well, and far earlier than dispositional trait inferences (age 9; Alvarez et al., 2001). Indeed, the distinction between generic categories and exemplars (a specific object or person) also develops early and is in place by the age of 2 (Gelman & Raman, 2003). Early development of significant-other representations is supported too, unsurprisingly, in re-

search on face recognition showing that infants can readily distinguish familiar and unfamiliar faces in the first several weeks of life (Meltzoff & Moore, 1992; Pascalis, de Schonen, & Morton, 1994).

Significant others are present early in life and such representations are thus acquired early. Beginning with rudimentary caretaker representations acquired in infancy as working models (Thompson, 1998), socialization and development enrich what is represented, for example, among preschoolers (as assessed through mommy or daddy narratives; Oppenheim, Emde, & Warren, 1997; Toth, Cicchetti, Macfie, & Emde, 1997). Additional significant-other representations also form throughout childhood (Rudolph, Hammen, & Burge, 1995; Ryan, Stiller, & Lynche, 1994) and adulthood (Brennan & Morris, 1995; Collins & Read, 1994; Hazan & Shaver, 1994; Mikulincer, 1995; Simpson & Rholes, 1998), and their acquisition plays a profound role in the self-concept (Aron, Paris, & Aron, 1995; Baldwin, 1992; Higgins, 1989). Hence, these representations constitute a constrained class of concepts in how they are represented and in the processes by which they are formed. This makes creating actual (but temporary) significant-other "relationships" in laboratory settings almost impossible or at least a poor analog for how they function in real life, limiting the kinds of research that can be done on acquisition of such representations. Once specific significant-other representations develop, however, they can be generalized to other people (e.g., Andersen et al., 1996). Even in 8- to 12-year-olds, caretaker representations predict adjustment and in fact mediate peer rejection (Shields, Ryan, & Cicchetti, 2001), although it remains to be seen whether or not parental representations literally serve as the basis for subsequent significant-other representations or for generic person prototypes.

In sum, the acquisition of social knowledge and its development in children and adults is what determines availability and heightened accessibility based on context. Existing evidence in social cognition, however, has offered relatively few insights thus far about how learning occurs in associative memory. In behavioral neuroscience, learning can be studied at the level of the individual synapse. When stimulus meets response in reliable fashion, weighting processes occur and an association is learned. The ability for such reorganization of synaptic strengths to occur is what defines plasticity and accounts for learning (e.g., Bailey & Kandel, 1995). When an association between cue and response is automated, it is presumably because a stable change in the pattern of synaptic weights has occurred (i.e., plasticity). Hence, if one wants to know how information is made available (learned), it may be in part to this level of analysis that one must turn. Of course, the study of individual synapses is typically done in nonhuman organisms. At the level of "mind" studied in social psychology, the simple assumption is made that experience allows for learning and enables those formed associations to become more accessible if contextually cued. In neuroscience, long-term, stable changes in neural activity in response to experience are studied in humans and higher-order mammals

through cortical mapping (e.g., Kaas, 1995; Wiesel & Hubel, 1965) and practice (e.g., Pascual-Leone, Amedi, Fregni, & Merabet, 2005). It is our view that more research on such problems is needed in our field that spans multiple levels of analysis—from child development to practice effects as well as to classical conditioning and cortical mapping (and even synaptic mechanisms if ever feasible). More extensive research on how social knowledge becomes available is necessary if automatic thought is to be fully understood.

CLOSING REMARKS

In conclusion, what is known about automatic thought has continued to grow at a steady clip over the last decade. It is ever more clear that vast reaches of what is commonly considered *thought* may not be particularly conscious, effortful, intentional, or controllable. The basic processes by which automatic thought transpires (availability, accessibility, application, and self-regulation) are rather well understood, and yet numerous questions about them remain. When seen through the lens of the wide swathe of research literatures we address, which range from traits, stereotypes, and attitudes, to interpersonal relationships and to motives and goals, there are emergent themes that come to light that we conclude constitute the basic principles of automaticity. In particular, we argue that automaticity and control are not discrete, opposing processes; that contexts and situations, and all the complex cues therein, are crucial to what automatic thought then arises; that self-regulation occurs without need of conscious will; that social knowledge becomes available to the perceiver (is learned) through practice and repetition in conjunction with developmental processes and also synaptic mechanisms in the brain, thus enabling the shift from conscious to automatic thought; and that automatic processes occur concurrently with the recruitment of specific brain regions and functions. We believe these principles add texture and nuance to existing knowledge and prompt us to probe pressing questions on which the research has thus far been silent. Much is known, but much remains to be seen.

NOTES

1. It is worth mentioning that we conceive mental representations simply as collections of related knowledge in memory.
2. We recognize that this is a working list of principles, and like any such summary or list, it is imperfect and in progress as science is. Still, it reflects a reasonable approximation of what is now known in the field.
3. Both spreading activation and connectionist models involve the manner in which relevant cues shift activation.
4. This effect is not reducible to what occurs based on social categories because the effect that occurs for categories is less pronounced (Andersen & Cole, 1990; Chen et al., 1999), nor is it reducible to what is observed for a nonsignificant other (e.g., Andersen et al., 1995).
5. Baseline measures of overlap between the overall self-

concept and the self-with-the-other (and of self-evaluation reflected in these features) are assessed so that they can be controlled when assessing increases in this overlap score in the experiment and in associated self-evaluation.

REFERENCES

Aarts, H., Chartrand, T. L., Custers, R., Danner, U., Dik, G., Jefferis, V., et al. (2005). Social stereotypes automatic goal pursuit. *Social Cognition*, *23*, 465–490.

Aarts, H., & Dijksterhuis, A. P. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, *78*, 53–63.

Aarts, H., & Dijksterhuis, A. P. (2003). The silence of the library: Environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, *84*, 18–28.

Aarts, H., Dijksterhuis, A. P., & De Vries, P. (2001). The psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology*, *92*, 631–642.

Aarts, H., Dijksterhuis, A. P., & Midden, C. (1999). To plan or not to plan? Goal achievement of interrupting the performance of mundane behaviors. *European Journal of Social Psychology*, *29*, 971–979.

Aarts, H., Gollwitzer, P. M., & Hassin, R. R. (2004). Goal contagion: Perceiving is for pursuing. *Journal of Personality and Social Psychology*, *87*, 23–37.

Aboud, F. E. (1988). *Children and prejudice*. New York: Blackwell.

Alvarez, J. M., Ruble, D. N., & Bolger, N. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development*, *72*(5), 1409–1425.

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652–661.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology*, *84*, 738–753.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15*, 88–93.

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (in press). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience*.

Andersen, S. M., & Baum, A. B. (1994). Transference in interpersonal relations: Inferences and affect based on significant-other representations. *Journal of Personality*, *62*(4), 460–497.

Andersen, S. M., & Chen, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review*, *109*, 619–645.

Andersen, S. M., & Cole, S. W. (1990). "Do I know you?": The role of significant others in general social perception. *Journal of Personality and Social Psychology*, *59*, 384–399.

Andersen, S. M., & Glassman, N. S. (1996). Responding to significant others when they are not there: Effects on interpersonal inference, motivation, and affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 3. The interpersonal context* (pp. 272–331). New York: Guilford Press.

Andersen, S. M., Glassman, N. S., Chen, S., & Cole, S. W. (1995). Transference in social perception: The role of chronic accessibility in significant-other representations. *Journal of Personality and Social Psychology*, *69*, 41–57.

Andersen, S. M., & Klatzky, R. L. (1987). Traits and social stereotypes: Levels of categorization in person perception. *Journal of Personality and Social Psychology*, *53*, 235–246.

Andersen, S. M., Klatzky, R. L., & Murray, J. (1990). Traits and social stereotypes: Efficiency differences in social information processing. *Journal of Personality and Social Psychology*, *59*, 192–201.

Andersen, S. M., Reznik, I., & Glassman, N. S. (2005). The unconscious relational self. In R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 421–481). New York: Oxford University Press.

Andersen, S. M., Reznik, I., & Manzella, L. M. (1996). Eliciting transient affect, motivation, and expectancies in transference: Significant-other representations and the self in social relations. *Journal of Personality and Social Psychology*, *71*, 1108–1129.

Anderson, C. A., & Deuser, W. E. (1993). The primacy of control in causal thinking and attributional style: An attributional functionalism perspective. In G. Weary, F. Gleicher, & K. L. Marsh (Eds.), *Control motivation and social cognition* (pp. 94–121). New York: Springer-Verlag.

Anderson, C. A., & Lindsay, J. J. (1998). The development, perseverance, and change of naive theories. *Social Cognition*, *16*, 8–30.

Ansfield, M. E., & Wegner, D. M., (1996). The feeling of doing. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 482–506). New York: Guilford Press.

Arndt, J., Greenberg, J., Pyszczynski, T., & Solomon, S. (1997). Subliminal exposure to death-related stimuli increases defense of the cultural worldview. *Psychological Science*, *8*, 379–385.

Aron, A, Paris, M., & Aron, E. N. (1995). Falling in love: Prospective studies of self-concept change. *Journal of Personality and Social Psychology*, *69*, 1102–1112.

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, *21*(1), 61–87.

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root. *Journal of Personality and Social Psychology*, *81*, 789–799.

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P. K., & Rodriguez, M. (2000). Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity. *Journal of Personality and Social Psychology*, *79*, 776–792.

Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, *15*(7), 498–502.

Bailey, C. H., & Kandel, E. R. (1995). Molecular and structural mechanisms underlying long-term memory. In M. S. Gazzaniga (Ed.-in-Chief), *The cognitive neurosciences* (pp. 19–36). Cambridge, MA: MIT Press.

Baldwin, M. W. (1992). Relational schemas and the processing of information. *Psychological Bulletin*, *112*, 461–484.

Baldwin, M. W., Carrell, S. E., & Lopez, D. F. (1990). Priming relationship schemas: My advisor and the Pope are watching me from the back of my mind. *Journal of Experimental Social Psychology*, *26*, 435–454.

Baldwin, M. W., Fehr, B., Keedian, E., Seidel, M., & Thompson, D. W. (1993). An exploration of the relational schemata underlying attachment styles: Self-report and lexical decision approaches. *Personality and Social Psychology Bulletin*, *19*, 746–754.

Baldwin, M. W., & Holmes, J. G. (1987). Salient private audiences and awareness of the self. *Journal of Personality and Social Psychology*, *52*, 1087–1098.

Baldwin, M. W., Keelan, J. P. R., Fehr, B., Enns, V., & Koh-Rangarajoo, E. (1996). Social-cognitive conceptualization of attachment working models: Availability and accessibility effects. *Journal of Personality and Social Psychology*, *71*, 94–109.

Banaji, M. R., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology*, *68*, 181–198.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, *7*, 136–141.

Banaji, M. R., Hardin, C. D., & Rothman, A. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*(2), 272–281.

Banse, R. (1999). Automatic evaluation of self and others: Affective priming in close relationships. *Journal of Social and Personal Relationships*, *16*(6), 803–821.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, *87*(1), 5–22.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 93–130). New York: Guilford Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169–183). New York: Guilford Press.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 1). Mahwah, NJ: Erlbaum.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford Press.

Bargh, J. A., & Barndollar, K. (1996). Automaticity in action: The unconscious as repository of chronic goals and motives. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 457–481). New York: Guilford Press.

Bargh, J. A., Bond, R. N., Lombardi, W. J., & Tota, M. E. (1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology*, *50*, 869–878.

Bargh, J. A. Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893–912.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditionally automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 185–210.

Bargh, J. A., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. In W. D. Spaulding (Ed.), *Integrative views of motivation, cognition, and emotion* (pp. 71–124). Lincoln: University of Nebraska Press.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trotschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*, 1014–1027.

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically accessible constructs in person x situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology*, *55*, 599–605.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437–449.

Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology*, *22*, 293–311.

Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology*, *49*, 1129–1146.

Baron, A. S., & Banaji, M. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10, and adulthood. *Psychological Science*, *17*, 53–58.

Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, *90*, 272–287.

Bassili, J. N. (1993). Procedural efficiency and the spontaneity of trait inference. *Personality and Social Psychology Bulletin*, *19*, 200–205.

Baum, A., & Andersen, S. M. (1999). Interpersonal roles in transference: Transient mood states under the condition of significant-other activation. *Social Cognition*, *17*, 161–185.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Berenson, K. R., & Andersen, S. M. (2006). Childhood physical and emotional abuse by a parent: Transference effects in adult interpersonal relations. *Personality and Social Psychology Bulletin*, *32*, 1509–1522.

Berk, M. S., & Andersen, S. M. (2000). The impact of past relationships on interpersonal behavior: Behavioral confirmation in the social-cognitive process of transference. *Journal of Personality and Social Psychology*, *79*, 546–562.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: Evaluation versus stereotype activation. *Social Cognition*, *18*(4), 329–353.

Bigler, R. S. (1995). The role classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development*, *66*(4), 1072–1087.

Bigler, R. S., Brown, C. S., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*(4), 1151–1162.

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*(3), 530–543.

Blair, I. V. (2001). Implicit stereotypes and prejudice. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 359–374). Mahwah, NJ: Erlbaum.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*, 242–261.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142–1163.

Blair, I. V, Chapleau, K. M., & Judd, C. M. (2005). The use of Afrocentric features as cues for judgment in the presence of diagnostic information. *European Journal of Social Psychology*, *35*(1), 59–68.

Blair, I. V, Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*(10), 674–679.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology*, *87*(6), 763–778.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology*, *83*(1), 5–25.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*(5), 828–841.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Brandstätter, V., Lengfelder, A., & Gollwitzer, P. M. (2001). Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology, 81*(5), 946–960.

Brennan, K. A., & Morris, K. A. (1995). Attachment styles, self-esteem, and patterns of seeking feedback from romantic partners. *Personality and Social Psychology Bulletin*, *23*, 23–31.

Bruner, J. S. (1957). Going beyond the information given. In H. E. Gruber, K. R. Hammond, & R. Jessor (Eds.), *Contemporary approaches to cognition* (pp. 41–69). Cambridge, MA: Harvard University Press.

Bugental, D. B., Lyon, J. E., Krantz, J., & Cortez, V. (1997). Who's the boss?: Differential accessibility of dominance ideation in parent–child relationships. *Journal of Personality and Social Psychology*, *79*, 1297–1309.

Cacioppo, J. T., & Berntson, G. C. (2001). The affect system and racial prejudice. In J. A. Bargh & D. K. Apsley (Eds.), *Unraveling the complexities of social life: A festschrift in honor of Robert B. Zajonc* (pp. 95–110). Washington, DC: American Psychological Association.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.

Cacioppo, J. T., Priester, J. R., & Bernston, G. G. (1993). Rudimentary determinants of attitudes: II: Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, *65*, 5–17.

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*, 840–856.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Carver, C. S., & Scheier, M. F. (1999). Themes and issues in the self regulation of behavior. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (pp. 1–105). Hillsdale, NJ: Erlbaum.

Centerbar, D. B., & Clore, G. L. (2006). Do approach–avoidance actions create attitudes? *Psychological Science*, *17*, 22–29.

Chaiken, S., & Bargh, J. A. (1993). Occurrence versus moderation of the automatic activation effect. *Journal of Personality and Social Psychology*, *64*, 759–765.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chartrand, T., & Bargh, J. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, *71*, 464–478.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, *33*, 541–560.

Chen, S., & Andersen, S. M. (1999). Relationships from the past in the present: Significant-other representations and transference in interpersonal life. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 31, pp. 123–190). San Diego, CA: Academic Press.

Chen, S., Andersen, S. M., & Hinkley, K. (1999). Triggering transference: Examining the role of applicability and use of significant-other representations in social perception. *Social Cognition*, *17*, 332–365.

Chen, S., Fitzsimons, G. M., & Andersen, S. M, (in press). Automaticity in close relationships. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior*. New York: Psychology Press.

Chen, S., Shechter, D., & Chaiken, S. (1996). Getting at the truth or getting along: Accuracy- vs. impression-motivated heuristic and systematic processing. *Journal of Personality and Social Psychology*, *71*, 262–275.

Collins, N. L., & Read, S. J. (1994). Cognitive representations of attachment: The structure and function of working models. In K. Bartholomew (Ed.), *Attachment processes in adulthood* (pp. 53–90). Bristol, PA: Jessica Kingsley.

Crites, S. L., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorizations: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology*, *68*, 997–1013.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, *85*, 639–649.

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (2004). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, *30*(10), 1332–1346.

Custers, R. (2006). *On the underlying mechanisms of nonconscious goal pursuit*. Unpublished doctoral dissertation, University of Utrecht.

Custers, R., & Aarts, H. (2005). Beyond priming effects: The role of positive affect and discrepancies in implicit processes of motivation and goal pursuit. *European Review of Social Psychology*, *16*, 257–300.

Dasgupta, N., & Asgari, S. (2004). *Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping*. Unpublished manuscript.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, *46*(5), 991–1004.

DeHart, T., Pelham, B., & Murray, S. (2004). Implicit dependency regulation: Self esteem, relationship closeness, and implicit evaluations of close others. *Social Cognition*, *22*(1), 126–146.

De Houwer, J., Baeyens, F., Vansteenwegen, D., & Eelen, P. (2000). Evaluative conditioning in the picture-picture paradigm with random assignment of conditioned stimuli to unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(2), 237–242.

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science*, *15*, 319–324.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 339–360). New York: Guilford Press.

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60*, 817–830.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2001). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*, 835–848.

Dijksterhuis, A. P. (2005). Why we are social animals: The high road to imitation as social clue. In S. Hurley & N. Charter (Eds.), *Perspectives on imitation: From neuroscience to social science: Vol. 2: Imitation, human development, and culture* (pp. 207–220). Cambridge, MA: MIT Press.

Dovidio, J. F., & Fazio, R. H. (1992). New technologies for the direct and indirect assessment of attitudes. In J. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 204–237). New York: Russell Sage.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*(1), 62–68.

Downey, G., & Feldman, S. (1996). Implications of rejection sensitivity for intimate relationships. *Journal of Personality and Social Psychology*, *70*, 1327–1343.

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, *13*, 513–519.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483.

Dunham, Y., Baron, A. S., & Banaji, M. (2006). From American city to Japanese village: Across-cultural investigation of implicit race attitudes. *Child Development*, *77*, 1268–1281.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204–243). New York: Guilford Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75–109.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, *73*, 31–44.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Finch, J. F., & Cialdini, R. B. (1989). Another indirect tactic of (self-)image management: Boosting. *Personality and Social Psychology Bulletin*, *15*(2), 222–232.

Fishbach, A., Friedman, R. S., & Kruglanski, A. W. (2003). Leading us not unto temptation: Momentary allurements elicit overriding goal activation. *Journal of Personality and Social Psychology*, *84*, 296–309.

Fiske, S. T. (2005). Social cognition and the normality of prejudgment. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 36–53). Malden, MA: Blackwell.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation from category-base to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Fitzsimons, G. M., & Bargh, J. A. (2003). Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology*, *84*, 148–164.

Förster, J. (2006, October). *Seven principles of self regulation*. Paper presented at the Society for Experimental Social Psychology conference, Philadelphia.

Förster, J., & Higgins, E. T. (2005). How global versus local perception fits regulatory focus. *Psychological Science*, *16*, 631–636.

Förster, J., Higgins, E. T., & Bianco, A. T. (2003). Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? *Organizational behavior and human decision processes*, *90*, 148–164.

Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the "goal looms larger" effect. *Journal of Personality and Social Psychology*, *75*, 1115–1131.

Förster, J., Higgins, E. T., & Strack, F. (2000). When stereotype disconfirmation is a personal threat: How prejudice and prevention focus moderate incongruency effects. *Social Cognition*, *18*, 178–197.

Fox, D. (1995). Integration of the cognitive and the psychodynamic unconscious: Comment. *American Psychologist*, *50*(9), 798–799.

Friedman, R. S., & Förster, J. (2000). The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, *79*, 477–492.

Friedman, R. S., & Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology*, *81*, 1001–1013.

Fyock, J., & Stangor, C. (1994). The role of memory biases in stereotype maintenance. *British Journal of Social Psychology*, *33*(3), 331–343.

Galinsky, A. D., & Moskowitz, G. B. (in press). Further ironies of suppression: Stereotype and counter-stereotype accessibility. *Journal of Experimental Social Psychology*.

Geen, R. G. (1995). *Human motivation: A social psychological approach*. Pacific Grove, CA: Brooks/Cole.

Gelman, S. A., & Markman, E. M. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. *Journal of Child Language*, *12*, 125–143.

Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development*, *74*, 308–325.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.

Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 3–11). New York: Guilford Press.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*, 509–517.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, *65*, 221–233.

Gillath, O., Mikulincer, M., Fitzsimons, G. M., Shaver, P. R., Schachner, D. A., & Bargh, J. A. (2006). Automatic activation of attachment-related goals. *Personality and Social Psychology Bulletin*, *32*, 1375–1388.

Glaser, J. (2005). Understanding prejudice and discrimination. *Analyses of Social Issues and Public Policy*, *5*(1), 277–279.

Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology*, *77*, 669–687.

Glassman, N. S., & Andersen, S. M. (1999a). Activating transference without consciousness: Using significant-other representations to go beyond what is subliminally given. *Journal of Personality and Social Psychology*, *77*, 1146–1162.

Glassman, N. S., & Andersen, S. M. (1999b). Transference in social cognition: Persistence and exacerbation of significant-other-based inferences over time. *Cognitive Therapy and Research*, *23*, 75–91.

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*, 128–1635.

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. *European Review of Social Psychology*, *4*, 141–185.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Gollwitzer, P. M., & Moskowitz, G. B. (1996). Goal effects on action and cognition. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399). New York: Guilford Press.

Gordon, P. C., & Holyoak, K. J. (1983). Implicit learning and generalization of the "mere exposure" effect. *Journal of Personality and Social Psychology*, *45*(3), 492–500.

Grant, H., & Higgins, E. T. (2003). Optimism, promotion pride, and prevention pride as predictors of well-being. *Personality and Social Psychology Bulletin*, *29*(12), 1521–1532.

Greenberg, J., & Pyszczynski, T. (1985). Compensatory self-

inflation: A response to the threat to self-regard of public failure. *Journal of Personality and Social Psychology*, *49*(1), 273–280.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellot, D. S. (2002). A unified theory of implicit attitudes, beliefs, self-esteem and self-concept. *Psychological Review*, *109*(1), 3–25.

Greenwald, A., Draine, S., & Abrams, R. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*, 1699–1702.

Greenwald, A. G., Pickrell, J. E., & Farnham, S. D. (2002). Implicit partisanship: Taking sides for no reason. *Journal of Personality and Social Psychology*, *83*(2), 367–379.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology*, *90*, 1–20.

Haddock, G., Rothman, A. J., Reber, R., & Schwarz, N. (1999). Forming judgments of attitude certainty, intensity, and importance: The role of subjective experiences. *Personality and Social Psychology Bulletin*, *25*(7), 771–782.

Hafer, C. L. (2000). Do innocent victims threaten the belief in a just world? Evidence from a modified Stroop task. *Journal of Personality and Social Psychology*, *79*, 165–173.

Hamilton, D. L., Katz, L. B., & Leirer, V. O. (1980). Cognitive representation of personality impressions: Organizational processes in first impression formation. *Journal of Personality and Social Psychology*, *39*, 1050–1063.

Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Is the production of aversive consequences necessary to create cognitive dissonance? *Journal of Personality and Social Psychology*, *70*(1), 5–16.

Hassin, R. R., Bargh, J. A., & Uleman, J. S. (2002). Spontaneous causal inferences. *Journal of Experimental Social Psychology*, *39*, 515–522.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behavior. *Journal of Personality and Social Psychology*, *37*, 25–38.

Hayes-Roth, B. (1977). Evolution of cognitive structures and processes. *Psychological Review*, *84*, 260–278.

Hazan, C., & Shaver, P. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry*, *5*, 1–22.

Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review*, *51*, 358–374.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hermans, D., Baeyens, F., & Eelen, P. (1998). Odours as affective-processing context for word evaluation: A case of cross-modal affective priming. *Cognition and Emotion*, *12*, 601–613.

Herr, P. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.

Hetts, J. J., & Pelham, B. W. (2001). A case for the nonconscious self-concept. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 105–124). Mahwah, NJ: Erlbaum.

Higgins, E. T. (1989). Continuities and discontinuities in self-regulatory and self-evaluative processes: A developmental theory relating self and affect. *Journal of Personality*, *57*, 407–444.

Higgins, E. T. (1996a). Emotional experiences: The pains and pleasures of distinct regulatory systems. In R. D. Kavanaugh, B. Zimmerberg, & S. Fein (Eds.), *Emotion: Interdisciplinary perspectives* (pp. 203–241). Mahwah, NJ: Erlbaum.

Higgins, E. T. (1996b). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1996c). The "self-digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*, 1062–1083.

Higgins, E. T. (1998). The aboutness principle: A pervasive influence on human inference. *Social Cognition* [Special Issue: Naive theories and social judgment], *16*(1), 173–198.

Higgins, E. T., Bargh, J. A., & Lombardi, W. (1985). Nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 59–69.

Higgins, E. T., & King, G. A. (1981). Accessibility of social constructs: Information-processing consequences of individual and contextual variability. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 69–122). Hillsdale, NJ: Erlbaum.

Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, *43*, 35–47.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Higgins, E. T., Roney, C. J., Crowe, E., & Hymes, C. (1994). Ideal versus ought predictions for approach and avoidance: Distinct self-regulatory systems. *Journal of Personality and Social Psychology*, *66*, 276–286.

Higgins, E. T., & Spiegel, S. (2004). Promotion and prevention strategies for self-regulation: A motivated cognition perspective. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 171–187). New York, NY: Guilford Press.

Hinkley, K., & Andersen, S. M. (1996). The working self-concept in transference: Significant-other activation and self-change. *Journal of Personality and Social Psychology*, *71*, 1279–1295.

Holland, R. W., Verplanken, B., & van Knippenberg, A. (2003). From repetition to conviction: Attitude accessibility as a determinant of attitude certainty. *Journal of Experimental Social Psychology*, *39*(6), 594–601.

Hommel, B., Muesseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–937.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology and Aging*, *17*(3), 482–495.

Idson, L., Chen, Liberman, N., & Higgins, E. T. (2000). Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective on hedonic intensity. *Journal of Experimental Social Psychology*, *36*, 252–274.

Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology*, *36*(6), 660–676.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*(3), 326–338.

Jastrow, J. (1906). *The subconscious*. Boston: Houghton-Mifflin.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107–117.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York: Academic Press.

Jones, J. T., Pelham, B. W., Carvallo, M., & Mirenberg, M. C. (2004). How do I love thee? Let me count the Js: Implicit egotism and interpersonal attraction. *Journal of Personality and Social Psychology*, *87*, 665–683.

Kaas, J. H. (1995). The reorganization of sensory and motor maps in adult mammals. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 51–70). Cambridge, MA: MIT Press.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the ne-

gation of stereotype associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871–888.

Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality and Social Psychology Bulletin*, *28*(1), 3–15.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24.

Kemp-Wheeler, S. M., & Hill, A. B. (1992). Semantic and emotional priming below objective detection threshold. *Cognition and Emotion*, *6*, 113–128.

Kircher, T. T. J., Senior, C., Phillips, M. L., Rabe-Hesketh, S., Benson, P. J., Bullmore, E. T., et al. (2001). Recognizing one's own face. *Cognition*, *78*, B1–B15.

Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin*, *23*, 736–742.

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, *80*(4), 669–685.

Koole, S. L., Smeets, K., van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology*, *77*, 111–125.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus response compatibility: A model and taxonomy. *Psychological Review*, *97*, 253–270.

Krosnick, J. A., Betz, A. L., Jussim, L. J., Lynn, A. R., & Stephens, L. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin*, *18*, 152–162.

Kruglanski, A. W. (1996). Motivated social cognition: Principles of the interface. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: A handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects of impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*, 448–468.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 331–378). San Diego, CA: Academic Press.

Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology*, *82*(3), 283–299.

Kunda, Z., & Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application, and inhibition. *Psychological Inquiry*, *10*(1), 12–22.

Kunda, Z., Sinclair, L., & Griffin, D. (1997). Equal ratings but separate meanings: Stereotypes and the construal of traits. *Journal of Personality and Social Psychology*, *72*(4), 720–734.

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*(4), 522–544.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*, 284–308.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, *207*, 557–558.

Lane, K., Mitchell, J. A., & Banaji, M. R. (2005). Me and my group: Cultural status can disrupt cognitive consistency. *Social Cognition*, *23*(4), 353–386.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Lefcourt, H. M. (1976). *Locus of control: Current trends in theory and research*. Oxford, UK: Erlbaum.

Leibenluft, E., Gobbini, M. I., Harrison, T., & Haxby, J. V. (2004). Mother's neural activation in response to pictures of their children and other children. *Biological Psychiatry*, *56*(4), 225–232.

Lenton, A. P., Blair, I. V., & Hastie, R. (2001). Illusions of gender: Stereotypes evoke false memories. *Journal of Experimental Social Psychology*, *37*, 3–14.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*, 275–287.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Leyens, J. Ph., Yzerbyt, V. Y., & Schadron, G. (1992). The social judgeability approach to stereotypes. *European Review of Social Psychology*, *3*, 91–120.

Li, P., & Moskowitz, G. B. (2005). *Two types of cognitive resources and their effects on stereotype activation*. Manuscript under review, University of Wisconsin.

Liberman, N., Molden, D. C., Idson, L. C., & Higgins, E. T. (2001). Promotion and prevention focus on alternative hypotheses: Implications for attributional functions. *Journal of Personality and Social Psychology*, *80*, 5–18.

Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdale activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, *8*, 720–722.

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African American's implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, *38*(4), 405–413.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology*, *82*, 5–18.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Logan, G. D. (1989). Automaticity and cognitive control. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 52–74). New York: Guilford Press.

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization. *Personality and Social Psychology Bulletin*, *13*, 411–429.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, *81*, 842–855.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835–1837.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, *69*, 397–407.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808–817.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, *76*, 200–213.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, *13*(5), 460–464.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37–47.

Macrae, C. N., Schloerscheidt, A. M., Bodenhausen, G. V., & Milne, A. B. (2002). Creating memory illusions: Expectancy-based processing and the generation of false memories. *Memory*, *10*, 63–80.

Maddox, K. B., & Gray, S. A. (2002). Cognitive representations of Black Americans: Re-exploring the role of skin tone. *Personality and Social Psychology Bulletin*, *28*, 250–259.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in im-

pression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, *59*, 27–37.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Social Psychology: General*, *114*, 159–188.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435–442.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278.

Meltzoff, A. N., & Moore, M. K. (1992). Early imitation within a functional framework: The importance of person identity, movement, and development. *Infant Behavior and Development*, *15*(4), 479–505.

Mikulincer, M. (1995). Attachment style and the mental representation of the self. *Journal of Personality and Social Psychology*, *69*, 1203–1215.

Mikulincer, M. (1998). Adult attachment style and affect regulation: Strategic variations in self-appraisals. *Journal of Personality and Social Psychology*, *75*, 420–435.

Mikulincer, M., & Arad, D. (1999). Attachment working models and cognitive openness in close relationships: A test of chronic and temporary accessibility effects. *Journal of Personality and Social Psychology*, *77*, 710–725.

Mikulincer, M., Gillath, O., & Shaver, P. R. (2002). Activation of the attachment system in adulthood: Threat-related primes increase the accessibility of mental representations of attachment figures. *Journal of Personality and Social Psychology*, *83*, 881–895.

Mikulincer, M., Hirschberger, G., Nachmias, O., & Gillath, O. (2001). The affective component of the secure base schema: Affective priming with representations of attachment security. *Journal of Personality and Social Psychology*, *81*, 305–321.

Mikulincer, M., Orbach, I., & Iavnieli, D. (1998). Adult attachment style and affect regulation: Strategic variations in subjective self-other similarity. *Journal of Personality and Social Psychology*, *75*, 436–448.

Mikulincer, M., & Shaver, P. R. (2004). Security-based self-representations in adulthood: Contents and processes. In W. S. Rholes & J. A. Simpson (Eds.), *Adult attachment: Theory, research, and clinical implications* (pp. 159–195). New York: Guilford Press.

Miller, D. T., Downs, J. S., & Prentice, D. A. (1998). Minimal conditions for the creation of a unit relationship: The social bond between birthdaymates. *European Journal of Social Psychology*, *28*(3), 475–481.

Miranda, R., & Andersen, S. M. (2005). *Induced processing efficiency in making pessimistic versus optimistic future-event predictions: Implications for depressive schemas*. Manuscript under review, New York University.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246–268.

Mitchell, J. A., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*(3), 455–469.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, *11*(6), 462–466.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology*, *65*, 469–485.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology*, *83*(5), 1029–1050.

Monteith, M. J., & Voils, C. (1998). Proneness to prejudiced responses: Toward understanding the authenticity of self-reported discrepancies. *Journal of Personality and Social Psychology, 75*, 901–916.

Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition, 19*(4), 395–417.

Moretti, M. M., & Higgins, E. T. (1999). Internal representations of others in self-regulation: A new look at a classic issue. *Social Cognition* [Special Issue: Social cognition and relationships], *17*(2), 186–208.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*, 467–470.

Moskowitz, G. B. (1993). Individual differences in social categorization: The effects of personal need for structure on spontaneous trait inferences. *Journal of Personality and Social Psychology*, *65*, 132–142.

Moskowitz, G. B. (2001). Preconscious control and compensatory cognition. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 333–358). Mahwah, NJ: Erlbaum.

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology*, *38*, 397–404.

Moskowitz, G. B. (2005). *Social cognition: Understanding self and others*. New York: Guilford Press.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167–184.

Moskowitz, G. B., Li, P., & Kirk, E. (2004). The implicit volition model: On the preconscious regulation of temporarily adopted goals. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 317–413). San Diego, CA: Academic Press.

Moskowitz, G. B., & Roman, R. J. (1992). Spontaneous trait inferences as self generated primes: Implications for conscious social judgment. *Journal of Personality and Social Psychology*, *62*, 728–738.

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Implicit control of stereotype activation through the preconscious operation of egalitarian goals. *Social Cognition*, *18*, 151–177.

Moskowitz, G. B., & Skurnik, I. W. (1999). Contrast effects as determined by the type of prime: Trait versus exemplar primes initiate processing strategies that differ in how accessible constructs are used. *Journal of Personality and Social Psychology*, *76*, 911–927.

Moskowitz, G. B., Skurnik, I., & Galinsky, A. D. (1999). The history of dual-process notions, and the future of preconscious control. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 12–36). New York: Guilford Press.

Murray, S. L., & Holmes, J. G. (1993). Seeing virtues in faults: Negativity and the transformation of interpersonal narratives in close relationships. *Journal of Personality and Social Psychology*, *65*, 707–722.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In F. N. Dempster & C. J. Brainerd (Eds.), *New perspectives on interference and inhibition in cognition* (pp. 207–261). New York: Academic Press.

Neuberg, S. L., & Newsom, J. (1993). Personal need for structure: Individual differences in the desire for simple structure. *Journal of Personality and Social Psychology*, *65*, 113–131.

Newman, L. S., & Uleman, J. S. (1990). Assimilation and contrast effects in spontaneous trait inferences. *Personality and Social Psychology Bulletin*, *16*, 224–240.

Niedenthal, P. M., Cantor, N., & Kihlstrom, J. F. (1985). Prototype matching: A strategy for social decision making. *Journal of Personality and Social Psychology*, *48*(3), 575–584.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*(6), 625–666.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002a). Harvesting intergroup implicit attitudes and beliefs from a demonstration Web site. *Group Dynamics, 6*(1), 101–115.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002b). Math = male, me = female, therefore math = me. *Journal of Personality and Social Psychology*, *83*(1), 44–59.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (in press). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior*. New York: Psychology Press.

Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, *15*, 353–361.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience, 14*, 1215–1229.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*(5), 413–417.

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science*, *309*, 785–787.

Oppenheim, D., Emde, R. N., & Warren, S. (1997). Children's narrative representations of mothers: Their development and associations with child and mother adaptation. *Child Development*, *68*(1), 127–138.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition*, *25*(3), 273–285.

Otten, S., & Moskowitz, G. B. (2000). Evidence for implicit evaluative in-group bias: Affect-biased spontaneous trait inference in a minimal group paradigm. *Journal of Experimental Social Psychology*, *36*, 77–89.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology*, *29*(8), 1049–1071.

Pascalis, O., de Schonen, S., & Morton, J. (1994). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, *18*(1), 79–85.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, *28*, 377–401.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293.

Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology*, *82*, 469–487.

Pendry, L., & Macrae, C. N. (1996). What the disinterested perceiver overlooks: Goal-directed social categorization. *Personality and Social Psychology Bulletin*, *22*, 249–256.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*(5), 729–738.

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience*, *4*, 437–441.

Pierce, T., & Lydon, J. E. (2001). Global and specific relational models in the experience of social interactions. *Journal of Personality and Social Psychology*, *80*(4), 613–631.

Pittman, T. S., & D'Agostino, P. R. (1989). Control deprivation and the nature of subsequent information processing. *Journal of Experimental Social Psychology*, *25*(6), 465–480.

Pittman, T. S., & Heller, J. F. (1987). Social motivation. *Annual Review of Psychology*, *38*, 461–489.

Pittman, T. S., & Pittman, N. L. (1980). Deprivation of control and the attribution process. *Journal of Personality and Social Psychology*, *39*(3), 377–389.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Plant, E. A., & Peruche, B. M. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, *16*(3), 180–183.

Poehlman, T. A., Uhlmann, E., Greenwald, A. G., & Banaji, M. R. (2005). *Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity*. Unpublished manuscript, Yale University, New Haven, CT.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcements. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: Vol. 2: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Reznik, I., & Andersen, S. M. (2004). *Becoming the dreaded self: Diminished self-worth with positive significant others in transference*. Unpublished manuscript, New York University.

Reznik, I., & Andersen, S. M (2006). *Agitation and despair in relation to parents: Activating emotional suffering in transference*. Manuscript under review, New York University.

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*, 550–560.

Richeson, J. A., & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, *39*(2), 177–183.

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, *6*, 1323–1328.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287–290.

Roskos-Ewoldsen, D. R., & Fazio, R. H. (1992). On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology*, *63*, 198–211.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 174–221). New York: Academic Press.

Rothbart, M., Sriram, N., & Davis-Stitt, C. (1996). The retrieval of typical and atypical category members. *Journal of Experimental Social Psychology*, *32*(4), 309–336.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*, 1–28.

Ruble, D. N., & Dweck, C. S. (1995). Self-conceptions, person conceptions and their development. In N. Eisenberg (Ed.), *Review of personality and social psychology: Vol. 15. Social development* (pp. 109–139). Thousand Oaks, CA: Sage.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*, 856–868.

Rudman, L. A., Feinberg, J., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition*, *20*(4), 294–320.

Rudolph, K. D., Hammen, C., & Burge, D. (1995). Cognitive representations of self, family, and peers in school-age children: Links with social competence and sociometric status. *Child Development*, *66*(5), 1385–1402.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognitive* (Vol. 1). Cambridge, MA: MIT Press.

Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of

academic motivation and self-esteem. *Journal of Early Adolescence*, *14*(2), 226–249.

Schaller, M., Park, J. H., & Mueller, A. (2003). Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes. *Personality and Social Psychology Bulletin*, *29*(5), 637–649.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing I: Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195–202.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? Informative functions of affective states. In K. Fiedler & J. P. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 44–62). Toronto: Hogrefe International.

Sczesny, S., & Kuhnen, U. (2004). Meta-cognition about biological sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence. *Personality and Social Psychology Bulletin*, *30*(1), 13–21.

Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology*, *80*, 645–654.

Shah, J. (2003a). Automatic for the people: How representations of significant others implicitly affect goal pursuit. *Journal of Personality and Social Psychology*, *84*, 661–681.

Shah, J. (2003b). The motivational looking glass: How significant others implicitly affect goal appraisals. *Journal of Personality and Social Psychology*, *85*, 424–439.

Shah, J. Y., Friedman, R., & Kruglanski, A. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology, 83*, 1261–1280.

Shah, J., & Higgins, E. T. (2001). Regulatory concerns and appraisal efficiency: The general impact of promotion and prevention. *Journal of Personality and Social Psychology*, *80*, 693–705.

Shah, J., Higgins, E. T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology*, *74*, 285–293.

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology, 38*, 368–383.

Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: Bottom-up priming of goals by means and its effects on self regulation. *Journal of Personality and Social Psychology*, *84*, 1109–1122.

Shields, A., Ryan, R. M., & Cicchetti, D. (2001). Narrative representations of caregivers and emotion dysregulation as predictors of maltreated children's rejection by peers. *Developmental Psychology*, *37*, 321–337.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

Simpson, J. A., & Rholes, W. S. (1998). *Attachment theory and close relationships*. New York: Guilford Press.

Sinclair, L., & Kunda, Z. (1999). Reactions to a black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*, 885–904.

Sinclair, L., & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychology Bulletin*, *26*(11), 1329–1342.

Sinclair, L., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: The role of affiliative motivation. *Journal of Personality and Social Psychology*, *89*(4), 583–592.

Skurnik, I., Moskowitz, G. B., & Johnson, M. (2006). *Biases in remembering true and false information: Illusions of truth and falseness*. Manuscript under review.

Skurnik, I., Schwarz, N., & Winkielman, P. (2000). Drawing inferences from feelings: The role of naive beliefs. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 162–175). New York: Psychology Press.

Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology*, *25*, 500–523.

Smith, E. R. (1996). What do connectionism and psychology have to offer each other? *Journal of Personality and Social Psychology*, *70*, 893–912.

Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist Interpretation of dual-process models. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 323–338). New York: Guilford Press.

Smith, E. R., & Mackie, D. M. (2005). Aggression, hatred, and other emotions. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 361–376). Malden, MA: Blackwell.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3–21.

Sorrentino, R. M., & Short, J. A. C. (1986). Uncertainty orientation, motivation, and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 379–403). New York: Guilford Press.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., & Dunn, M. A. (1998). Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin*, *24*, 1139–1152.

Spencer, S. J., Fein, S., Zanna, M. P., & Olson, J. M. (Eds.). (2003). *Motivated social perception: The Ontario Symposium* (Vol. 9). Mahwah, NJ: Erlbaum.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, *37*, 1660–1672.

Srull, T. K., & Wyer, R. S., Jr. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgment. *Journal of Personality and Social Psychology*, *38*, 841–856.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review*, *96*, 58–83.

Stapel, D. A., & Koomen, W. (2001a). The impact of interpretation versus comparison mindsets on knowledge accessibility effects. *Journal of Experimental Social Psychology*, *37*(2), 134–149.

Stapel, D. A., & Koomen, W. (2001b). Let's not forget the past when we go to the future: On our knowledge of knowledge accessibility. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 229–246). Mahwah, NJ: Erlbaum.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1996). The referents of trait inferences: The impact of trait concepts versus actor-trait links on subsequent judgments. *Journal of Personality and Social Psychology*, *70*, 437–450.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1997). Categories of category accessibility: The impact of trait concept versus exemplar priming on person judgments. *Journal of Experimental Social Psychology*, *33*, 47–76.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology: Vol. 21. Social psychological studies of the self: Perspectives and programs* (pp. 261–302). San Diego, CA: Academic Press.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*, 5–19.

Stone, J. (2001). Behavioral discrepancies and the role of construal processes in cognitive dissonance. In G. B. Moskowitz (Ed.), *Cognitive social psychology*: *The Princeton Symposium on the legacy and future of social cognition* (pp.41–58). Mahwah, NJ: Lawrence Erlbaum Associates.

Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology*, *37*, 228–243.

Stone, J., & Cooper, J. (2003). The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes. *Journal of Experimental Social Psychology*, *39*, 508–515.

Storbeck, J., & Clore, G. L. (2005). *Mood governs implicit processes: Affect, priming, and false memories*. Unpublished manuscript.

Storbeck, J. S., & Robinson, M. D. (2004). Preferences and inferences in encoding visual objects: A systematic comparison of semantic and affective priming. *Personality and Social Psychology Bulletin*, *30*(1), 81–93.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.

Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 579–596). New York: Guilford Press.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Strahan, E. J., Spencer, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: Striking while the iron is hot. *Journal of Experimental Social Psychology, 38*, 556–568.

Strauman, T. J., & Higgins, E. T. (1987). Automatic activation of self-discrepancies and emotional syndromes: When cognitive structures influence affect. *Journal of Personality and Social Psychology* [Special Issue: Integrating personality and social psychology], *53*(6), 1004–1014.

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin*, *95*(1), 3–28.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*, 96–102.

Tajfel, H., Billig, M. G., Bundy, R. F., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Psychology*, *1*, 149–177.

Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S., & Bargh, J. A. (1994). Accuracy motivation attenuates covert priming effects: The systematic *re*processing of social information. *Journal of Personality and Social Psychology*, *66*, 474–489.

Thompson, R. A. (1998). Early sociopersonality development. In W. Damon (Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 25–104). New York: Wiley.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology*, *37A*, 571–590.

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*, 1051–1065.

Tormala, Z. L., Petty, R. E., & Briñol, P. (2002). Ease of retrieval effects in persuasion: A self-validation analysis. *Personality and Social Psychology Bulletin*, *28*(12), 1700–1712.

Toth, S. L., Cicchetti, D., Macfie, J., & Emde, R. N. (1997). Representations of self and other in the narratives of neglected, physically abused, and sexually abused preschoolers. *Development and Psychopathology*, *9*(4), 781–796.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology*, *73*, 662–674.

Trope, Y., & Liberman, A. (1993). The use of trait conceptions to identify other people's behavior and to draw inferences about their personalities. *Personality and Social Psychology Bulletin*, *19*, 553–562.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Uhlmann, E., Dasgupta, N., Elgueta, A., Greenwald, A. G., & Swanson, J. (2002). Subgroup prejudice based on skin color among hispanics in the United States and Latin America. *Social Cognition*, *20*, 198–225.

Uleman, J. S., Hon, A., Roman, R., & Moskowitz, G. B. (1996). On-line evidence for spontaneous trait inferences. *Personality and Social Psychology Bulletin*, *22*, 337–394.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 211–280). San Diego, CA: Academic Press.

Uleman, J. S., Winborne, W. C., Winter, L., & Shechter, D. (1986). Personality differences in spontaneous personality inferences at encoding. *Journal of Personality and Social Psychology*, *51*, 396–403.

Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behaviour: Is habit an empty construct or an interesting case of automaticity? *European Review of Social Psychology*, *10*, 101–134.

Vohs, K. D., & Baumeister, R. F. (2004). Ego depletion, self-control, and choice. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 398–410). New York: Guilford Press.

Weary, G., Jacobson, J. A., & Edwards, J. A. (2001). Chronic and temporarily activated causal uncertainty beliefs and stereotype usage. *Journal of Personality and Social Psychology*, *81*, 206–219.

Wegener, D. T., Dunn, M., & Tokusato, D. (2001). The flexible correction model: Phenomenology and the use of naive theories in avoiding or removing bias. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 277–291). Mahwah, NJ: Erlbaum.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Mahwah, NJ: Erlbaum.

Wegener, D. T., Petty, R. E., & Dunn, M. (1998). The metacognition of bias correction: Naive theories of bias and the flexible correction model. In V. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Metacognition: Cognitive and social dimensions* (pp. 202–227). London: Sage.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 446–496). Boston: McGraw-Hill.

Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, *63*(6), 903–912.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., NcInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*, 411–418.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16*(1), 56–63.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Erlbaum.

Wiesel, T. N., & Hubel, D. H. (1965). Comparison of the effects of unilateral and bilateral eye closure on cortical unit response in kittens. *Journal of Neurophysiology*, *28*, 1029–1040.

Wilson, T., Lindsey, S., & Schooler, T. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117–142.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation increases positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*, 237–252.

Wittenbrink, B., Judd, C., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*, 815–827.

Woike, B. A., Lavezzary, E., & Barsky, J. (2001). The influence of implicit motives on memory processes. *Journal of Personality and Social Psychology, 81*, 935–945.

Wyer, R. S., Jr., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review*, *93*, 322–359.

Yovetich, N. A., & Rusbult, C. E. (1994). Accommodative behavior in close relationships: Exploring transformation of motivation. *Journal of Experimental Social Psychology, 30*(2), 138–164.

Yzerbyt, V. Y., Schadron, G., Leyens, J. P., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, *66*, 48–55.

Zajonc, R. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.

Zayas, V., & Shoda, Y. (2005). Do automatic reactions elicited by thoughts of a romantic partner, mother, and self relate to adult romantic attachment? *Personality and Social Psychology Bulletin, 31*, 1011–1025.

Zeigarnik, B. (1927). Das Behalten erledigter und unerledigter Handlungen [The retention of completed and uncompleted actions]. *Psychologische Forschung*, *9*, 1–85.

CHAPTER 8

Information Ecology
and the Explanation
of Social Cognition and Behavior

KLAUS FIEDLER

To understand the structure of the human mind and the psychological processes within the individual, we first have to consider the texture and contents of the stimulus environment that impinges on the individual's mind. Some of the most influential theorists—Lewin (1951), Brunswik (1956), Simon (1956), and Gibson (1979)—have articulated this fundamental insight many decades ago. In current scientific psychology, these claims have been widely accepted and regularly cited. However, in reality the confession that ecology matters comes more like a truism that hardly needs to be substantiated than like a necessary ingredient for sound research and theorizing. No doubt, the ecological part in the inseparable interaction of the mind and the environment has been widely neglected. Half a century after Lewin, Simon, and Brunswik, and 25 years after Gibson, it must be admitted that there is little room for ecological concepts in contemporary social psychology. No major textbook or handbook so far has covered the topic of the present chapter, information ecology. When a textbook chapter is exceptionally devoted to the "environment," the term is meant in a different way, referring to stress, crowding, or environmentalist behavior rather than ecological constraints on information processing. The same negative conclusion—it should be added quickly—can be drawn for the entire discipline of scientific psychology as it has rapidly grown and flourished throughout the last five decades.

To be sure, there have always been notable exceptions to this negative appraisal, which are consensually cited as testimony for an ecological approach (Neisser, 1985; Zebrowitz-McArthur & Baron, 1983). However, these contributions, important and prominent as they do appear, can be characterized as unusual, alternative approaches that supplement the mainstream paradigms, rather than merging with cognitive theories to yield new models and research designs that take both parts into account, the mind and the environment. The current social psychology was hardly influenced by Simon's (1956) metaphor that mind and environment have to collaborate like the two blades of a pair of scissors, well adjusted to each other and equally important (Gigerenzer, 2004b). There is no conceptual role played by ecological variables in the most influential theories and paradigms, such as theories of implicit cognition and priming, social comparison, intergroup relations, stereotype threat, mortality salience, motivated processes, aggression, and prosocial behavior. One might conjecture that, as a matter of fact, theorists claim all the time that stimulus learning underlies implicit cognition, that aggression reflects the media environment, that stereotypes depend on culture, or that social cognition is domain specific. However, granting this truism that all behavior has its original learning input, which depends on historical, cultural, and contextual moderators, the leading theories no longer seem to require ecological variables, which are deferred to the background.

PREVIEW OF THE PRESENT CHAPTER

Given the wide neglect of the present chapter's topic, or the mere background role assigned to ecological factors, the format and organization of this chapter will slightly

deviate from a common handbook frame. Both the explicit, acknowledged role of information ecology in social psychology as well as its implicit, unrecognized role played "under different conceptual cover" are examined in the following sections:

- A brief historical overview of major conceptual origins of genuinely ecological ideas provides a starting point, along with a short discussion of how they map onto modern social and cognitive psychology.
- As the conceptual underpinnings required for ecological theories have not been developed to a similar degree as cognitive, motivational, or personality theories, considerable space is devoted to introducing such common ground.
- Drawing on this conceptual framework, a review of pertinent research covers both explicitly ecological research and implicit contributions traditionally treated under different cover.
- Particular attention is given to alternative explanations of well-established findings, which have to be revisited as ecological factors are taken into account.
- Quite uncommon for a handbook chapter, but appropriate to the present topic, the review of research is supplemented with previews of research that has not yet been done but that strongly suggests itself for future research.

MAJOR ORIGINS, IN HISTORICAL PERSPECTIVE

Shared Ecological Premises in Psychological Science

The conviction that human beings—like all living creatures—must reflect to some degree the social and physical environment in which they are embedded is uncontested. The whole behaviorist research program (Kimmel, 1970; Skinner, 1984) with its radical emphasis on environmental stimuli as determinants of behavior and its neglect of constructive, creative, self-generative cognitive processes within the organism's mind can be considered a historical antecedent of ecological theorizing. However, apart from the central role assigned to the stimulus, assumptions about the structure of the stimulus environment in the real world are virtually missing in prominent behaviorist models (cf. Rescorla & Wagner, 1972). Learning and conditioning have been explained in terms of organisms' performance rather than ecological constraints, such as the distribution of reinforcement schedules in different parts of the world, the problem of how to segment the world into discrete stimuli, the construction of complex compound stimuli, or the measurement of the adaptive value of learned responses.

Not too remote from behaviorism is the biological and ethological approach to behavioral science (Eibl-Eibesfeldt, 1975; Lorenz, 1978). Just as animal behavior is attuned to the restrictions of the environment and the filters of phylogenetic learning and evolution, human behavior may be understood in broad biological terms. From such a perspective, it seems clear that aggression relates to territories, that attraction is limited by incest

barriers preventing young animals from choosing their mating partners among close relatives, or that specific signals can strongly influence instinctive reactions. Although such insights have been adopted from ethology, they continue to exert little influence on modern social psychology. These notions are not central for current theories of aggression, attraction, or emotion, with very few exceptions (Bischof, 1972).

Similar points could be made for social psychology's interface with macrosociology, as exemplified in the notion of social–economical status (Bornstein & Bradley, 2003), or microsociology, as exemplified in symbolic interactionism (Goffman, 1990; Mead, 1934). Both perspectives highlight the truism that individuals are but small pieces in a larger whole of society and societal games. But although celebrated in their respective areas, both perspectives did not contribute much to the development of a comprehensive conceptual framework of ecological variables. To be sure, symbolic interactionism reappears in Snyder's (1984) influential work on social hypothesis testing, or in Schlenker (1985) work on self-presentation, but it had little systematic impact on methods, research designs, and the structure of modern theories.

Seminal Contributions to Ecological Theorizing

Aside from those general precursors in psychology and neighboring disciplines, the roots of ecological theorizing can be found in the seminal work of a few pioneer scholars. One of them, Kurt Lewin, became very prominent in social psychology and inspired many leading researchers and paradigms. Another pioneer, Egon Brunswik, remained a tragic figure. After committing suicide, he continued to be widely ignored in social psychology. He had a stronger influence, though, in some areas of decision making and personality assessment.

The dominant metaphor for Lewin's (1951) ecological approach was the notion of *field theory*, conceived quite analogous to magnetic, electrical, or gravity fields. Accordingly, behavior is not only a function of organismic factors but also of the surrounding environment. Social-impact theory (Latané, 1981; Latané & Wolf, 1981) provides a prototypical example of how the field metaphor influenced psychological theorizing. Just as the illumination of a given point in space is a joint function of the number, the distance, and the strength of surrounding lightbulbs, the social impact that impinges on an individual is conceived as a function of the number, the (social or physical) distance, and the power (status, attraction, amount of resources) of other people. Social-impact theory accounts for many empirical phenomena, such as diffusion of responsibility (Pruitt & Teger, 1969), or the relative influence of majorities versus minorities (Nowak, Szamrej, & Latané, 1990; Latané & Wolf, 1981; Wolf, 1987).

The leading metaphor in Brunswik's (1956) work is his *lens model*, which describes the process by which ecological entities that are not themselves amenable to direct perception (e.g., distance of a stimulus object) can be inferred from more proximal, observable cues (disparity of retina images, surface texture, lightness, echo, etc.).

These cues provide the lens, as it were, through which the organism perceives the environment. The "ecological validities," or diagnosticities, of individual cues are typically rather modest. No single cue is indispensable, but cue substitution is crucial for Brunswik's probabilistic functionalism; if some person has but one eye, the central disparity cue in depth perception is compensated by other cues. Nevertheless, through parallel processing of multiple cues, one can reach a remarkable degree of accuracy (Funder, 1995; Kenney, Horber, Kashy, & Chu, 1992), even in highly fallible, probabilistic environments.

Gibson's (1979) seminal work was originally confined to visual perception, where it is perfectly clear that stimulus attributes (e.g., color) have to fit the properties of receptor organs (sensitivity to specific wave length). However, Gibson's approach turned out to be applicable more broadly to the adaptive fit between organisms and their physical and social environment (Valenti, Hofstra, & Good, 1991; Zebrowitz, 2002).

CONCEPTUAL UNDERPINNINGS FOR COGNITIVE–ECOLOGICAL RESEARCH

No doubt, the famous ecologists introduced in this historical sketch testify to the existence of ecological perspectives. Yet, their systematic impact on current theories and research methods remained modest, as already noted. Indeed, the topic of the present chapter—the information ecology providing the input to all cognitive processes—is so uncommon that a conceptual framework has to be developed first, before pertinent research can be reviewed. The framework to be outlined in this section consists of three facets: (1) the distribution of stimulus information; (2) the topology of environments experienced from different perspectives; and (3) the constraints imposed on the sampling of environmental information. Within all three facets, or basic principles of the information environment, more fine-graded distinctions between important features of the information ecology can be drawn (cf. Table 8.1).

Distribution of Stimulus Information

Three aspects of the distribution of stimulus information are of principal importance for the cognitive–ecological interface, the *density* of information in different parts of the world, the *variance* of information, or change in density, and the *redundancy* structure of the available stimulus data. It has become common sense that social cognition is a function of what subset of information happens to be accessible in memory, but the more fundamental question of what information the stimulus ecology renders available is rarely considered in models of judgment, decision making, attribution, and behavior.

Although Internet and mass media render most sources of information available, the information to which a particular person is actually exposed is multiply restricted in terms of language, culture, highly selective mass-media coverage, education, specific music, food, advertising, sports news, and societal topics experienced from a particular political viewpoint. What renders this selectivity of environmental input important and influential is that strong cognitive and behavioral biases can arise in completely unbiased people, just because the input is biased. It is crucial to recognize that the samples of information to which individuals are exposed are virtually never random. On one hand, the individual receiver exerts a selective bias on the information input due to her location in time, space, and social context (e.g., the subculture of other undergraduates). On the other hand, the providers of information impose restrictions on the information conveyed. Thus, consumers are unlikely to learn about negative aspects of products, because producers won't reveal such data, the media won't be paid for such information, and rival producers will not dare to publish negative information about others' products for legal reasons.

Similarly, social interaction will show a positivity bias (Kanouse & Hansen, 1972; Parducci, 1968) in that interaction (discussion, exchange, negotiation, dating etc.) with others is likely to stop when it is unpleasant but likely to continue when it is pleasant (Denrell, 2005). The opposite can also be true, though, as expressed in Ditto,

TABLE 8.1. Conceptual Framework and Terminology Required to Delineate the Role of Information Ecology

Aspects of the distribution of stimulus information	Topologies creating asymmetries	Sampling strategies
Density of information about different categories	Self–other asymmetry	Natural (unconditional) random sampling
	Ingroup–outgroup relation	Output-bound sampling
Variability	Old–young asymmetry	Input-bound sampling
Redundancy	Principal-agent relation	Sampling by availability
	Distant versus close in time	Sampling by social rules
	Distant versus close in space	Emergent sampling in group interaction
	Distant versus close in culture	Self-generated sampling
	Status asymmetry	Sampling of aggregate versus individuating information
	Actor versus observer perspective	

Munro, Apanovitch, Scepansky, and Lockhart's (2003) evolutionary thesis that organisms may invest more in processing aversive than pleasant stimuli (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998). Other biases in the information supply are equally sensible: Simplifying and commonly shared information is more likely to be broadcast than complicated (though often more accurate) information. News that is consistent with prior knowledge will be encountered, and actively solicited, more often than expectancy-inconsistent information (Frey, 1986; Jonas, Schulz-Hardt, Frey, & Thelen, 2001). Or, shared information is more likely to be raised in group discussions than new, unshared knowledge (Stasser & Titus, 1985; Wittenbaum, Hubbell, & Zuckerman, 1999), even when the latter is more informative.

As a consequence of such pervasive biases in the information ecology, the resulting information input is constrained in three psychologically important respects: density, variability, and redundancy. All three aspects have a profound impact on environmental learning. *Density* refers to the frequency of pertinent observations, thus constraining the number of learning trials on a topic afforded by the environment. Of two equivalent stimulus objects (e.g., two consumer products of equally high quality), the one object with the higher frequency or density of exposure is likely to be preferred just because there are more opportunities to learn about this object's assets (Fiedler, 2000a; Fiedler, Walther, Freytag, & Plessner, 2002; Shavitt, Sanbonmatsu, Smittipatana, & Posavac 1999). Similarly, of two equally aversive stimuli, the one that is met with higher density is likely to be feared and avoided more, as we know from aversive-conditioning experiments, in which the strength of an escape or avoidance reaction increases with the number of trials.

To understand the impact of *variability*, the analogy to learning experiments is again helpful. Just as a partial reinforcement schedule, in which an unconditioned stimulus (US) does not occur on all trials, makes learning resistant to extinction, a stimulus that varies (in terms of presence, quality, and appearance) keeps the environmental learning process active and alerted. It is a well-established empirical fact (Kamin, 1968; Rescorla & Wagner, 1972) that learning decreases and ceases as a stimulus can be predicted with perfect certainty and loses all surprise value. Variability can also influence adaptive-behavior strategies. To keep within the consumer example, a variable market environment not only keeps the consumer alert but also prevents her from developing too rigid routines and facilitates her trying out different strategies, which is important for adaptive learning in changing environments.

Finally, the notion of *redundancy* refers to environmental correlations that hold between different stimulus aspects. Social individuals—like organisms in general—are highly sensitive to all kinds of stimulus redundancies that help them to detect regularities in the environment (Garner, 1974; Melara & Algom, 2003), such as contingencies between causes and effects, signals and dangerous events, or stimulus aspects that seem to go together, like price and quality of consumer products. In a complex ecology, redundancy facilitates discrimination (Eiser & Stroebe, 1972; Tajfel, 1957). Thus, having noted the (apparent) correlation between price and quality, the consumer can use price as a proxy for inferring the quality of products, which is not visible at first sight.

All three aspects of the distribution of environmental stimuli, their density, variability, and redundancy, can have a profound influence on learning and adaptation. All three aspects will thus play an important role in the interpretation of the empirical findings to be reviewed in the next section. At the conceptual level, the abstract notions of density, variability, and redundancy, which refer to environmental structures, are related to corresponding cognitive notions, just as in psychophysics sound pressure corresponds to loudness, or wave length to perceived colors. In the same vein, the notion of density is related to the more familiar subjective experience of (high base rate) expectation. Variability relates to uncertainty or risk experience, and ecological redundancy corresponds to perceived or learned contingencies. It is typical of an ecological approach to cognitive information processing that familiar cognitive variables have to be complemented with less familiar ecological variables.

Topology of Environments from Different Perspectives

The very location and movement of individuals and groups in the environment creates asymmetries and distinct topological relations, which are at the heart of many phenomena in social psychology. One particularly prominent topological feature of the information environment is the *self–other asymmetry*—one of the first distinctions acquired in infant learning. Similarly, the *ingroup–outgroup asymmetry* is central to the social psychology of stereotyping, discrimination, and intergroup relations (Pettigrew, 1986; Tajfel & Turner, 1986). The amount and structure of information that is available about one's ingroup (and self) differs markedly from the information one has about outgroups (others) in terms of all three aspects—density, variability, and redundancy. As a matter of rule, both the density and the variability of ingroup-related and self-related information are higher than for outgroup-related and other-related information, just because one is closer to the self and one's own group than to others and other groups. This asymmetry contributes to various phenomena, such as the self-serving bias (Pronin, Gilovich, & Ross, 2004; Zuckerman, 1979), the ingroup-serving bias (Brewer, 1979), the actor-observer bias in attribution (Jones & Nisbett, 1972; Watson, 1982), or the outgroup homogeneity effect (Judd & Park, 1988; Park & Rothbart, 1982). Assuming the same prevalence of desirable over undesirable behavior in an ingroup (self) and an outgroup (other), the unequal density will make the prevalent desirability more apparent for the ingroup—thus resulting in ingroup favoritism. Likewise, because of the relative paucity of outgroup

(other-related) information, its variability will not be as apparent as for the ingroup—thus producing the illusion of outgroup homogeneity (Judd & Park, 1988).

The asymmetry in redundancy is slightly more complicated. The correlational structure of traits attributed to an outgroup, or a distant group, tends to be simplified, suggesting a halo effect (strong correlations among traits sharing the same valence) and a low-dimensional factor pattern (Liberman, Sagristano, & Trope, 2002; Linville, Fischer, & Yoon, 1996). Nevertheless, more complex patterns of actually existing correlations may be more evident for an ingroup, due to richer input information (Fiedler, Kemmelmeier, & Freytag, 1999; Linville, Fischer, & Salovey, 1989). In any case, the asymmetries of judgments and attributions that inevitably arise between self and others, between ingroups and outgroups, or more generally between close and distant targets, cannot be explained in terms of biased cognitive and motivational processes alone. To understand the nature of these phenomena, the density, variability, and redundancy of the pertinent information has to be taken into account.

One particularly prominent consequence of the self–other asymmetry is the actor–observer effect (Jones & Nisbett, 1972; Malle & Knobe, 1997; Watson, 1982), that is, the tendency to attribute others' behavior to internal dispositions and personality factors but to attribute one's own behavior to external, situational circumstances. Although a number of different factors contribute to this bias, one major determinant lies in the respective information environments. The perceptual field of actors judging their own behavior is directed toward other people and the surrounding situation, whereas observers typically focus on the actor whose behavior they are to judge, as the vivid figure against the ground (Storms, 1973). Conversely, actors possess more background information about the history and the external stimuli and motives that have triggered their behavior than observers do.

Other, less prominent asymmetries can have similar effects. For instance, old people know more about being young than young people know about being old. Judgments and stereotypes of old versus young people are likely to be influenced by this asymmetry in the information ecology. Comparable to the *old–young asymmetry*, migrants (unlike immobile residents) are exposed to other people and cultures from more than one perspective, or androgynous people experience life from both masculine and feminine perspectives. A whole variety of asymmetries and topologies can be generated from the concept of *psychological distance*, as raised by Lewin (1951) and recently refined by Trope and Libermann (2003). In addition to the perspectival distance between self and others, ingroups and outgroups, familiar people and strangers, psychological distance can refer to space, past and future time, status and power, or affective involvement. Common to all these distance dimensions is that information input about remote targets differs systematically from information about close targets, for ecological reasons alone.

Sampling Constraints

Whereas the topological relations considered so far place hard constraints on the distribution (i.e., density, variability, and redundancy) of stimulus input, the remaining facet of the information ecology entails a number of soft constraints that are imposed on the sampling of information. The process of information sampling involves a genuine interaction between human beings (or, more generally, organisms) and their social and physical environment. This interaction is not fully determined by the structure of the external world alone but is to a large extent contingent upon the organisms' actively chosen standpoints, movements, strategies of information search, and the ways in which organisms shape their own environment. A considerable part of the information that impinges upon human's sensory system—through conversation, literature, mass media, internet, teaching, or advertising—is self-made, reflecting human products or artifacts.

Thus, the environment is not a purely independent variable that determines the organism conceived as a purely dependent variable. Rather, manifold sampling processes reflect a variety of intriguing ways in which systematic perspectives and biases arise from the manner in which people solicit information from environments that differ in density, variability, and correlational structure. Let the term "sampling" refer to such active information search, which always involves an interaction between the individual and the environment. For example, for a teacher to reach a fair evaluation of student achievement, it is important to collect a representative sample of all students' performance. However, real samples are hardly ever random. The density (i.e., number) and variety of observations (i.e., learning trials) the teacher gets about individual students depends on environmental constraints (presence vs. absenteeism of students, how often they participate and raise their hand, their sitting position and salience, etc.) as well as the teacher's own information search (i.e., selective attention to different students, her testing hypotheses about specific students, and her prejudice and prior beliefs). Redundancy can facilitate the teacher's complex memory task or exaggerate differences that in reality are not that clear-cut. For example, a stereotypical correlation such that girls raise their hand more often in language lessons whereas boys participate more in science may accentuate the impact of gender stereotypes on the teacher's grading decisions.

Sampling is a key concept to understanding the potential and the limits of human information processing. Figure 8.1 illustrates the role of sampling as a crucial interface between cognition and the environment. The *latent environment* on the left refers to the "actual" states of nature that are the focus of judgments and behavioral decisions. The quotation marks here suggest that the states of nature need not be "true" in a logical sense, or "real" in an ontological sense. There is wide agreement that a student's *body size* is a really existing attribute, less agreement as to whether his *math ability* is an objectively existing trait, and little agreement as to his *likability*. However,
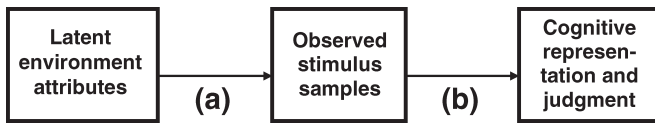
FIGURE 8.1. Two-stage model of the cognitive representation of the environment: (a) translation of environment into stimulus samples; (b) cognitive assessment of characteristics in the stimulus sample.

all three attributes are estimated all the time, based on samples of observations, just as population parameters in statistics are estimated from sample statistics. By analogy, samples of infinite or very large size may be conceived as approximations of the "true" attributes in the latent environment. A student's "actual" ability in math can be understood as the proportion of correct responses given in an infinite sample of math tasks. Similarly, his "actual" likability can be defined through the overall consensus of all persons who interact with this student.

However the latent attributes are defined—physically, logically, or through social consensus—they are never amenable to direct perception or assessment. We have neither sense organs to directly perceive the actual values of environmental attributes, such as lethal risks, dangerousness of action, honesty of people, ability of students or applicants, or quality of products, nor access to infinite samples or entire populations of relevant observations. Virtually all cognitive assessment is contingent upon finite samples, which for the reasons depicted earlier are virtually never random. Samples reflect both the topological constraints of the environment (i.e., distance, asymmetries) and sampling strategies of the individual, resulting in overrepresentation of some features and underrepresentation of others. Moreover, samples are characterized by differences in grain size (Goldsmith, Koriat, & Weinberg-Eliezer, 2002), biases toward salient features (McArthur, 1981), and accentuation of redundant aspects (Tajfel, 1957).

In any case, the effective input to the organism's psychological processes is not the environment per se but finite, selective samples reflecting specific perspectives, topologies, and sampling strategies. Traditional research in psychology has been mainly concerned with the *utilization* of stimulus samples controlled by experimenters. Theories and explanations of psychological phenomena have mostly, if not almost exclusively, referred to that utilization process, that is, the manner in which organisms process and transform the given stimulus input, as function of goals, habits, mood states, capacity restrictions, and learned or inherited automatic reactions (Bargh, 1996). In contrast, the *generation* of the stimulus samples that trigger such organismic functions was considered only exceptionally in theoretical accounts of stereotyping (cf. Eagly, 1987; Kashima, 2000), intergroup relations (Linville et al., 1989), or group decision making (Mojzisch & Schulz-Hardt, 2005). One prominent goal of a genuinely cognitive–ecological approach to social psychology is to correct for this neglect and to revisit empiri-

cal research in terms of the constraints placed on cognitive processes by the environment.

REVIEW OF PERTINENT EMPIRICAL RESEARCH

Toward this goal, the first part of the review of pertinent research is devoted to alternative accounts of well-known empirical phenomena, which typically originate in cognitive research treating individuals as the unit of analysis. The next part is also concerned with traditional research topics but at levels higher than isolated individuals, such as groups and dyads. The third major part then addresses paradigms that are explicitly built on cognitive–environmental theories. The final part addresses some dynamic aspects of information environments that characterize modern societies.

Prominent Research Findings Revisited

Although little research was planned deliberately to test the impact of density, variability, and redundancy, numerous findings can be explained alternatively in terms of these aspects of the information ecology. The present subsection is devoted to such reinterpretations of well-established findings.

With regard to density, to start with, a basic law says that learning increases monotonically with the number of learning trials. Existing trends become more and more visible as the frequency or density of relevant observations increases. Prior to reporting relevant empirical data, I illustrate this basic principle with an example. In an operant-conditioning experiment, pigeons may learn to discriminate between two response options, A and B, producing reward on 75% versus 25% of the trials, respectively. Before the first learning trial, the pigeon has to start with equal expectations; pecking rates for the two response keys have to be 50% versus 50% (assuming no bias). As the number of trials increases, pigeons will learn gradually that A is superior to B, inducing a gradual shift toward the superior option, A, up to a point when the learning process is complete and the learning curve has reached an asymptote.

The same gradual curves characterize the learning process in complex human environments, as evident in the following data from a simulated classroom (Fiedler et al., 2002). Taking the role of a teacher, participants had to learn the rates of correct responses given by 16 students of a school class that was graphically represented on the computer screen. Each student's performance was controlled by a parameter specifying the probability of correct responses. Students' correctness parameter could vary from .2 to .5 to .8 (i.e., from 20% to 80% correct responses). At the end of each of eight lessons, teachers had to estimate all 16 students' correctness proportions. Figure 8.2 provides the resulting curves over four blocks of two lessons, averaging over all participants and students sharing the same performance parameters. As in the animal-learning sketch, learning to discriminate be-
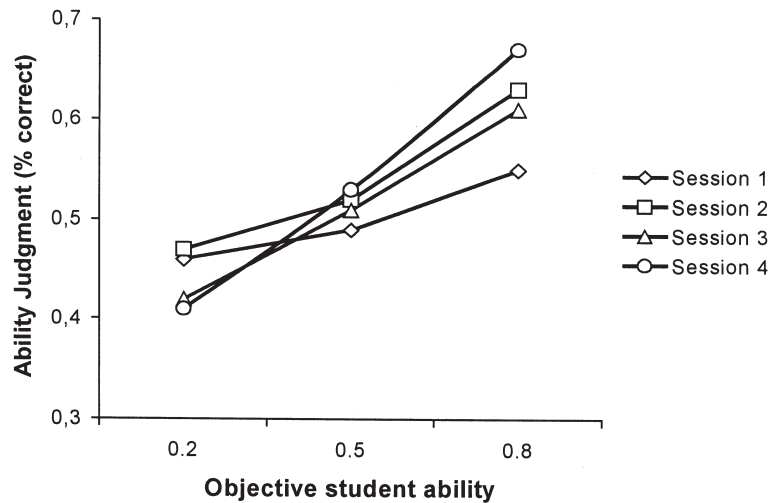
FIGURE 8.2. The influence of increasing learning experience, across four sessions of a simulated-classroom experiment (Fiedler et al., 2002), on the ability to discriminate between high and low student ability. From sessions 1 to 4, judgments are becoming less regressive (i.e., closer to the diagonal indicating perfect discrimination).

tween superior and inferior students took time. From the first to the last block, the judges' ability to discriminate smart and poor students increased gradually, and the curves became less regressive. However, typical of a complex probabilistic ecology, learning was never complete; even after the last session the estimates of the correctness rates of smart versus poor students remained regressive. Such regressiveness is a central feature of all probabilistic environments (Fiedler, 1996; Nesselroade, Stigler, & Baltes, 1980; Sedlmeier, 1999).

As a matter of principle, an existing (probabilistic) trend in an environment becomes more and more visible as the density of observations increases. This principle is evident in many other findings. In the simulated classroom mentioned earlier, the student ratings could also be analyzed as a function of sample size, that is, whether few or many observations had been gathered about individual students' performance (i.e., whether the rate of observations about a student was above (large samples) versus below the median (small) of all students. When the density of observations was high, smart students were more likely to be recognized as smart and poor students were more likely to be recognized as poor than when the density was low.

More pronounced judgments with increasing sample size—when the trend to be captured is held constant—are also evident from an experiment by Shavitt and colleagues (1999). In two experimental conditions, the number of observations about a group was varied, presenting either 24 positive and 0 negative behaviors or 12 positive and 0 negative behaviors. Although the positivity rate was invariantly 100%, the density of observations varied and resulted in more positive impressions of the group when sample size was large rather than small—a normal reflection of the density principle. In general, the so-called density bias is a well-established, general finding in research on contingency assessment (Allan, 1993; Fiedler, 2000b). The same contingency (e.g., the difference in the occur-

rence rate of an effect given the presence minus the absence of a cause) appears stronger when the absolute occurrence rate (or density) is high rather than low.

Ingroup Favoritism

By the same token, the density principle offers an obvious alternative interpretation for naturally occurring differences in sample size. The ingroup–outgroup asymmetry affords a striking paradigm. The density of observations that are available about one's ingroup is generally higher than for the outgroup, for obvious reasons. As a consequence, any trend, or multiple trends, assumed to be equally prevalent in both groups should be more readily detected for the ingroup than for the outgroup. Granting that trends normally favor desirable, norm-conforming behavior rather than undesirable, norm-violent behavior (Kanouse & Hansen, 1972), the density principle alone affords a sufficient account of ingroup favoritism (Brewer, 1979; Tajfel & Turner, 1986). The higher density of ingroup than outgroup information implies that this common trend toward desirable behavior is more apparent for the ingroup. These differences can be predicted in the absence of any cognitive or affective processing bias. A simple, monotonically increasing learning function, which is sensitive to number of trials, is sufficient to predict an ingroup advantage. Self-evident as this consideration may appear, it is missing in traditional accounts of ingroup favoritism or its complement, outgroup derogation.

Self-Serving Bias

An analogous argument applies to the self-serving bias, the bias to perceive and judge oneself in more positive terms than others (Krueger & Gilovich, 2004; Kunda, 1990; Mullen & Riordan, 1988; Pronin et al., 2004; Wood, Michela, & Giordano, 2000). More observations

can be assumed to be available about oneself than about others. This difference in density alone implies that the prevalent trend toward positivity in social behavior should be more apparent for oneself than for others, regardless of whether a genuine self-serving processing bias or unequal processing of self-related versus other-related information may also contribute to the phenomenon. By the same token, it can be explained that when more negative behavior is available about oneself than about others, negative self-judgments can be exaggerated too (Blanton, Axsom, McClive, & Price, 2001; Kruger, 1999; Moore, in press; Ross & Sicoly, 1979), contrary to a self-serving bias.

Outgroup Homogeneity and Personal Distinctiveness

The same assumption—that is, higher density of self-related and ingroup-related information than other-related and outgroup-related information—can explain two other prominent findings, the so-called outgroup-homogeneity effect (Judd & Park, 1988; Linville et al., 1989; Mullen & Hu, 1989; Quattrone, 1986; Simon, 1992) and the need for personal distinctiveness (Brewer, 1993, 2003). The former finding says that outgroups are represented in a less differentiated, more simplified fashion than ingroups, whereas the latter finding means that people want to set their personal, distinct identity apart from other people of their own group as well as from other groups.

An obvious ecological account relies on the same simple notion as the aforementioned accounts of ingroup-serving and self-serving biases. If the natural asymmetry of stimulus input renders learning about the self and one's ingroup more complete than learning about others and outgroups, then this should affect the learning not only of singular predominant trends but also of multiple trends that can be found in a person or group. Thus, of many attributes A, B, C, . . . , X assumed to exist in either person or group, a higher subset is likely to be noticed—above some threshold—for the self than for others, and for ingroups than for outgroups. The resulting representation of the self and of one's ingroup should thus be more distinct and differentiated than representations of others and outgroups.

Thus, the density rule offers a parsimonious account of the two major empirical phenomena of intergroup research—ingroup-serving bias and outgroup homogeneity—and, at the same time, two major findings of the social psychology of the self—the self-serving bias and the need for self-distinctiveness. The rationale underlying this ecological account seems plausible and easy to understand, but direct empirical evidence is scarce. Most pertinent evidence comes from computer simulations that demonstrate that divergent ingroup and outgroup judgments can be expected to evolve if two groups differ in only one aspect, the number or density of observations. Such simulation models (Fiedler, Kemmelmeier, & Freytag, 1999; Linville et al., 1989) start from the assumption that two groups, an ingroup and an outgroup, are equally represented on some attribute dimension (e.g., the degree of extraversion, or honesty).

Samples of values are then drawn at random from this distribution, reflecting the scale value of group behaviors on the attribute dimension. The samples supposed to represent the ingroup are larger (e.g., 24) than the samples supposed to represent the outgroup (e.g., 12). To simulate memory loss or forgetting, part of the samples' attribute values are erased or obscured, and the scale values of the remaining sample values are then transformed to categories on some judgment scale. Because of the reduced variance of small samples, the simulated judgments of "outgroups" show less variation than "ingroup" judgments. Simulated outgroup judgments are also less differentiated in terms of the number of different attributes exceeding some threshold.

Such simulation findings are corroborated in experiments in which fictitious groups are described through large versus small samples of stimulus behaviors. Subsequent judgments of the smaller group turn out to show less variation and differentiation (Linville et al., 1989), and the correlation pattern for multiple attribute judgments is simpler than for the larger group (Linville et al., 1996).

Using synthetic ingroups and outgroups that only differ in sample size, as in all the experiments and simulation studies considered so far, has the advantage that the impact of a single factor can be isolated. The alternative research strategy—relying on naturally existing groups or persons—would inevitably confound the variable of interest (i.e., density) with many other aspects (e.g., liking, identification, and familiarity). Nevertheless, a few studies have found ways to deal with this problem. Sande, Goethals, and Radloff (1988) asked their participants to judge themselves as well as other people on various attribute dimensions, holding the self–other comparison constant and merely varying the format of judgment. In one condition, self and other ratings had to be given on bipolar scales, the poles of which represented antonyms on attribute dimensions (e.g., extravert vs. introvert). In another condition, separate unipolar ratings for both antonyms (e.g., one rating for extravert and one for introvert) of all dimensions were called for. Using this paradigm, other persons received more extreme ratings on bipolar scales—reflecting more simplified, one-sided representations—whereas the sum of both antonym ratings was more pronounced for the self—reflecting that both opposites characterized the self at the same time. An analogous set of findings, using the same method as Sande and colleagues, was obtained for ingroup versus outgroup ratings by Fiedler, Kemmelmeier, and Freytag (1999). Thus, with increasing density of information, group representations become more complicated and less simplified. Simon's (1992) finding that outgroup homogeneity is often reversed for minority groups is also consistent with the important role attributed to numerosity (i.e., density) as a crucial distinctive feature between ingroups and outgroups.

Mere Thinking

The same ecological principle can account for seemingly unrelated phenomena. For instance, the mere-thinking

effect (Tesser, 1978; Tesser & Leone, 1977), quite detached from self and intergroup research, can be understood as reflecting the same density principle. Merely asking participants to extend thinking about an attitude target can be shown to increase the extremity of attitude judgments, provided the participants' knowledge base is rich enough to support extended sampling of arguments in memory. Thus, men develop more extreme attitudes through continued thinking about sports, whereas women's attitudes get more extreme when they continue thinking about fashion (cf. Tesser & Leone, 1977). Simply increasing the number of thoughts can render an attitude more extreme (cf. Downing, Judd, & Brauer, 1992); no selective bias toward attitude-consistent thoughts is necessary.

Biases and Heuristics

A number of other biases in judgments and decision making, which have been traditionally attributed to biased heuristic processing in the judge's mind, can be explained alternatively just by considering the information environment. One striking example to start with is *unrealistic optimism* (Armor & Taylor, 2002; Weinstein, 1980). Most people consider themselves to be above average in terms of ability and below average in terms of risk. This phenomenon is usually treated like a prototype of a motivated, self-serving bias. However, the distribution of environmental data is not at all incompatible with the conclusion that most people are better than average. Consider, for example, the conviction that "I am less likely than the average person to be involved in a car accident." Indeed, the distribution of the number accidents per person is clearly right-skewed; there are many people who had no accident or only one or two, whereas an increasingly smaller number of people have records of three, four or more accidents. Given such a skewed distribution, the median is well below the mean, implying that a majority of people are indeed better than average (Gigerenzer & Fiedler, 2007).

This is not to say that skewed distributions can account for all manifestations of unrealistic optimism (Price, 2001) or that motivated processes such as wishful thinking cannot also contribute. However, the normal ecological input distribution alone is sufficient to predict an optimistic bias, in the absence of any cognitive or affective bias in the judge's mind. Moreover, the higher risk attributed to the average other than to the self can again reflect a density effect. As Price (2001) has shown thinking about the risk of the average other involves a larger sample of people and events than thinking about the single self. Accordingly, the bias disappears when risk judgments of several others are averaged rather than considering judgments of the average other.

The *overconfidence* bias figures prominently in textbooks and in many applied areas of judgment and decision making (Klayman, Soll, González-Vallejo, & Barlas, 1999; Lichtenstein, Fischhoff, & Phillips, 1982; Soll, 1996) and in metacognition research (Koriat, Lichtenstein, & Fischhoff, 1980). When judges are asked to indicate the subjective confidence (in percent) that their answers to knowledge questions are correct, they prove to be ill-calibrated. That is, their percentage estimates tend to be higher than the actual proportion of correct responses on such questions. Moreover, the discrepancy increases with the extremity of subjective confidence (Juslin, Winman, & Olsson, 2000).

Given the prominent status of the overconfidence effect, a major challenge arises from recent findings showing that the phenomenon largely disappears when the task ecology is taken into account. The typical strong overconfidence effect is for the most part confined to studies that use arbitrary samples of stimulus tasks, rather than representative samples. In such studies, researchers tend to select judgment tasks that intuitively promise to induce more confidence than is warranted, thus creating a task environment that yields the predicted overconfidence effect. In contrast, a minority of studies using representative samples of knowledge tasks from a clearly circumscribed domain (e.g., population size of all cities of a country) yield drastically reduced effects or even null findings (Gigerenzer, Hoffrage, & Kleinbölting, 1991). A recent meta-analysis of Juslin and colleagues (2000) testifies to the stability of this ecological influence on overconfidence. Moreover, research by Winman, Juslin, and Björkman (1998) shows that, like the overconfidence effect, the prominent *hindsight bias* (Fischhoff, 1975) can also be eliminated or even reversed through representative sampling of stimuli or tasks.

Such provocative discoveries highlight the failure of both researchers and participants to take the information ecology into account. Researchers have overlooked the impact of arbitrary task sampling on research findings (Wells & Windschitl, 1999), whereas participants exposed to selected tasks fail to recognize that the experimental task environment is more difficult than the learning environment in which their subjective confidence feeling was calibrated.

Similar findings have been obtained in the classic heuristics-and-biases research program (Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982). Evidence on the availability heuristic (Tversky & Kahneman, 1973) also turned out to depend on researcher's intuitive generation of biased research environments. One of the most often cited findings—the overestimation of the frequency of words with the latter "k" in the first than the third position, supposed to reflect enhanced availability in memory—was shown by Sedlmeier, Hertwig, and Gigerenzer (1998) to hold only for the few letters that have been used intuitively in successful studies. For the rest of the alphabet, the phenomenon did not work. Although this demonstration need not undermine many other availability findings, it highlights the need to control for the stimulus ecology of psychological experiments.

Base-rate neglect (Bar-Hillel, 1990; Kahneman & Tversky, 1972), perhaps the most famous paradigm of the heuristics-and-biases program, refers to erroneous probability judgments of criterion events whose base rates differ from intuitive predictions. For example, the conditional probability that someone has HIV given a

positive test T_{HIV+} is drastically overestimated (Gigerenzer & Hoffrage, 1995; Swets, Dawes, & Monahan, 2000). As HIV is the cause most closely associated with T_{HIV+}, inferences from a positive T_{HIV+} result to HIV are readily made, neglecting the low base rate of HIV. Although—in the absence of further symptoms—the actual probability of the virus given a positive test, $p(HIV/T_{HIV+})$, is under 20% for the U.S. medical environment (cf. Swets et al., 2000), subjective estimates are extremely high, suggesting a strong base-rate fallacy.

This prominent bias is at least partly due to an information environment that invites output-bound judgment processes (Fiedler, Brinkmann, Betsch, & Wild, 2000). When judging $p(HIV/T_{HIV+})$, the logically appropriate reference set of all people having a positive test result T_{HIV+} is unknown. Available statistics and memory samples typically focus on the critical output events (e.g., the actual HIV cases), thereby inflating the likelihood of HIV cases in the effective judgment sample. When this ecological bias is eliminated (by inducing input-bound sampling with a focus on T_{HIV+} cases), the base-rate fallacy is greatly reduced (Fiedler et al., 2000; Gigerenzer & Hoffrage, 1995), as described in more detail in the section on biased hypothesis testing.

Dyads, Groups, and Distinct Social Topologies

The research revisited so far was not designed to study ecological influences. The general goal of the preceding section was to demonstrate that many well-known social cognition findings, obtained in intrapersonal task settings, can be explained in radically new ways as soon as the information environment is taken into account. The present section moves from single-person tasks to behaviors in dyads, small groups, or other settings that already constitute some interpersonal ecology. Nevertheless, the primary goal of these studies was still not environmental.

The explicit and implicit rules of social interaction among people in dyads and groups—as vividly described in early writings on symbolic interactionism (Goffman, 1990; Mead, 1934)—restrict and shape the social stimulus environment in distinct ways. What is likely to be said, written, and expressed; what behaviors are likely to be exhibited or hidden; and what information is sought or avoided depends on rules of conduct. Rather than revealing private motives and traits, people engage in self-presentation strategies and display rules, exaggerating normative constraints and reducing variability in the information ecology. The presence of an audience serves to reduce variability but to increase self-consistency (Duval & Wicklund, 1972; Schlenker, 1985) and thereby redundancy (i.e., correlations between different aspects of social behavior). Cooperation and social exchange rules oblige people to reciprocate actions shown by their interaction partners, thus imposing periodic distributive-justice rules on social interaction.

One particularly simple and straightforward assumption about the shaping of social input is as old as Thorndyke's (1916) law of effect (Nevin & Grace, 2000). Pleasant actions are more likely to be repeated and continued than unpleasant ones. Following this very simple

reinforcement law, it can be predicted that social interaction tends to be truncated when the input is negative in valence, whereas it tends to be continued as long as it is pleasant (Denrell, 2005). The resulting bias in the valence of social information input can account for a number of well-known phenomena. As already mentioned, the density distribution of stimulus behaviors is normally skewed toward more positive, desirable behaviors, which are the norm. Negative, undesirable behavior, which is the exception, is likely to be truncated and unlikely to be corrected resulting in more enduring negative than positive attitudes.

The same avoidance of unpleasant, conflict-prone experience characterizes *communication ecologies*, as evident in cognitive tuning (Zajonc, 1960). When communicating a message to somebody who is known to like (dislike) the message target, communicators tend to frame the message in positive (negative) ways. In the communication-game approach, Higgins (1981) has extended this notion, showing that communicators adapt in several ways to recipient needs. The communicators' own subsequent memory representations are then biased toward the message format tailored for the recipient's needs. For instance, if the recipient's prior knowledge is low, the message is supposed to be rich in detail, and so will the communicators own memory representation be characterized by the same grain size.

Grice's (1975) principle of cooperation and the four maxims of verbal communication—quality, quantity, relatedness, and manner—afford a sensible framework for understanding communication environments. The default assumption between speaker and listener (or writer and reader) is that all conveyed information validly reflects some aspect of reality (maxim of quality), is neither too scarce nor overly long (quantity), is related clearly to other aspects in the message (relatedness), and is not using strange and awkward words and style (manner). As a result, communicated information is characterized by a truth bias (maxim of quality; Granhag & Stroemwall, 2001; Rassin, 2000), a specifically chosen level of abstractness versus detail (quantity; Maass, 1999; Vallacher & Wegner, 1985), a level of internal coherence that exceeds the actual coherence of the world (relatedness; Kashima, 2000), and an avoidance of contents that cannot be expressed simply and clearly (manner; Conway & Schaller, 2007).

Using Allport's (1954) serial-reproduction paradigm, Kashima and colleagues (Kashima, 2000; Lyons & Kashima, 2003) let their participants communicate descriptions of stereotyped targets (e.g., gender targets) to others, who were then to retell the message to still other people and so forth, up to a fifth serial stage. From the beginning to the end of the communication chain, the ratio of stereotype-consistent to stereotype-inconsistent information increased steadily, thus reducing the complexity (quantity) and increasing the internal consistency (relatedness) of text. Further experiments showed that this was particularly the case when the task said communicators should communicate to others, but less so when the same task was labeled a (nonsocial) memory task. Thus, the choice of content did not reflect the communicators' re-

stricted capacity, but the impact of rules that shape the communication ecology (cf. Semin & Fiedler, 1988). This suggests an interpretation distinct from traditional stereotype theories, which emphasize intergroup conflicts, biased processing, or cognitive economy (Bodenhausen, 1990). Ironically, stereotyping can be fostered by the prosocial motive of cooperative communication.

The same holds for the linguistic-intergroup bias (Maass, 1999), that is, the tendency to communicate expected information (i.e., positive ingroup behaviors and negative outgroup behaviors) in more abstract terms than unexpected behaviors (i.e., negative ingroup behaviors and positive outgroup behaviors). This often replicated phenomenon (Karpinski & von Hippel, 1996; von Hippel, Sekaquaptewa, & Vargas, 1995; Wigboldus, Semin, & Spears, 2000) serves to stabilize and multiply stereotypical expectancies in media, literature, and conversation ecologies. However, recent studies using more refined methods clarified that this pervasive aspect of the communication environment need not arise from a self-serving or ingroup-serving motive. When motives and vested interests are pitted against expectancies, it turns out that expected information will be communicated abstractly even when it has negative implications for the self and the ingroup (Maass, Milesi, Zabbini, & Stahlberg, 1995; Wigboldus et al., 2000). Again, the prosocial motive to tune one's message toward the maxim of quantity (i.e., avoiding too much detail) suggests an ironic ecological source of social stereotyping. Such irony—pointing to unbiased or "innocent" causes of biased or undesirable results—is quite typical of the interplay of cognitive and environmental factors.

Group Decision Making

The maxim of quantity is immediately relevant to understanding one of the most prominent phenomena in group decision making, the *shared-information effect* (Larson, Foster-Fishman, & Keys, 1994; Stasser & Titus, 1985). In decision-oriented group discussions, thoughts and arguments that are shared by other group members and therefore add little new evidence are more likely to be uttered than unshared information, which has the potential to add something new. Decision-group environments thus tend to reduce the variability and to increase the redundancy of the information utilized. As a consequence, group decisions are systematically flawed when a so-called hidden profile is given, that is, when the shared information does not reveal the optimal solution, which is hidden in the unshared knowledge of individual specialists.

The shared-information effect seems to violate the maxim of quantity, because useful unshared information goes unmentioned, but it may actually reflect a compromise of the maxims of quantity and quality. Although shared information adds little new evidence, it is usually considered more valid and less idiosyncratic than unshared information (Greitemeier & Schulz-Hardt, 2003). Consensus among group members helps to fulfill the maxim of relatedness. Although a fully comprehensible explanation is not yet available (cf. Mojzisch & Schulz-

Hardt, 2005), the shared-information effect highlights the significance of information ecologies for social problem solving. Stasser and Titus's (1985) sampling model offers a viable account in terms of unbiased processes, by simply assuming that every group member samples arguments from his or her own knowledge.

A related and actually much older paradigm is the *group-polarization effect* (Lamm & Myers, 1978; Moscovici & Zavalloni, 1969; Myers, 1978). Group members tend to support and accentuate the predominant tendency in the group, giving rise to group decisions that are more polarized than the average individual decision. More recent research (Brauer, Judd, & Gliner, 1995) has shown that this is not merely due to the selectivity or reduced variability of the arguments raised in group discussions but can result from simply repeating the arguments representing the dominant tendency, that is, through increased information density alone. Again, from an ecological theory perspective, group polarization can be considered an interpersonal analog of the mere-thinking effect mentioned previously. Extended elaboration on a topic, whether in group discussion or by thinking alone, serves to increase the density of learning trials supporting a predominant information trend and hence leads to more extreme responses.

Social Hypothesis Testing

Since the pioneering work of Snyder and his colleagues (Snyder, 1984; Snyder & Swann, 1978; Swann, Giuliano, & Wegner, 1982), the process of social hypothesis testing is known to exhibit a systematic confirmation bias, reflecting the impact of distinct interaction rules on the information environment. Just as Grice's (1975) pragmatic approach, Snyder's (1984) theoretical analysis—deeply rooted in symbolic interactionism—emphasized the inherently cooperative nature of social-interaction environments. Thus, in order to test the hypothesis that one's interview partner is extravert, an interviewer is obliged to ask diagnostic questions related to extraversion and to refrain from asking less relevant questions about introversion and other traits. Such an information search strategy is called *positive testing* (Klayman & Ha, 1987); it serves to increase the density of information about the attribute that is in the focus of hypothesis. The interviewee, conversely, is obliged to cooperate with the interviewer's question focus. As the interviewer was interested in extraverted behavior, the interviewee is supposed to think and talk first about her extravert behaviors. (Had she been interviewed about introversion, she might have been able to talk about introversion as well.) This generalized tendency toward higher density of confirming yes responses than disconfirming no responses is well known in survey research as an *acquiescence effect* (Blau & Katerberg, 1982; Ray, 1983; Zuckerman, Knee, Hodgins, & Myake, 1995). Both tendencies together, positive testing in the interviewer (hypothesis tester) and acquiescence in the interviewee (target person), create an information ecology that fosters a confirmation bias. Even when the high acquiescence rate of confirming responses is equally high for extraversion and introversion ques-

tions, the enhanced density of observations for the focal attribute (extraversion) can explain a confirmation bias.

That density or sample size, rather than prior expectancies or the demands of the hypothesis-testing task, is the crucial factor that was tested by Fiedler, Walther, and Nickel (1999). Participants were free to ask questions about overt aggression and covert aggression observed in a male and a female target person, to test the hypothesis that male aggression tends to be overt whereas female aggression tends to be covert. In fact, the computer whose database they could use for information search was programmed to provide a constantly high proportion of 75% confirmative feedback, regardless of target gender and aggression type. Nevertheless, there was a marked bias toward the focal hypothesis. Frequency estimates and trait ratings revealed that the male target was judged to be higher in overt aggression and the female target was rated to be higher in covert aggression. The best predictor of this autoverification effect was the degree of positive testing (i.e., enhanced frequency of questions about male overt and female covert aggression).

However, importantly, neither prior beliefs nor the demands of the task focus could account for the full pattern of findings. When participants were asked to test the counterstereotypical hypothesis that the male target shows covert aggression and the female target overt aggression, a full reversal was obtained. Whether the focus was on a stereotype-consistent or -inconsistent hypothesis, the crucial factor was the density of stimulus observations. When judges witnessed the information search of another judge, it was possible to induce *negative testing* experimentally. Participants testing whether male aggression is overt could be exposed to a higher number of observations about female overt aggression. Density was shown to override both the hypothesis focus and the influence of stereotypical expectancies.

Analogous results were obtained in the aforementioned simulated classroom (Fiedler et al., 2002). Teachers supposed to test the hypothesis that girls are better in language whereas boys are better in science engaged in marked positive testing (asking many questions to girls in language lesson and to boys in science lessons), and the proportion of correct answers was constantly high. Consequently, they found support for the sexist stereotype. In contrast, when teachers were asked to test the reverse hypothesis—that in this class girls dominate in science and boys in language—they again engaged in (reverse) positive testing and actually found support for the counterstereotype. These findings only pertained to those gender targets who actually exhibited the necessary acquiescence effect (i.e., the high confirmation rate). Altogether, these results provide convergent evidence for differential information density as an important ecological factor.

The illusions resulting from hypothesis test strategies are particularly severe when judges engage in output-bound information search, as noted before. Participants in the study by Fiedler and colleagues (2000) were given a computerized data file to sample data relevant to estimating the probability $p(BC/M+)$ that a women has breast cancer (BC) given a positive mammogram (M+). In the database representing the ecological reality, the joint frequencies were distributed as follows:

4 BC and M+ cases
1 BC and M– cases
20 No BC and M+ cases
120 No BC and M– cases

Thus, the database was quite close to reality in that the base rate of BC was low (5 out of 145) but the hit rate of M+ given BC was clearly higher (i.e., 4/5 = 80%) than the false alarm rate of M+ given NoBC (20/140 = 14.3%). Due to the low BC base rate, the correct Bayesian solution in this task setting was $p(BC/M+) = 16.7\%$. The expectation that a woman with a positive mammogram actually has breast cancer is only 16.7%. Whether subjective estimates were highly inflated (as in previous research using such tasks; Eddy, 1982; Gigerenzer & Hoffrage, 1995) or quite accurate depended strongly on the search strategy. When sampling was output bound, contingent upon the criterion event BC, judges would typically draw all cases with BC and about the same number of cases without BC, receiving feedback as to whether the test result of these women was M+ or M–. Their resulting estimates reflected the BC proportion among M+ women in the sample quite accurately. But because the BC base rate was drastically inflated (now about one-half), these proportions were inflated too (64.0 %). In contrast, when sampling was input bound, considering women with M+ and M– and receiving feedback as to whether they had BC or not, the resulting estimates came rather close to the correct value (21.8%).

In general, to the extent that output-bound sampling overrepresents a low-base-rate event (e.g., BC) in the available sample, the actual rate will be highly overestimated. The ecological point here is that information sources (media, databanks, statistics archives) often enforce output-bound information search. For pragmatic reasons, the critical outcomes in question (e.g., HIV and car accidents) are much more visible and much more likely to be assessed in statistics than the relevant predictors (HIV test results, drinking alcohol). In Gavanski and Hui's (1992) terms, outcomes often afford a more natural and more visible sampling space than input factors, or independent variables. As a consequence, estimates of risks, diseases, and many other low-base-rate events are based on output-bound samples and therefore tend to be grossly inaccurate.

Memory as an Internalized Environment

A considerable part of the information ecology that informs behavioral decisions is internalized in memory (Anderson & Schooler, 1991). Most real problems call for the use of both externally provided and internally memorized knowledge. Hypothesis-testing processes use internal sources in the same fashion as externally provided information. Analogous to a group discussion, memory search can be conceived as a discussion with oneself or with one's autobiography. Not surprisingly, memory-based judgments and hypothesis tests have also

been shown to exhibit positive testing and confirmation biases (Koriat et al., 1980; Mussweiler, 2003).

Moreover, memory is subject to the same topological constraints as the external environment. Because memory provides more self-referent information than information about others, the impact or contribution of the self on social behavior tends to be overestimated (Ross & Sicoly, 1979). Self-referent information is not only high in density but high in resolution, or fine-grain size, reflecting that the distance to oneself is closer than to other persons (Fiedler, Semin, Finkenauer, & Berkel, 1995).

A distinct feature of memory, the internal part of the ecology, is its generative power. Merely prompting memory for particular contents—that is, considering, imagining, or mentally simulating and rehearsing possible contents—can instigate constructive processes that actively generate those contents (Koehler, 1991; Loftus, 1979; Wells & Gavanski, 1989). Merely imagining or explaining a possible event (accident, success experience) can raise the subjective belief in the truth of this event (Koehler, 1991). Or, merely focusing on an arbitrary target person identified only by her initials can let that focal person appear as above average on various judgment tasks (Klar & Giladi, 1997), presumably due to an increased virtual density of imagined or self-generated thoughts about the focal person or group (Price, 2001).

Although this active, generative property is most characteristic of internalized knowledge in memory, similar generative processes may be found in the external ecology. The notion of self-fulfilling prophecies (Jussim, 1986; Kukla, 1993; Rosenthal & Rubin, 1978) refers to the tendency to generate in active social interaction the very kind of information that is the focus of a question, quite comparable to the impact of imagination on memory. People who expect their interaction partners to possess some attribute will often treat them in a way that actually induces that attribute in the interaction partner (Snyder, Tanke, & Berscheid, 1977; Trope & Thompson, 1997).

The Dynamics of Social Comparisons

One particularly effective device for generating an environment that accords to one's motives, goals, and hypotheses is the selection of social comparison targets. As outlined in Festinger's (1954) old theory of social comparison processes, in order to determine our own identity, our abilities, relative attitudes, and the meaning of our sensations, we constantly need to compare ourselves with physical and social comparison targets. Physical comparisons help us to determine how strong, fast, or persistent we are (i.e., by comparing ourselves with weights or distances per time unit). But most of the time, when objective, physical comparisons are not available, we have to resort to social comparisons, using other persons and groups as graduators for our own attitudes and attributes.

Of course, the outcome of this process of self-definition depends crucially on the comparison environment, which in turn depends in part on the individual's motives and hypotheses that actively create an essential part of the information environment. People who are low in self-esteem or who belong to a stigmatized group are more likely to engage in safe downward comparisons that promise to foster one's self-esteem. In contrast, high self-esteem or membership in a successful group render upward comparisons more likely. In a similar vein, comparison direction may vary with the individual's regulatory focus (Higgins, 1998). Whereas a prevention focus induces a strategy to avoid threatening comparisons, a promotion focus may support engaging in exploring or even adventurous comparisons. Thus, ecological factors must not be considered purely independent variables, but in a genuine cognitive–ecological framework, ecological conditions are themselves shaped to a considerable degree by the individual's behaviors toward the environment (Higgins & Molden, 2003).

However, crucially, the choice of comparison is not determined by the individual's goals and motives alone (Wood, 1989). The range and number of available comparison targets can be severely restricted. For example, the intellectual development of a child in a school or family context may be hindered due to a lack of high-level comparisons (Zajonc, Markus, & Markus, 1979; Zajonc & Mullally, 1997). Moral rules of fidelity and incest taboo restrict the opportunities for social comparisons with sexual partners (Bischoff, 1972). Credentialist rules reduce the comparison and competition between people having different educational backgrounds. And segregation or discrimination practices may reduce the chances for comparisons with female leaders, gay child raisers, or members of different cultures.

Social Roles and Task Demands

Although the concept of social roles belongs to common sense, it only plays a minor role in current psychological theorizing. The most prominent exception here is Eagly's (1987) role account of gender stereotypes. As vividly illustrated in Hoffman and Hurst's (1990) experimental analog of Eagly's theory, individuals who play female as opposed to male roles will be perceived differently on relevant attribute dimensions, even when their actual position on those attributes has been carefully equated.

Genuinely Environmental Approaches

Ironically, when looking out for those most advanced theoretical approaches that come closest to the ideal of genuinely cognitive–environmental theories, one has to resort to some of the oldest sources. This section provides a cursory review of those seminal approaches, and how they map onto contemporary research.

Probabilistic Functionalism: Brunswik's Approach

Most prominently, Brunswik's (1955, 1956) probabilistic functionalism starts from the assumption that the distal entities that are the target of significant decisions and adaptive behavior—such as danger, risk, honesty, ability, or intention—are not amenable to direct perception.

Rather, they have to be actively inferred from more proximal cues that bear only probabilistic relations to the distal entity. For example, we have no sense organs for detecting the truth. Rather, lie detection and veracity judgments use multiple cues (like gaze, speech hesitations, pitch of the voice, etc.) that are at best weakly correlated with the objective truth (Vrij, 2000; Zuckerman, DePaulo, & Rosenthal, 1981). Nevertheless, because the brain uses multiple cues simultaneously, the aggregate accuracy of the entire cue system can be much higher than for individual cues. The process through which distal realities are actively reconstructed by the organism is extremely flexible and adaptive, characterized by cue substitution and "vicarious functioning." Thus, when particular cues are not available (e.g., the visual cues in lie detection when communications are presented only auditively), the inference process switches smoothly to using other, available cues. This occurs without effort and awareness. Thus, we continue to have stable depth perception when we cover one eye, although this eliminates the leading cue, disparity of the two retina images. Such adaptive and efficient cue utilization, leading to remarkable accuracy with even minimal information input (Ambady, Hallahan, & Rosenthal, 1995; Borkenau & Liebler, 1992; Funder, 1995; Kenney et al., 1992), holds for social judgments of likability, honesty, and many other personality attributes.

Brunswik's ecological approach, after a long period of being ignored, has finally given rise to Hammond, Stewart, Brehmer, and Steinman's (1975) social judgment theory (Cooksey, 1996), Gigerenzer and colleagues' (1991) probabilistic mental model approach, and several other developments in research on judgment and decision making (Juslin, 1994; Juslin & Olsson, 1997; Winman et al., 1998). In social psychology, it has inspired research on verbal and nonverbal communication (Zuckerman et al., 1981) and the study of accuracy and consensus in personality judgments (Funder, 1995). In all these domains, the description and analysis of the cue systems used for communication may be as central a research topic as the processes occurring within the individual.

Whether social judgments are accurate and correct depends not so much on the judge's effort, intelligence, or accuracy motivation but on whether the ecological input offers appropriate cues. Often, refraining from using too many cues and relying on a single, dominant cue can inform equally accurate or even better decisions than trying to exploit too many cues (Ambady & Rosenthal, 1992; Forest & Feldman, 2000; Gigerenzer & Goldstein, 1996). This less-is-more effect is at variance with the traditional tenet of dual-process models in social cognition, which presuppose that extended, more elaborated processing and effort expenditure are the keys to enhanced accuracy.

Affordances: Gibson's Approach

The emphasis on external stimuli and signals as determinants of human (and animal) behavior is also central to another environmental approach advocated by Gibson (1979). Though mainly concerned with visual percep-

tion, this approach is also relevant to the study of social behavior. Gibson's core concept of an *affordance* highlights the notion that the stimuli and behavioral options offered by the information ecology shape and restrict one's action space (Gaver, 1996; Valenti et al., 1991; Zebrowitz, 2002). Being offered an attractive job at the right time, meeting a romantic love partner, or being confronted with criminal temptations can determine a person's life forever. Affordances—conceived as environmental invitations, or opportunities hard to miss—are at the heart of social psychological topics such as role behavior (Eagly, 1987), demand characteristics (Orne, 1962; Shimp, Hyatt, & Snyder, 1991), and the impact of salience on attributions and behavior (Taylor, Fiske, Etcoff, & Ruderman, 1978). Salient stimuli and persons have been shown to attract more attention and more effective memory encoding, and are more likely to be perceived as causal origins than less salient sources (McArthur, 1981; Nothdurft, 1993).

However, notably, in a cognitive–environmental framework, affordances should not be misunderstood as purely external explanations. For an affordance to be effective, it has to meet a predisposition, motive, or propensity of the organism to respond accordingly. For example, the very preparedness of human observers to respond so sensitively to faces complements the affordance role that faces can play in triggering social behavior (Zebrowitz, Fellous, Mignault, & Andreoletti, 2003; Zebrowitz-McArthur & Baron, 1983). The schema of a baby face, for instance, has been shown to influence judicial judgments of guilt and punishment (Berry & Zebrowitz-McArthur, 1988). The same description of a crime led to different ratings of guilt and appropriate punishment when a photo showing the defendant's face was manipulated. Given a negligent crime due to naivety, less severe punishments were recommended for baby-faced than mature-faced defendants. In contrast, for an intentional offense, more severe punishments were recommended for baby-faced than mature-faced defendants. Apparently, guilt judgments depend on the fit between the nature of the crime and the surplus meaning of the face.

Domain Specificity: Simon's Approach

One intrinsic feature of ecological models is domain specificity. The notion that the environment can codetermine behavior entails, logically, the assumption that different environments, or domains, can drive behavior in different directions. Arguing from an evolutionary point of view, Cosmides and Tooby (1994) coined the metaphor of a Swiss Army knife to explain the way in which phylogenetic learning has equipped organisms with specific skills or devices for specific situations. Quite in line with Simon's (1956) bounded rationality notion, humans may not acquire general rules of logical reasoning, but they may rather apply specialized tools that can solve practical problems within specific domains. For example, a long tradition of research on Wason's (1966) selection task shows that people have no general understanding of logical implications of the form "if p, then q."

However, in the biologically important domain of detecting cheaters (Gigerenzer & Hug, 1992), they fully understand the implications of logically equivalent social contracts (Cheng & Holyoak, 1985; Cosmides, 1989), such as "If one drives a car, she must have a licence." Given such a rule, the cheater detector knows it is not merely important to look at people who drive a car (i.e., the p clause in "if p, then q") but also to monitor people who have no driving licence (the non-q clause).

The notion of domain specificity entails, logically, the assumption that adaptive intelligence is not striving for cognitive algorithms that provide optimal and universally correct solutions under all conditions. Rather, intelligence is content with satisficing algorithms or heuristics that provide viable local solutions in a given problem context, in accordance with Simon's (1956) notion of bounded rationality (restricted by limited capacity). An alternative, slightly different criterion for adaptive behavior is ecological reality (tuning organisms to learn those skills that fit the problem structure of the environment; Gigerenzer, Todd, & ABC Research Group, 1999).

In current social psychology, the most prominent variant of domain specificity is culture-specific variation (Markus & Kitayama, 1991; Nisbett, Peng, Choi, & Norenzayan, 2001; Triandis, 1995). Like other ecological factors, different cultures are characterized by information distributions that differ in density, variability, and redundancy. For instance, in the language input of people living in collectivistic cultures, the density and variety of verbs and process-centered terms is higher, whereas the density and variety of adjectives, nouns, and object-centered terms is lower than in individualistic cultures (Choi, Nisbett, & Norenzayan, 1999; Kashima, 2001; Rhee, Uleman, Lee, & Roman, 1995). This difference in linguistic input offers an explanation for why, in language development, Asian children show a relative advantage in understanding and using verbs as compared with nouns and object-centered terms (Choi & Gopnik, 1995; Tarif, Gelman, & Xu, 1999). This culture-specific impact on the ontogenetic learning input can in turn be assumed to trigger intercultural differences in thinking and reasoning, such as the lower rate of trait attributions (Choi et al., 1999; Miller, 1984) and of the correspondence bias (Miyamoto & Kitayama, 2002) in Asian (collectivistic) than in Western (individualistic) cultures, or their reduced use of static, logical categories in thinking and reasoning (Nisbett et al., 2001). A more extensive review of intercultural research can be found in Chapter 34 (Chiu and Hong, this volume).

Psychological Distance: Lewin's Approach

In Lewin's (1951) field-theoretic approach, psychological distance was a concept of central importance. In spite of the huge impact that Lewin had on the growth of social psychology after World War II, his notion of distance received little attention until very recently. Originally, the psychodynamics of distance had been contaminated with valence. As the distance from a reinforcing event decreases, the unpleasant, aversive aspects loom larger, whereas the pleasant, appetitive aspects are less apparent. Therefore, avoidance tendencies increase at a faster rate than approach tendencies with decreasing distance (Miller, 1944).

This basic asymmetry of positive and negative stimuli was given up by Trope and Liberman (2003), whose construal-level theory entails a new comprehensive conception of psychological distance, with rich theoretical and practical implications. Most of the research conducted within this framework refers to temporal distance, although the theory is also applicable to other dimensions, such as social, affective, spatial, or status distance. The central assumption underlying Trope and Liberman's (2000, 2003) theory says that behaviors and events in the far future tend to be construed at a higher level of abstraction, and more context independent, than behaviors and events in the near future. Thus, when we are making plans for next year, we think of doing more sports because it is inherently healthy, going to Vietnam for vacation because Vietnam is inherently attractive, or buying a convertible car that gives pleasure in summer. In contrast, when making decisions for tomorrow, we decide against sports because time is scarce, we recognize that other countries but Vietnam may be more familiar and easier to reach, and that the old used car will do it one more year, before the convertible dream can be realized in another far future. Far-future options tend to be represented by their inherent value, whereas the representation of present and near future is more sensitive to external constraints and circumstances that distract one from following the intrinsic value of goals and objects. As Liberman and Trope (1998) phrase it, the high-level construal induced from a distant perspective focuses on the desirability of goals, whereas the low-level construal induced from a proximal perspective revolves around the feasibility of actions.

Table 8.2 summarizes a number of findings about preference reversals in decision and choice that can be derived from, or explained within, construal-level theory. For example, the theory predicts, correctly, that of two decision options with the same expected value, the one with the higher outcome in case of success is preferred in the long run, but the one option with the higher likelihood of success will be preferred in the short run (Sagristano, Trope, & Liberman, 2002). For another example, when risk-related information is based on first-hand experience or direct observation (low distance), rare events are underestimated (consistent with the enhanced sensitivity to probabilistic information). However, when indirect, secondhand information is presented verbally or numerically (high distance), rare events are overestimated (Hertwig, Barron, Weber, & Erev, 2004; Weber, Shafir, & Blais, 2004). By the same token, construal-level theory can assimilate strong influences on decision making, and preference reversals, obtained when decision options are presented affectively (low distance) or purely descriptively (high distance; Hsee & Rottenstreich, 2004), when outcomes are decomposed or unpacked into their components (low distance) or considered globally (high distance; Fiedler & Armbruster, 1994; Tversky & Koehler, 1994), or when the decision process is at the beginning (high distance) or

TABLE 8.2. Influences on Decision Making and Sources of Preference Reversal That Can Be Derived from, or Assimilated within, Trope and Liberman's (2003) Construal-Level Theory

Distance dimension	Empirical evidence obtained with increasing psychological distance:	Sample reference
Temporal	Stronger impact of value component, lesser impact of probability component	Sagristano et al. (2002)
Affective	Increased consideration of scope (number of cases affected by decision) rather than affective value of a single outcome	Hsee & Rottenstreich (in press)
Grain size	Stronger planning fallacy, reflecting focus on planned option and failure to unpack multitude of alternative options	Kruger & Evans (2004)
Inclusiveness	Reduced attraction of a chosen option when all alternative outcomes are packed as one large category rather than many small categories	Windschitl & Wells (1998)
Investment	Reduced sunk-cost effect as long as commitment to decision is not too high	Arkes & Blumer (1985)
Commitment	More weight given to payoff as opposed to probability when task focuses on pricing (desirability) rather than choice (feasibility)	Slovic (1995)
Participation	Overweighting of rare events when infrequency is not experienced (firsthand information) but stated in summary statistic or verbal quantifier (secondhand information)	Hertwig et al. (2004)

has already reached a high degree of commitment (low distance, as in a so-called sunk-cost situation; Arkes & Blumer, 1985).

Cognitive–Ecological Rivalry: Trivariate Reasoning

Understanding the interplay of personal and (ecological) determinants of behavior can itself become an objective of psychological experiments. In many attribution experiments, judges are presented with behaviors that can be attributed to internal personality factors or to external factors in the environment. Let this paradigm be called *trivariate reasoning*, as the task involves three variables; two independent variables, X and Z, competing for the explanation of one dependent variable Y (cf. Spellman, 1996). Typically, Y is the behavior to be explained (e.g., a hostile act), X is a personal cause (e.g., aggressiveness trait), and Z is an ecological condition (e.g., media violence).

The tendency to give more weight to personal than to ecological factors in this trivariate paradigm is well known as the fundamental attribution bias (Ross, 1977). Moreover, the failure to take both determinants into account is evident in another major finding, the discounting effect (Kelley, 1973; Kruglanski, Schwartz, Maides, & Hamel, 1978; McClure, 1998; Morris & Larrick, 1995). To the extent that one factor is known to exert an influence, the judged impact of the other factor decreases. The discounting principle assumes that the impact of the primary factor, X, on the dependent variable Y, decreases when the secondary factor, Z, is known to be present. As illustrated in Figure 8.3, this is the model of a *spurious correlation*, in which Z can account for (part of the) influence of X on Y. As Z is partialed out (i.e., the X–Y relation is considered at separate levels of Z), the impact decreases or disappears.

However, discounting is but one way in which personal and ecological factors can codetermine an effect. Figure 8.3 reveals several other possibilities. The influence of the person (X) and the situation (Z) on behavior (Y) could be independent (orthogonal). Or, the inclusion of a secondary (external) factor, Z, could increase the relation of Y to the primary (internal) factor X—a constellation known as a *suppressor effect* (Conger & Jackson, 1972). This may occur when Z is correlated with that variance component of X that is uncorrelated with Y, thus absorbing "error" variance. (For example, Z might be a law that prohibits the individual from acting out her trait.) In such a suppressor situation, partialing out Z increases, rather than decreases, the relation of X and Y.

Trivariate problems as in Figure 8.3, calling for the integration of two or more determinants (e.g., internal and external; personal and environmental), are hard to solve, for researchers and research participants alike. This inherent difficulty of trivariate reasoning is emphasized in Simpson's paradox (Schaller, 1992a, 1992b; Spellman, Price, & Logan, 2001), which is based on the logic of a spurious correlation. In a typical task, participants are presented with a series of observations showing that more male than female applicants are accepted for a graduate program. The default explanation in terms of personality factors suggests that applicant gender (X) is a

Uni-causal influence. z has no impact

x —0.50— y 0.00   0.00 z $r^{*}_{xy.z} = r_{xy}$

Orthogonal pattern. z has independent influence

x —0.50— y 0.00   0.50 z $r^{*}_{xy.z} = r_{xy}$

Suppressor. z suppresses error variance in x-y relation

x —0.50— y −0.50   0.00 z $r^{*}_{xy.z} > r_{xy}$

Spurious correlation. z accounts for x-y relation

x —0.25— y 0.50   0.50 z $r^{*}_{xy.z} < r_{xy}$
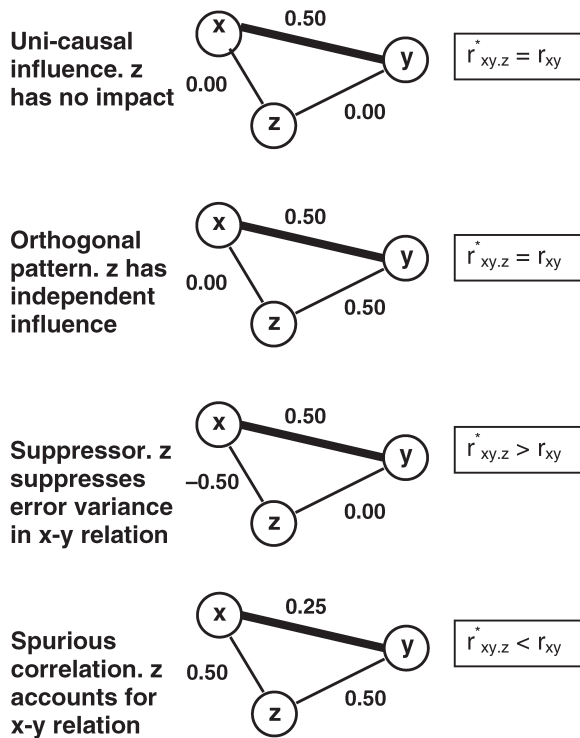
FIGURE 8.3. Different trivariate relations describing the way in which behavioral effects on a dependent variable y are codetermined by an individual factor x and an ecological factor z. The ecological factor, z, may be ineffective, independent, increase, or decrease the impact of x on y, as evident in a partial correlation $r^{*}_{xy.z}$ is equal, larger, or smaller than the focal zero-order correlation r_{xy} between x and y.

significant determinant of graduate admission (Y). However, in fact, a second, ecological determinant (Z) is involved, the distinction between two different universities. If this factor is taken into account, it turns out that within both universities, the proportion of accepted women is actually higher. The solution of this paradox is that one university is much more prestigious and has a much higher rejection rate than the other. As more women than men apply to the prestigious university, it is possible that the absolute success rate for women is reduced although the relative proportion of successful women is higher at both universities. According to Schaller (1992a, 1992b), feministic motivation facilitates the recognition of a female advantage, as opposed to the global male advantage. Other findings (Waldmann & Hagmayer, 1995) demonstrate that the hidden female advantage is more likely to be noticed when a plausible causal model is available.

However, crucially, participants have a hard time refraining from discounting and to understand that both conclusions can be true at the same time. Although females are superior at university level, it remains true that male applicants are superior at the aggregate level. The partial correlation between gender and achievement, with universities partialed out (favoring females), is no more real than the zero-order correlation across universi-

ties (favoring males). After all, the rejection rates of the two universities can be interpreted differently. It is possible of course that the high rejection rate of one university is simply due to higher standards, but it is also possible that the high rejection rate reflects the high number of (inferior) females joining this university. So the ultimate solution of Simpson's paradox (as of many other ecological problems) is to recognize that both opposing conclusions hold at the same time. This level of cognitive–ecological understanding is extremely hard to meet (Fiedler, Walther, Freytag, & Nickel, 2003).

Ecological Correlations

The seemingly paradoxical relation between ecological and personal explanations is highlighted in the famous notion of an ecological correlation (Robinson, 1950). The correlation, say, between skin color (being Black) and illiteracy across a large representative sample of U.S. individuals amounted to +.203. However, the correlation between the proportion of Blacks and the proportion illiteracy across the Census Bureau's nine geographic divisions was as high as +.906. Thus, depending on what level of aggregation is considered, individuals or ecologies, quite different conclusions about the same correlation are valid. An implication, which suggests itself, is that social stereotypes may to an unknown degree reflect the existence of such ecological correlations (Fiedler & Freytag, 2004). The neglect of this exciting phenomenon in modern social psychology is symptomatic for the lack of theoretical models that integrate personal as well as ecological laws, within the same framework.

Ecological correlations may be involved in some of the most prominent research topics that draw on ecological variables. For example, with regard to gender stereotypes, the relationship between leadership and female participation is certainly negative across vocational ecologies (i.e., in high-leadership areas, the prevalence of women is low), even though individual females may be strong leaders within specific professions. Correlations at different aggregation levels may not only diverge in size but even reverse their sign. A stereotypical relation may hold at aggregate level while disappearing or even reversing at individual level. Because the statistical boundary conditions that produce or eliminate ecological bias are well understood (Hammond, 1973; Hannan, 1970), this approach has promising implications for stereotyping research.

In a similar vein, when social roles (Eagly, 1987; Hoffman & Hurst, 1990) are used as constructs to explain individual behavior, a high-level aggregation construct (roles) is used to explain individual behavior, again raising a problem of ecological correlations. The same holds for intercultural studies, when differences that exist between cultures (Kühnen, Hannover, & Schubert, 2001; Markus & Kitayama, 1991; Nisbett et al., 2001) are used to explain differences in individual behaviors. Equally prominent are evolutionary accounts (Cosmides & Tooby, 1992; Tooby & Cosmides, 1992) assuming that correlations that hold over large temporal periods can explain the behavior of individual persons nested in one

generation (for an enlightening critique, see Lewontin, 1979). The point here is not that analogies and explanations in terms of roles, cultures, or evolution must be misleading. The point is, rather, that the ecological-correlation framework—which specifies the limiting conditions under which such aggregate constructs can be supposed to predict individual behaviors—is known for decades but is still waiting to be applied in social psychology (cf. Slatin, 1969).

Dynamic and Complex Features of Modern Information Societies

Among the most interesting cognitive–ecological research designs are those written by the modern information society itself. The following research topics promise to be content valid and highly pertinent to the interaction of environmental stimulus input and the preparedness of individual cognition: advertising and consumer behavior (Miller & Berry, 1998), risk assessment and the interpretation of statistics (Gigerenzer et al., 1999; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000), health-related decisions in light of health-related publications (Lindsey, Hertwig, & Gigerenzer, 2003), sampling bias in criminality assessment (Slatin, 1969), the impact of the Internet and other modern communication media (Gaver, 1996), the role of mass media in the formation of political attitudes, affordances and hindrances of globalization, communication problems related to worldwide migration movements, international law and translation problems, memory and cognition in elderly people, ecological and economical factors in education, political attitudes and action related to such existential topics as birth control, energy consumption, or AIDS, and various other highly prominent topics. A common denominator of all these challenges and threats of the new millennium is that psychological aspects of information transmission are of ultimate importance (Combs & Slovic, 1979). Hardly any other scientific discipline has to play a similarly responsible, crucial role as psychology, because the communication of problem-relevant information can be as important as the investment into medical, physical, geological, or astrophysical solutions, which are often far out of reach.

Publicizing Scientific Information

For example, the mad-cow disease in Europe, which had an incisive impact on economics and behavior around the turn of the century, has been almost forgotten and no longer influences consumer behavior, although the potential danger remains largely unchanged, just because the mass media have dropped this favorite topic. A number of recently published studies testify to the claim that dealing with information about problems and risk can be equally important as the problems themselves. Gigerenzer (2004a) provides compelling evidence showing that the number of people dying from highway accidents due to the media reactions to September 11, 2001, rose before the end of the year 2001 to a number higher than the primary victims of the terror attack. Equally pertinent is the Swets and colleagues (2000) finding mentioned earlier that given a positive result on both biologically independent HIV tests on the U.S. market and the likelihood of HIV is only about 15%. As physicians are not trained in statistics and unlikely to communicate risk information appropriately, the secondary risks in terms of suicide, stress, and disparate action may be in the same range as the original problem, HIV.

Legal decisions in the courtroom may be misled by the inability to understand inherently statistical proofs. Having learned that a perfect copy of a perpetrator's DNA occurs in only 1 of a million cases, judges and jurors may infer that a matching DNA yields strong evidence that a defendant is guilty. However, a closer examination of the prosecution environment may reveal that the matching DNA has been found in a computerized search of more than 2 million cases. So the probability that an innocent person has been misidentified may be twice as high (two cases expected) than the probability that the DNA belongs to the original (one case) perpetrator (Lindsey et al., 2003).

Dealing with the quality of statistical, technical, and political information is one task in modern information societies; dealing with the quantity of information overload is another task. Social psychologists have long been interested in selective exposure (Frey, 1986), selective information search (Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Schulz-Hardt, Frey, Lüthgens, & Moscovici, 2000), and related strategies of self-determined reduction of information input. Modern information societies call for the study of external, ecological determinants of information reduction. On one hand, the media and the advertising industry determine the saliency and frequency (density) with which products, brand names, symbols, topics, role models, and threats are imposed on the people. On the other hand, rules of social conduct and cooperation (Clark & Schober, 1992; Grice, 1975) constrain the information that is likely to be published, broadcasted, or uttered in conversation. Thus, for political arguments to be raised in a political campaign, they have to be simplistic rather than difficult and refined. For a message to be distributed in advertising, it has to be financed by a company having vested interests. To be multiplied on the Internet, it has to appear in the upper positions of major search engines. In contrast, information that is "politically incorrect" or that could cause legal sanctions is unlikely to be publicized. In any case, the inevitable reduction of information jointly depends on deliberate selection decisions and externally determined restrictions of information flow.

That the outcome of this eminent ecological process can strongly impact social behavior is vividly demonstrated in the influence of media on aggression (Anderson et al., 2003; Bushman & Anderson, 2001). Other research shows that prudent media consumption can be the key to effective education (Hokanso & Hooper, 2000). Offensive advertising, say, for cellular phones, can change an entire generation's communication style. And the association with major brand names has become a significant part of social identity (Miller & Berry, 1998).

Less Can Be More

Given the tremendous problem of information overload, a major task for adaptive intelligence is optimal selection of information (Oaksford & Chater, 1994), and an intriguing insight from cognitive–ecological research is that less information can be more. The quality and accuracy of decisions and actions can decrease as the amount of input information increases. Based on both developmental data and computer simulation, Elman (1993) showed that the extremely complex task of language learning is facilitated by working-memory limitations and restricted length of input utterances. A similar demonstration was made by Kareev (1995, 2000) for contingency-based decisions, showing that small samples can inform better decisions than large samples. Too much thinking and information search can dissociate decisions from unconscious attitude components (Wilson & Schooler, 1991), causing inconsistency and disruption of behavior. And lexicographic, single-cue strategies can inform more accurate judgments than utilizing many cues for complex inference strategies (Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 2002).

CONCLUSIONS

Aside from those research traditions that can be traced back to a few famous environmentalists—such as Brunswik, Lewin, or Gibson—the roots of a modern cognitive–ecological approach can be found in an old question that has intrigued psychologists since the beginning of the discipline: Is behavior determined by the stimulus environment (Behaviorist position) or by the organism's built-in (and often innate) schemas and propensities (Gestalt position)? Does individual cognition reflect language as a sign system (Benjamin Whorf) or is language but a product of individuals' cognition (Roger Brown)? Is social cognition driven by affordances of the information society, including the mass media (Gaver, 1996) and the immersive virtual world of computers (Alley, 1990; Blascovich, Loomis, & Beall, 2002; Zebrowitz, 2002)? Or, do constraints arise in the cognitive processes of individual persons (Martin, 1990)?

In experimental social psychology, this debate about the external versus internal locus of causation is at the heart of some most prominent paradigms, such as attribution (Jones & McGillis, 1976; Kelley, 1967, 1973), the fundamental attribution bias (Ross, 1977), and the related correspondence bias (Jones, Riggs, & Quattrone, 1979). The basic finding here was that everyday social judgments tend to be biased toward internal causes (i.e., traits and intentions within the person), whereas the external constraints of the environment are often neglected and only considered under auspicious conditions. It appears that behavioral scientists' thinking and theorizing is subject to the same basic restriction and facing the same developmental task, to overcome this fundamental barrier. The task is not merely to admit that external, ecological factors can exert an influence but to develop refined models within which one can understand the joint operation of both classes of determinants. Just as in Inhelder and Piaget's (1958) ontogenetic approach to the development of higher-order (i.e., operational) intelligence is marked by the ability to understand the simultaneous influence of two factors (e.g., that the amount of water in a glass depends on both its height and diameter), current (social) psychology has to learn how to integrate individual and ecological influences within the same models. Up to now, the appearance of such models is conspicuously missing or lacking behind. With a few notable exceptions (e.g., Trope, 1986), the attribution to internal versus external causes is conceived as an either–or problem, a forced choice between either one or the other determinant. The discounting principle (i.e., tendency to allow for only one effective cause) seems to govern not only everyday attributions but scientific explanations as well. This is vividly evident in the popularity of numerous dual-process models in social cognition (Chaiken & Trope, 1999; Smith & DeCoster, 2000), which all tend to assume that either one or the other process is at work at a given time, but not both (for an exception, see Kruglanski & Thompson, 1999).

Given this persisting inclination toward monadic, monocausal explanations, the theoretical discounting of ecological processes would appear to represent the natural consequence of social psychology's focus on cognitive, affective, and motivational processes within the individual. Facing this situation, the present handbook review of research on information ecology had to some extent to be a preview of undone research. The major points conveyed in this review/preview can be summarized in the following basic principles:

1. To understand the structure of psychological processes within the individual, it is first of all necessary to analyze the structure and distribution of the stimulus environment that impinges upon the individual's mind.
2. In particular, the structure of the information ecology entails three aspects, the *density*, the *variance*, and the *redundancy* structure of stimulus distributions.
3. The location and movement of individuals and groups in the environment creates asymmetries and distinct topological relations (i.e., the self–other topology), which are at the heart of many phenomena in social psychology.
4. When individuals search for samples of information for decision making and problem solving, the sampling process is virtually never random but subject to sampling constraints leading to biases and shortcomings.
5. A considerable part of the information that impinges upon human's sensory system—through conversation, literature, mass media, Internet, teaching, or advertising—is self-made, reflecting a genuine interaction of both cognitive and ecological factors.

by Michaela Wänke, Yaakov Kareev, and Henning Plessner on a draft of this chapter are gratefully acknowledged.

REFERENCES

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, *114*, 435–448.

Alley, T. R. (1990). The ecological approach to person perception. *Contemporary Social Psychology, 14*, 153–158.

Alley, T. R., & Clemson, U. (1990). The ecological approach to person perception. *Contemporary Social Psychology*, *14*, 153–158.

Allport, G.W. (1954). *The nature of prejudice*. Oxford, UK: Addison-Wesley.

Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology*, *69*, 518–529.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256–274.

Anderson, C. A. Berkowitz, L., Donnerstein, E., Huesmann, L., Johnson, J. D., Linz, D., et al. (2003). The influence of media violence on youth. *Psychological Science in the Public Interest*, *4*, 81–110.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396–408.

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*, 124–140.

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 334–347). Cambridge, UK: Cambridge University Press.

Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 200–216). Chicago: University of Chicago Press.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169–183). New York: Guilford Press.

Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a face? Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin*, *14*, 23–33.

Bischof, N. (1972). The biological foundations of the incest taboo. *Social Science Information*, *11*, 7–36.

Blanton, H., Axsom, D., McClive, K. P., & Price, S. (2001). Pessimistic bias in comparative evaluations: A case of perceived vulnerability to the effects of negative life events. *Personality and Social Psychology Bulletin*, *27*(12), 1627–1636.

Blascovich, J., Loomis, J., & Beall, A. C. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*, 103–124.

Blau, G., & Katerberg, R. (1982). Agreeing responses set: Statistical nuisance or meaningful personality concept? *Perceptual and Motor Skills*, *54*, 851–857.

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, *1*, 319–322.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, *62*, 645–657.

Bornstein, M. H., & Bradley, R. H. (Ed.). (2003). *Socioeconomic status, parenting, and child development* (Monographs in Parenting Series, pp. 13–27). Mahwah, NJ: Erlbaum.

Brauer, M., Judd, C. M., & Gliner, M. D. (1995). The effects of repeated expressions on attitude polarization during group discussions. *Journal of Personality and Social Psychology*, *68*, 1014–1029.

Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive motivational analysis. *Psychological Bulletin*, *86*, 301–324.

Brewer, M. B. (1993). Social identity, distinctiveness, and in-group homogeneity. *Social Cognition*, *11*, 150–164.

Brewer, M. B. (2003). Optimal distinctiveness, social identity, and the self. In M. R. Leary & J. P. Tangney (Ed.), *Handbook of self and identity* (pp. 480–491). New York: Guilford Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*, 193–217.

Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkeley: University of California Press.

Bushman, B., & Anderson, C. (2001). Media violence and the American public. *American Psychologist*, *56*, 477–489.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford Press.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, *22*, 497–529.

Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 15–47). New York: Russell Sage.

Combs, B., & Slovic, P. (1979). Newspaper coverage of causes of death. *Public Opinion Quarterly*, *56*, 837–843.

Conger, A. J., & Jackson, D. N. (1972). Suppressor variables, prediction, and the interpretation of psychological relationships. *Educational and Psychological Measurement*, *32*, 579–599.

Conway, L. G. III, & Schaller, M. (2007). How communication shapes culture. In K. Fiedler (Ed.), *Frontiers in social psychology: Social communication* (pp. 107–127). New York: Psychology Press.

Cooksey, R. W. (1996). The methodology of social judgement theory. *Thinking and Reasoning*, *2*, 141–173.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187–276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptions for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adaptive mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 85–116). New York: Cambridge University Press.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*, 951–978.

Ditto, P. H., Munro, G. D., Apanovitch, A. M., Scepansky, J. A., & Lockhart, L. K. (2003). Spontaneous skepticism: The interplay of motivation and expectation in responses to favorable and unfavorable medical diagnoses. *Personality and Social Psychology Bulletin*, *29*, 1120–1132.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, *75*, 53–69.

Downing, J. W., Judd, C. M., & Brauer, M. (1992). Effects of repeated expressions on attitude extremity. *Journal of Personality and Social Psychology*, *63*, 17–29.

Duval, T. S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge, UK: Cambridge University Press.

Eibl-Eibesfeldt, I. (1975). *Ethology. The biology of behavior* (2nd ed.) (E. Klinghammer, Trans.). New York: Holt, Rinehart & Winston.

Eiser, J. R., & Stroebe, W. (1972). *Categorization and social judgment*. New York: Academic Press.

Elman, J. L. (1993). Learning and development in neural networks: The importance of staring small. *Cognition*, *48*, 71–99.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review*, *103*, 193–214.

Fiedler, K. (2000a). Beware of samples! A cognitive–ecological sampling approach to judgment biases. *Psychological Review*, *107*, 659–676.

Fiedler, K. (2000b). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology*, *4*, 25–58.

Fiedler, K., & Armbruster, T. (1994). Two halfs may be more than one whole: Category-split effects on frequency illusions. *Journal of Personality and Social Psychology*, *66*, 633–645.

Fiedler, K., Brinkmann, B., Betsch, R., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond baserate neglect and statistical format. *Journal of Experimental Psychology: General*, *129*, 1–20.

Fiedler, K., & Freytag, P. (2004). Pseudocontingencies. *Journal of Personality and Social Psychology*, *87*, 453–467.

Fiedler, K., Kemmelmeier, M., & Freytag, P. (1999). Explaining asymmetric intergroup judgments through differential aggregation: Computer simulations and some new evidence. *European Review of Social Psychology*, *10*, 1–40.

Fiedler, K., Semin, G. R., Finkenauer, C., & Berkel, I. (1995). Actor-observer bias in close relationships: The role of self-knowledge and self-related language. *Personality and Social Psychology Bulletin*, *21*, 525–538.

Fiedler, K., Walther, E., Freytag, P., & Nickel, S. (2003). Inductive reasoning and judgment interference: Experiments on Simpson's paradox. *Personality and Social Psychology Bulletin*, *29*, 14–27.

Fiedler, K., Walther, E., Freytag, P., & Plessner, H. (2002). Judgment biases in a simulated classroom—A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes*, *88*, 527–561.

Fiedler, K., Walther, E., & Nickel, S. (1999). The autoverification of social hypotheses: Stereotyping and the power of sample size. *Journal of Personality and Social Psychology*, *77*, 5–18.

Fischhoff, B. (1975). Hindsight Foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Forest, J. A., & Feldman, R. S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin*, *26*, 118–125.

Förster, J., Friedman, R. S., & Lieberman, N. (2004). Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology*, *87*, 177–189.

Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41–80). New York: Academic Press.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652–670.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Gavanski, I., & Hui, C. (1992). Natural sample spaces and uncertain belief. *Journal of Personality and Social Psychology*, *63*, 766–780.

Gaver, W. W. (1996). Situating action II: Affordances for interaction: The social is material for design. *Ecological Psychology*, *8*, 111–129.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gigerenzer, G. (2004a). Dread risk, September 11, and fatal traffic accidents. *Psychological Science*, *15*, 286–287.

Gigerenzer, G. (2004b). The irrationality paradox: Comment. *Behavioral and Brain Sciences*, *27*, 336–338.

Gigerenzer, G., & Fiedler, K. (2007). *Minds in environments: The potential of an ecological approach to cognition*. Manuscript submitted for publication.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*, 127–171.

Gigerenzer, G., Todd, P. M., & ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment* (pp. 686–715). New York: Cambridge University Press.

Goffman, E. (1990). The presentation of self. In D. Brissett & C. Edgley (Eds.), *Life as theater: A dramaturgical sourcebook. Communication and social order* (2nd ed., pp. 129–139). Hawthorne, NY: Aldine de Gruyter.

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size memory reporting. *Journal of Experimental Psychology: General*, *131*, 73–95.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.

Granhag, P. A., & Stroemwall, L. A. (2001). Deception detection based on repeated interrogations. *Legal and Criminological Psychology*, *6*, 85–101.

Greitemeyer, T., & Schulz-Hardt, S. (2003). Preference-consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. *Journal of Personality and Social Psychology*, *84*, 322–339.

Grice, H. P. (1975). Logic of conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic Press.

Hammond, J. L. (1973). Two sources of error in ecological correlations. *American Sociological Review*, *38, 764–777*.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgement theory. In M. Kaplan & S. Schwartz (Eds.), *Human judgement and decision processes* (pp. 271–312). New York: Academic Press.

Hannan, M. T. (1970). *Problems of aggregation and disaggregation in sociological research*. Chapel Hill: University of North Carolina Institute for Research in Social Science.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539.

Higgins, E. T. (1981). The "communication game": Implications for social cognition and persuasion. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (pp. 343–392). Hillsdale, NJ: Erlbaum.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press.

Higgins, E. T., & Molden, D. C. (2003). How strategies for making judgments and decisions affect cognition: Motivated cognition revisited. In G. V. Bodenhausen & A. J. Lambert (Eds.), *Founda-*

tions of social cognition: A festschrift in honor of Robert Wyer, Jr. (pp. 211–236). Mahwah, NJ: Erlbaum.

Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, *58*, 197–208.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, *290*, 2261–2262.

Hokanso, B., & Hooper, S. (2000). Computers as cognitive media: Examining the potential of computers in education. *Computers in Human Behavior*, *16*, 537–552.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, *133*, 23–30.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on "selective exposure to information." *Journal of Personality and Social Psychology*, *80*, 557–571.

Jones, E. E., & McGillis, D. (1976). Correspondent inferences and the attribution cube: A comparative reappraisal. *New directions in attribution research* (Vol. 1, pp. 389–420). Hillsdale, NJ: Erlbaum.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Jones, E. E., Riggs, J. M., & Quattrone, G. (1979). Observer bias in the attitude attribution paradigm: Effect of time and information order. *Journal of Personality and Social Psychology*, *37*, 1230–1238.

Judd, C. M., & Park, B. (1988). Outgroup-homogeneity: Judgments of variability at the individual and the group levels. *Journal of Personality and Social Psychology*, *54*, 778–788.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter—Guided selection of almanac items. *Organizational Behavior and Human Decision Processes*, *57*, 226–246.

Juslin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: A sampling model of confidence in sensory discrimination. *Psychological Review*, *104*, 344–366.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*, 384–396.

Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, *93*, 429–445.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–453.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior 1967: Aversive stimulation* (pp. 9–31). Coral Gables, FL: University of Miami Press.

Kanouse, D. E., & Hansen, L. R., Jr. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 47–62). Morristown, NJ: General Learning Press.

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, *56*, 263–269.

Kareev, Y. (2000). Seven (indeed, plus minus two) and the detection of correlations. *Psychological Review*, *107*, 397–402.

Karpinski, A. T., & von Hippel, W. (1996). The role of the linguistic intergroup bias in expectancy maintenance. *Social Cognition*, *14*, 141–163.

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin*, *26*, 594–604.

Kashima, Y. (2001). Culture and social cognition: Toward a social psychology of cultural dynamics. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 325–360). New York: Oxford University Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–238). Lincoln: University of Nebraska Press.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107–128.

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, *62*, 88–97.

Kimmel, H. D. (1970). Essential events in the acquisition of classical conditioning *Conditional Reflex*, *5*, 156–164.

Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, *73*, 885–901.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Klayman, J., Soll, J. B., González-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, *79*, 216–247.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110*, 499–519.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118.

Krueger, J., & Gilovich, T. (2004). Actions, intentions, and self-assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin*, *30*, 328–339.

Kruger, J. (1999). Lake Wobegon be gone!: The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77(2), 221–232.

Krueger, J., & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology, 40*, 586–598.

Kruglanski, A. W., Schwartz, J., Maides, S., & Hamel, I. (1978). Covariation, discounting and augmentation: Toward a clarification of attribution principles. *Journal of Personality*, *46*, 176–199.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Kühnen, U., Hannover, B., & Schubert, B. (2001). The Semantic-Procedural Interface Model of the self: The role of self-knowledge for context-dependent versus context-independent modes of thinking. *Journal of Personality and Social Psychology*, *80*, 397–409.

Kukla, A. (1993). The structure of self-fulfilling and self-negating prophecies. *Theory and Psychology*, *4*, 5–33.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Lamm, H., & Myers, D. E. (1978). Group-induced polarization of attitudes and behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 145–195). New York: Academic Press.

Larson, J. R., Jr., Foster-Fishman, P. G., & Keys, C. B. (1994). Discussion of shared and unshared information in decision-making groups. *Journal of Personality and Social Psychology*, *67*, 446–461.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*, 343–356.

Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, *88*, 438–453.

Lewin, K. (1951). *Field theory in social science*. New York: Harper.

Lewontin, R. C. (1979). Sociobiology as an adaptionist program. *Behavioral Science*, *24*, 5–14.

Liberman, N., Sagristano, M., & Trope, Y. (2002). The effect of temporal distance on level of construal. *Journal of Experimental Social Psychology*, *38*, 523–535.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*, 5–18.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of subjective probabilities: The state of the art up to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

Lindsey, S., Hertwig, R., & Gigerenzer, G. (2003). Communicating statistical DANN evidence. *Jurimetrics*, *43*, 147–163.

Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: Empirical evidence and a computer simulation. *Journal of Personality and Social Psychology*, *57*, 165–188.

Linville, P. W., Fischer, G. W., & Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. *Journal of Personality and Social Psychology*, *70*, 421–436.

Loftus, E. F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

Lorenz, K. Z. (1978). *The foundations of ethology*. New York: Springer-Verlag.

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication?: The influence of stereotype sharedness? *Journal of Personality and Social Psychology*, *85*, 989–1005.

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. *Advances in Experimental Social Psychology*, *31*, 79–121.

Maass, A., Milesi, A., Zabbini, S., & Stahlberg, D. (1995). Linguistic intergroup bias: Differential expectancies or in-group protection. *Journal of Personality and Social Psychology*, *68*, 116–126.

Malle, B. F., & Knobe, J. (1997). Which behaviors do people explain? A basic actor–observer asymmetry. *Journal of Personality and Social Psychology*, *72*, 288–304.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Martin, L. L. (1990). Comments on the presentations of Alley, Baron, and Berry: What do their data really afford us? *Contemporary Social Psychology*, *14*, 159–163.

McArthur, L. Z. (1981). What grabs you?: The role of attention in impression formation and causal attribution. In E. T. Higgings, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1, pp. 201–246). Hillsdale, NJ: Erlbaum.

McClure, J. (1998). Discounting of causes of behavior: Are two reasons better than one? *Journal of Personality and Social Psychology*, *74*, 7–20.

Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422–471.

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*, 961–978.

Miller, N. E. (1944). Experimental studies of conflict. In McV. Hunt (Ed.), *Personality and the behavior disorders* (pp. 431–456). New York: Ronald Press.

Miller, S., & Berry, L. (1998). Brand salience versus brand image: Two theories of advertising effectiveness. *Journal of Advertising Research*, *38*, 77–82.

Miyamoto, Y., & Kitayama, S. (2002). Cultural variation in correspondence bias: The critical role of attitude diagnosticity of socially constrained behavior. *Journal of Personality and Social Psychology*, *83*, 1239–1248.

Mojzisch, A., & Schulz-Hardt, S. (2005). Information sampling in group decision making: Sampling biases and their consequences. In K. Fiedler & P. Juslin (Eds.), *Information sampling as a key to understanding adaptive cognition* (pp. 299–326). New York: Cambridge University Press.

Moore, D. A. (in press). Not so above average after all: When people believe they are worse than average and its implications for theories of bias in social comparison. *Organizational Behavior and Human Decision Processes*.

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of Discounting in causal attribution. *Psychological Review*, *102*, 331–355.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, *12*, 125–135.

Mullen, B., & Hu, L. (1989). Perceptions of ingroup and outgroup variability: A meta-analytic integration. *Basic and Applied Social Psychology*, *10*, 233–252.

Mullen, B., & Riordan, C. A. (1988). Self-serving attributions for performance in naturalistic settings: A meta-analytic review. *Journal of Applied Social Psychology*, *18*, 3–22.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.

Myers, D. G. (1978). Polarization effects of social comparison. *Journal of Experimental Social Psychology*, *14*, 554–563.

Neisser, U. (1985). The role of theory in the ecological study of memory: Comment on Bruce. *Journal of Experimental Psychology: General*, *114*, 272–276.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin*, *88*, 622–637.

Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, *23*, 73–130.

Nisbett, R. E., Peng, K. P., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*, 291–310.

Nothdurft, H. C. (1993). Saliency effects across dimensions in visual search. *Vision Research*, *33*, 839–844.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, *97*, 362–376.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608–631.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776–783.

Parducci, A. (1968). The relativism of absolute judgment. *Scientific American*, *19*, 84–90.

Park, B., & Rothbart, M. (1982). Perception of outgroup homogeneity and levels of social categorization: Memory for the subordinate attributes of ingroup and outgroup members. *Journal of Personality and Social Psychology*, *42*, 1051–1068.

Pettigrew, T. F. (1986). The intergroup contact hypothesis reconsidered. In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 169–195). Oxford, UK: Basil Blackwell.

Price, P. C. (2001). A group-size effect on personal risk judgments: Implications for unrealistic optimism. *Memory and Cognition*, *29*, 578–586.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*, 781–799.

Pruitt, D. G., & Teger, A. I. (1969). The risky shift in group betting. *Journal of Experimental Social Psychology*, *5*, 115–126.

Quattrone, G. A. (1986). On the perception of a group's variability. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 25–48). Chicago: Nelson-Hall.

Rassin, E. (2000). Criteria based content analysis: The less scientific road to truth. *Expert Evidence*, *7*, 265–278.

Ray, J. J. (1983). Reviving the problem of acquiescence response bias. *Journal of Social Psychology*, *121*, 81–96.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rhee, E., Uleman, J. S., Lee, H. K., & Roman, R. J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. *Journal of Personality and Social Psychology*, *69*, 142–152.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*, 351–357.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, *3*, 377–386.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, *37*, 322–336.

Sagristano, M. D., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: Odds now, money later. *Journal of Experimental Psychology: General*, *131*, 364–376.

Sande, G. N., Goethals, G. R., & Radloff, C. E. (1988). Perceiving one's own traits and others': The multifaceted self. *Journal of Personality and Social Psychology*, *54*, 13–20.

Schaller, M. (1992a). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of Personality and Social Psychology*, *63*, 61–74.

Schaller, M. (1992b). Sample size, aggregation, and statistical reasoning in social inference. *Journal of Experimental Social Psychology*, *28*, 65–85.

Schlenker, B. (1985). *The self and social life* (B. R. Schlenker, Ed.). New York: McGraw-Hill.

Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, *78*, 655–669.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications.* Mahwah, NJ: Erlbaum.

Sedlmeier, P., Hertwig, R., & Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 754–770.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558–568.

Shavitt, S., Sanbonmatsu, D. M., Smittipatana, S., & Posavac, S. S. (1999). Broadening the conditions for illusory correlation formation: Implications for judging minority groups. *Basic and Applied Social Psychology*, *21*, 263–279.

Shimp, T. A., Hyatt, T. M., & Snyder, D. J. (1991). A critical appraisal of demand artifacts in consumer research. *Journal of Consumer Research*, *18*, 273–283.

Simon, B. (1992). The perception of ingroup and outgroup homogeneity: Reintroducing the intergroup context. *European Review of Social Psychology*, *3*, 1–30.

Simon, H. A. (1956). Rational choice and the structure of environments. *Psychological Review*, *63*, 129–138.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B*, *13*, 238–241.

Skinner, B. F. (1984). The operational analysis of psychological terms. *Behavioral and Brain Sciences*, *7*, 547–581.

Slatin, G. T. (1969). Ecological analysis of delinquency: aggregation effects. *American Sociological Review*, *34*, 894–907.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*, 364–371.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131.

Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 247–305). New York: Academic Press.

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing strategies in social interaction. *Journal of Personality and Social Psychology*, *36*, 1202–1212.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of stereotypes. *Journal of Personality and Social Psychology*, *36*, 941–950.

Soll, J. B. (1996). Determinants of overconfidence and miscalibration: The roles of random error and ecological structure. *Organizational Behavior and Human Decision Processes*, *65*, 117–137.

Spellman, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, *7*, 337–342.

Spellman, B. A., Price, C. M., & Logan, J. M. (2001). How two causes are different from one: The use of (un)conditional information in Simpson's Paradox. *Memory and Cognition*, *29*, 193–208.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*, 1467–1478.

Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, *27*, 165–175.

Swann, W. B., Giuliano, T., & Wegner, D. M. (1982). Where leading questions can lead: The power of conjecture in social interaction. *Journal of Personality and Social Psychology*, *42*, 1025–1035.

Swets, J., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*(Whole No. 1).

Tajfel, H. (1957). Value and the perceptual judgment of magnitude. *Psychological Review*, *64*, 192–204.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.

Tarif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context. *Child Development*, *70*, 620–635.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*, 778–793.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 289–338). New York: Academic Press.

Tesser, A., & Leone, C. (1977). Cognitive schemas and thought as determinants of attitude change. *Journal of Experimental Social Psychology*, *13*, 340–356.

Thorndyke, E. L. (1916). Notes on practice, improvability, and the curve of work. *American Journal of Psychology*, *27*, 550–565.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). New York: Oxford University Press.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.

Trope, Y. (1986). Identification and inference processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology*, *79*, 876–889.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–421.

Trope, Y., & Thompson, E. P. (1997). Looking for truth in all the

wrong places? Asymmetric search of individuating information about stereotyped group members. *Journal of Personality and Social Psychology*, *73*, 229–241.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Tversky, A., & Koehler, D. J. (1994). Support theory: A non-extensional representation of subjective probability. *Psychological Review*, *101*, 547–567.

Valenti, S. S., Hofstra, U., & Good, J. M. (1991). Social affordances and interaction: I. Introduction. *Ecological Psychology*, *3*, 77–98.

Vallacher, R. R., & Wegner, D. M. (1985). *A theory of action identification*. Hillsdale, NJ: Erlbaum.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology*, *27*, 177–254.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and its implications for professional practice*. Chichester, UK: Wiley.

Waldmann, M., & Hagmayer, Y. (1995). Causal paradox: When a cause simultaneously produces and prevent an effect. *Proceedings of the 17th annual conference of the Cognitive Science Society*, Hillsdale, NJ.

Wason, P. C. (1996). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). London: Penguin.

Watson, D. (1982). The actor and the observer: How are their perceptions of causality divergent? *Psychological Bulletin*, *92*, 682–700.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk-sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, *111*, 430–445.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*, 806–820.

Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, *56*, 161–169.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*, 1115–1125.

Wicklund, R. A., & Braun, O. L. (1987). Incompetence and the concern with human categories. *Journal of Personality and Social Psychology*, *53*, 373–382.

Wigboldus, D. H. J., Semin, G. R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, *78*, 5–18.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*, 181–192.

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, *75*, 1411–1423.

Winman, A., Juslin, P., & Björkman, M. (1998). The confidence-hindsight mirror effect in judgment: An accuracy-assessment model for the knew-it-all-along phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 415–431.

Wittenbaum, G. M., Hubbell, A. P., & Zuckerman, C. (1999). Mutual enhancement: Toward an understanding of the collective preference for shared information. *Journal of Personality and Social Psychology*, *77*, 967–978.

Wolf, S. (1987). Majority and minority influence: A social impact analysis. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium on Personality and Social Psychology* (Vol. 5, pp. 207–235). Hillsdale, NJ: Erlbaum.

Wolf, S., & Latané, B. (1983). Majority and minority influences on restaurant preferences *Journal of Personality and Social Psychology*, *45*, 282–292.

Wood, J. (1989). Theory and research concerning social comparisons and personal attributes. *Psychological Bulletin*, *106*, 231–248.

Wood, J. V., Michela, J. L., & Giordano, C. (2000). Downward comparison in everyday life: Reconciling self-enhancement models with the mood-cognition priming model. *Journal of Personality and Social Psychology*, *79*, 563–579.

Zajonc, R. B. (1960). The process of cognitive tuning and communication. *Journal of Abnormal and Social Psychology*, *61*, 159–167.

Zajonc, R. B., Markus, H., & Markus, G. B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology*, *37*, 1325–1341.

Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist*, *52*, 685–699.

Zebrowitz, L. A. (2002). The affordances of immersive virtual environmental technology for studying social affordances. *Psychological Inquiry*, *13*, 143–145.

Zebrowitz, L. A., Fellous, J.-M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*, 194–215.

Zebrowitz-McArthur, L. A., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, *90*, 215–238.

Zuckerman, M. (1979). Attribution of success and failure revisited, or: the motivational bias is alive and well in attribution theory. *Journal of Personality*, *47*, 245–287.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, *14*, 1–57.

Zuckerman, M., Knee, C. R., Hodgins, H. S., & Miyake, K. (1995). Hypothesis confirmation: The joint effect of positive test strategy and acquiescence response set. *Journal of Personality and Social Psychology*, *68*, 52–60.

CHAPTER 9

Knowledge Activation

Jens Förster
Nira Liberman

In 1977, Higgins, Rholes, and Jones invited participants to an experiment that consisted of two apparently unrelated studies. The first study was a Stroop color-naming task, for part of which participants had to memorize words, supposedly to make the task more demanding; for some participants these were positive words (adventurous, self-confident, independent, persistent), whereas for other participants these were negative words (reckless, conceited, aloof, stubborn). The second study was an impression formation task in which participants read a description of Donald who performed a series of ambiguous behaviors that could be regarded as adventurous or reckless (e.g., Donald thought about crossing the Atlantic in a sailboat). Participants then wrote a free description of the target person and indicated how much they liked him on a rating scale. Results indicated that Donald was rated as more adventurous and less reckless by the positive priming group than by the negative priming group.

A typical priming experiment has two phases: (1) a priming phase in which participants are exposed to some information and (2) an ostensibly unrelated perception, memory, or judgment task in which participants' responses to a target stimulus are examined. The *afferent stage* of a priming experiment consists of a stimulus presentation (e.g., reckless or adventurous) that leads to activation of associated memory structures. The *efferent stage* of a priming experiment consists of activation of a memory structure in the process of responding to related target stimuli (Fiedler, 2003).

From the perspective of the participant, there is no relation between the two phases (the procedure described above is often referred to as the *unrelated task paradigm*). Thus, any effect of the first stage on the second stage is not noticed by the participant, ruling out conversational effects (e.g., thinking that the information of the first stage is useful) or motivational effects (e.g., wanting to conform to a specific behavior pattern).

It is argued that the first stage enhances the accessibility of the primed construct and that constructs with higher accessibility are more likely to be used than those with lower accessibility. In the study by Higgins and colleagues (1977), the behavior "thinking about crossing the Atlantic in a sailboat" could be perceived as both adventurous and reckless. A higher accessibility of one of these constructs leads to the perception of the target in terms of this construct.

In a typical priming task, people are not aware that they were affected by the first task. If directly asked how they formed their judgments, they would most likely reply that they based their judgments on the stimulus itself—they would say that the target person was, indeed, involved in reckless or adventurous activities (i.e., the judgment is "about" the target's behavior and not about the priming event). In logical terms, the influence of the prime is not justifiable—nobody would maintain that the target should be perceived differently because of a previous task of color naming. Yet, these effects occur reliably and have been replicated and used in many social psychological experiments.

In this chapter, we first propose some working definitions of priming effects as distinct from other effects of prior information processing on further target perception. Then, in the second part we review the main findings within the vast field of social psychological research on knowledge accessibility, including perceptual readiness, retrieval from memory, disambiguation, the effects of priming related to behavior and motivation, procedural priming, affective and evaluative priming, and chronic accessibility. The third part examines some theoretical principles that emerge from this review—including factors affecting decay rates and intensity of priming, how accessibility from different sources combines, and how applicability and accessibility compensate for each other. The fourth part examines metacognitive processes related to priming, and the fifth part examines the possible function of different priming effects—what do they serve and what would moderate such effects. The final part examines some applications of the principles of accessibility in social psychology. Specifically, we discuss using affective priming to measure attitudes, to assess motivations and personal concerns, and to understand processes of thought suppression. We do not review models of accessibility as these have been examined in detail elsewhere (see, e.g., Higgins, 1996; Huber, Shiffrin, Lyle, & Ruys, 2001; Wyer, 2004; Wyer & Radvansky, 1999).

DEFINITIONS

Let us begin with introducing a number of terms, starting with an attempt to define the effects of accessibility. As we will see, there are difficulties in many of these definitions; however, we believe that these difficulties, some of which originate from the fuzzy boundaries of these theoretical constructs, should not detract from their usefulness.

What Are the Effects of Accessibility?

First, we would like to delineate the domain of our discussion—which effects would be termed "effects of accessibility" and which effects will be left outside this definition. Generally, *accessibility is a temporary state that is produced by prior processing of a stimulus and thus activates knowledge*, be it semantic, procedural, experiential, or any other form (e.g., Tulving, 1983). By frequently activating certain knowledge structures accessibility can also become chronic. Generally, we are talking about the effect of prior processing on subsequent processing of a target stimulus. Not all such effects, however, should be labeled "accessibility." For example, priming (e.g., the experiment described at the outset of this chapter) is an accessibility effect, but persuasion per se is not an effect of accessibility, or at least not only an effect of accessibility. So, what is the theoretical principle that defines accessibility?

We would like to suggest that accessibility effects occur when effects of prior processing happen regardless of people's beliefs about the relevance of the prime to the

target, and regardless of their beliefs about the suitability of its influence on processing the target. For example, in subliminal priming, when people are unaware of processing the prime, they may not consider the relevance of the prime to the target. This is similar to the unrelated tasks procedure in which people are aware of processing the prime but do not consider it relevant to the second phase when the target is presented.[1] By being independent of the questions of relevance and suitability of influence, the effects of knowledge accessibility are insensitive to the propositional content of the prime (e.g., to its truth value, see Fiedler, 2003). Notably, unlike in processes of attitude change via persuasion, increased accessibility from priming would occur with both negated and affirmed primes (Fiedler, 2003; Mayo, Schul, & Burnstein, 2004; Strack & Deutsch, 2004), or regardless of the credibility of its source (Schul, Mayo, & Burnstein, 2004).[2]

Sometimes, effects of accessibility occur in conjunction with other effects. For example, people who watch a commercial may unknowingly be affected by it. They may rightly acknowledge the relevance of the entire commercial to their behavior and resist its effects (e.g., "They are trying to persuade me to buy diet toothpaste, but I do not need it and won't yield to that attempt.") and at the same time be affected by features of the commercial that they would not judge relevant for the focal judgment (e.g., they might end up thinking that the toothpaste is healthier because the commercial paired it with health-related stimuli). In this example, the part of the process that occurs regardless of the person's beliefs in its relevance and suitability (the effect of irrelevant health-related stimuli) would be termed an "effect of accessibility." This example illustrates the fact that sometimes accessibility effects may interact in complex ways with more deliberate, belief-based processes (e.g., persuasion). We return to this point many times when we discuss in more detail the various effects of accessibility and again when we discuss metacognitive processes.

By our definition, while some types of classical conditioning would not count as effects of accessibility, other types of conditioning may be viewed as effects of accessibility. Namely, when the unconditioned stimulus is diagnostic with respect to the conditioned stimulus (e.g., when the bell signifies that food is about to be served), then conditioning is not analogous to effects of accessibility because the association is based on a true informational, predictive value of the prime. However, if toothpaste becomes associated with health because it is shown on a commercial together with fresh fruits the association is formed irrespective of the person's belief in the relation between toothpaste and health and thus may be considered an effect of accessibility (see Walther, Nagengast, & Trasselli, 2005).

We should emphasize that belief-driven (conscious or unconscious) processes may alter the effects of accessibility. For example, if I believe that my judgments could have been affected by prior processing, I might want to counteract them. Such correction is not, by itself, an effect of accessibility, but it is related to accessibility and metacognitive knowledge. This is discussed further in the fourth part of the chapter.

There are other definitions of accessibility effects that, we believe, are not useful. One equates effects of accessibility with automatic processes and distinguishes them from controlled processes. Any process, including belief-based propositional computations, may be automated and become efficient, unaware, and unintended. For example, people may automatically yield to messages from an authority as a result of a reason-based process of learning and believing that doing so is useful. This will produce an automatic effect. One could argue that such processes had to be performed repeatedly in a nonautomatic manner in order to achieve automaticity, whereas effects of accessibility happen automatically (i.e., with no intention and with no awareness) with no prior learning. The latter claim, however, cannot be endorsed with certainty because it is possible that some learning process did, in fact, bring about at least some automatic effects of accessibility. For example, it is possible that people learned over time that stimuli in close temporal or spatial proximity are related to each other and therefore began using co-occurring stimuli to disambiguate each other. Although accessibility effects typically happen outside awareness, they are not unique in that other processes also may require little effort and occur with no awareness or intention. The reverse is also true—accessibility need not happen outside awareness. A person could be aware of an influence of a prime on subsequent processing but still have an effect of its accessibility. Part of the problem here is that automaticity itself has more than one property, with awareness being only one of them (see Bargh, 1994b). Second, it has also been proposed that accessibility does not alter knowledge structures (i.e., does not change availability) but, rather, alters only momentary activation potential, whereas other processes, such as persuasion, introduce a more long-lasting change in long-term memory. This distinction is problematic because momentary changes may amount over time to new associations and create new meanings (e.g., form associations between stimuli, such as between toothpaste and health). As we see later, procedural priming and chronic accessibility, in particular, involve changes in long-term memory (see Wyer, 2004). We thus propose to define effects of accessibility as effects that are insensitive to beliefs (but are, as we will see next, sensitive to associations). We are aware of the fact that the boundaries of this definition may be unclear. However, there are clear examples of effects that are produced by accessibility of knowledge and others that are not. Thus, the effect of priming on judging a target is a clear example of a knowledge accessibility effect, whereas when one is persuaded to try a product it is a clear example of an effect that is not driven by accessibility of knowledge. In the latter case it is not the mere accessibility of the idea that leads to trying the product but rather that some more complex processes are involved.

There are effects that are more difficult to classify one way or another, such as some types of conditioning, or procedural priming. Nevertheless, the existence of twilight does not mean that people are unable to distinguish between day and night, and thus we proceed to explore the category of accessibility with a number of further distinctions and terms that are used throughout this chapter. Generally, our terminology is consistent with that of Higgins (1996) and with Decoster and Claypool (2004).

- *Construct use* is a response in terms of a construct. For example, rating an ambiguous target as aggressive involves using the aggressive construct.
- *Availability* is presence in memory. People can only use a construct and be primed with a construct if it is available. For example, if one never associated "creativity" with the category of professors, priming "professor" would not bring "creativity" to mind. Availability is a necessary condition for accessibility.
- *Accessibility* is the activation potential of available knowledge (Higgins, 1996). Knowledge activation (e.g., by priming) increases its potential for further activation—that is, it increases accessibility (Bruner, 1957).
- *Applicability.* Accessibility is not sufficient to produce priming effects. Rather, accessible constructs affect perception and judgment only if they are applicable to the target (Higgins, 1989). For example, when primed with "dependency," participants found a female target but not a male target to be more dependent than neutrally primed control participants (Banaji, Hardin, & Rothman, 1993). The reverse was found for priming aggression. Notably, the target stimuli, which were behavioral descriptions of individuals, were the same for male and female targets. Yet, the primed construct was used more when the prime matched the social category ("dependent" for a female target, "aggressive" for a male target; for a recent meta-analysis of the effect of applicability on assimilative priming effects, see Decoster & Claypool, 2004, for an application to survey research see Todorov, 2000). Applicability should be distinguished from *consistency* (Higgins, 1996). Inconsistent constructs are applicable to a target that has an opposite meaning (e.g., the prime "dishonest" is applicable to an honest person, although inconsistent with it) more so than completely unrelated constructs (e.g., the prime "stupid" to an honest person). Thus, inconsistency naturally implies some applicability.
- *Ambiguity and vagueness.* A stimulus description is *ambiguous* when at least two alternative constructs are equally applicable to it (e.g., a target may be described as both adventurous and reckless), whereas a behavioral description is *vague* when no construct has more than weak applicability to it (e.g., when a behavior only weakly implies intelligence, e.g., "Brett read a newspaper," Higgins & Brendl, 1995). Perceptions and evaluations of ambiguous and vague stimuli, more than those of clear, unequivocal stimuli, are influenced by priming (and, in fact, by any manipulation, such as persuasion, conditioning, etc.).
- *Judged usability* is the perceived appropriateness of applying some knowledge to a stimulus (Higgins, 1989, 1996). Information that is both accessible and applicable may not be used if it is perceived as irrelevant or inappropriate. For example, when a person thinks a racist thought, such as "African Americans are hostile," is untrue or politically incorrect she might refrain from judging an African American person as hostile, even if the

construct of hostility is highly accessible to her (Devine, 1989). Judged usability is often viewed as a conscious process, but it is entirely possible that unconscious judgments of usability also occur; judgments may become automated, as any overlearned process can be. Regardless of whether it is conscious or not, judged usability is a metacognitive process and we discuss it and related metacognitive processes in more detail in the fourth part of the chapter.

EFFECTS OF ACCESSIBILITY

We turn now to review the different effects of accessibility—effects on perceptual readiness, memory, disambiguation, priming related to motivation and goals, procedural priming, affective and evaluative priming, and chronic accessibility.

Perceptual Readiness

In a classic demonstration of semantic priming, Meyer and Schvaneveldt (1971) measured the time it took participants to decide if a letter string was a word or not. This task, called the *lexical decision task*, often serves to measure perceptual readiness as produced by accessibility. It is assumed that a decision about a particular word would be facilitated by accessibility, and thus, lexical decisions should be shorter for more accessible constructs. The authors found facilitation of lexical decision when words were preceded by semantically related primes. For example, when the target stimulus was the word "building," lexical decision time was reduced if the semantically related word "house" preceded it. It is beyond the scope of this chapter to review the vast literature on semantic priming in cognitive psychology. Rather, we present only a few core findings that we found useful for a better understanding of related findings in social psychology.

Perceptual Readiness for Visual Stimuli

Conceptual priming may facilitate perception of visual features. Zwaan, Stanfield, and Yaxley (2002) presented participants with sentences such as "The ranger saw the eagle in the nest," or "The ranger saw an eagle in the sky," and then asked participants to view pictures and decide as quickly as possible whether the object in the picture was mentioned in the sentence. The pictures contained either features that matched the implied shape of the object or not (e.g., stretched wings or folded wings). Decisions were faster when there was a match between the sentence and the picture. These results suggest that sentences trigger perceptual simulations and not only activate semantic knowledge (see also Wyer, Adaval, & Colcombe, 2002).

Decay Rates of Priming Effects

In many of the studies in cognitive psychology, effects of semantic priming were found to decay extremely rapidly. For example, the effect of semantic priming on lexical decisions disappears if the target is presented a few seconds after the prime, or if another word intervenes between the prime and the target. Slower decay rates, however, appear with deeper processing of the prime and the target. For example, Becker, Moscovitch, Behrmann, and Joordens (1997) found that a decision on animosity (i.e., "Is it a living thing?") for both the prime and the target created a facilitation that lasted for much longer than a lexical decision (i.e., "Is it a word?"). With 10 intervening words the effects of priming were still evident with deep semantic processing but not with lexical decisions. The authors suggest that stronger effects after a delay (i.e., slower decay) stem from a larger degree of incremental learning that is characteristic of deeper processing. In our view, deeper processing of the target makes the task more similar to disambiguation (see below) than to pure perceptual readiness. Thus, we would expect slower decay and stronger effects for disambiguation than for perceptual readiness tasks.

Directional Associations

Concepts can be activated by interrelated networks, which are formed if two concepts are frequently activated together (Hebb, 1948). For example, Mussweiler and Förster (2000) reasoned that the concepts of "sex" and "aggression" became interrelated over socialization and that the strength of this association may depend on the direction (e.g., sex → aggression vs. aggression → sex). Specifically, because sexual situations often involve aggressive components (see, e.g., Koss, Gidyez, & Wisniewski, 1987), an automatic sex–aggression link could have developed. Consequently, activating concepts associated with "sex" should also activate concepts associated with "aggression." By contrast, many aggressive situations are free of sexual components and therefore "sex" would not be contiguous with "aggression," meaning that activation of aggression concepts should not activate sex concepts. A study using subliminal priming demonstrated this effect with a lexical decision task. When primed with mild sex concepts (e.g., the word "wet") participants were faster in identifying aggression-related words (e.g., "brutal"). However, this was not true for sex-related words after priming of aggression-related words, a finding that is consistent with the idea of a unidirectional link.

Inhibition of Competing Constructs

Primes facilitate related constructs, but they also inhibit competing constructs (see Förster & Liberman, 2005, for a review). For example, Ratcliff and McCoon (1996) presented participants with pictures of mundane objects (e.g., a lightbulb) and recorded naming latencies after being primed 1 week earlier with the same object or a graphically similar but semantically different object (e.g., a balloon). A no-priming control group was also included. The authors found that whereas semantically similar primes produced facilitation, graphically similar primes produced inhibition by increasing the response time relative to the control group.

Similarly, in linguistic processing, it has been found that homonyms initially activate all the semantic meanings (e.g., the word "bank" activates both the meaning of a financial institution and the meaning of a riverside). However, once a context is determined, the context-irrelevant meanings become inhibited (Swinney, Prather, & Love, 2000).

In Neely's (1977) classic study, participants performed lexical decisions in which the target words were preceded by signal words. In one condition, the signal "body" appeared consistently with targets that signified building parts (e.g., door) and, therefore, was reasoned to semantically prime body parts and simultaneously induce an "active set" for building parts. The study found that at longer delays (1,000 msec) between the prime and the target (stimulus onset asynchrony), the prime inhibited decisions about semantic associates (e.g., body parts) and facilitated decisions related to the active set (e.g., building parts). However, at shorter delays (250 msec), it facilitated decisions about the semantic associates but not about the active set. This has been interpreted to mean that novel associations have longer onset times, whereas more habitual associations are faster to exert influence. We should expect, therefore, that practicing an association would shorten onset times of its influence as a prime.

Macrae, Bodenhausen, and Milne (1995) exposed participants to targets that belonged to multiple stereotyped groups (e.g., an Asian woman). Beforehand, participants were subliminally primed with one of the category labels (women, Chinese). Using a lexical decision task, the study showed that such priming led to inhibition of concepts that were related to the other, nonprimed category (see also Fiedler & Schenck, 2001). In sum, it seems that primes have onset times that get shorter with practice. It also seems that priming causes inhibition of prime-associated constructs that may impede performance of the task at hand and may be delayed.

Retrieval from Memory

Cued Recall

After participants memorize a list of words (e.g., a list of objects from a number of categories), exposure to semantic associates of the learned items (e.g., the name of some of the categories) typically improves recall, a phenomenon that has been termed "cued recall." Thus, retrieval from long-term memory is assisted by priming, namely, by exposure to a semantically related item or a cue (see, e.g., Santa & Lamwers, 1974; Shimamura & Squire, 1984).

Part-Set Cuing Inhibition and Retrieval-Induced Forgetting

Sometimes retrieval cues impede rather than assist memory, a phenomenon that has been termed "part-set cuing inhibition" (for reviews of the empirical findings, see Anderson & Neely, 1996; Nickerson, 1984; Roediger & Neely, 1982). The paradigm typically requires partici-

pants to memorize a list of words from a small number of semantic categories (e.g., 40 words comprised of colors, trees, fish, and musical instruments). At recall, participants are presented with a different number of words from each category (e.g., one musical instrument, four colors, and four trees) as cues. The classic finding is that participants recall fewer of the remaining items in a given category when *more* words from that category are presented as cues (see, e.g., Rundus, 1973; Slamecka, 1968; Watkins, 1975). Anderson and colleagues (Anderson & Neely, 1996; Anderson & Spellman, 1995) have suggested that selective attention toward the rehearsed items inhibits accessibility of the other items.

Retrieval-induced forgetting refers to a somewhat related phenomenon. In these experiments, during the study phase, participants study several categories, each composed of several exemplars in a category-exemplar format (e.g., fruit-orange). After the study phase, participants engage in directed "retrieval practice" for half of the items from half of the studied categories. In this phase, they complete category-plus-exemplar stem cue tests (e.g., fruit-or___). After a retention interval, a final and unexpected category-cued recall test is administered in which participants are cued with each category name and asked to reproduce any exemplar of that category that they remember from any point in the experiment. It is typically found that performance on practiced items improves relative to the baseline condition (i.e., items from unpracticed categories), whereas performance on the unpracticed items from practiced categories falls below this baseline. Anderson and Bjork (1994) explained these effects by suggesting that during the retrieval practice phase the practiced items compete with the unpracticed items from the same category and thus inhibit them (see also Anderson & Neely, 1996; Anderson & Spellman, 1995).

It is worth emphasizing that semantically related cues typically enhance memory, but as demonstrated by the phenomena of part-set cuing inhibition and retrieval-induced forgetting, they may also reduce memory. Specifically, if the cue and the target compete with each other as potentially relevant responses, then the nonchosen response is inhibited, becoming less accessible than before.

Disambiguation

Perhaps of particular relevance to social psychology are studies on accessibility effects on perception of ambiguous social stimuli, as social stimuli are often ambiguous and evaluation and labeling constitute important parts of social perception. The example at the beginning of this chapter is a classic demonstration of priming effects on disambiguation of a social stimulus. Stemming from that research, a great variety of priming effects on disambiguation have been documented with varying priming procedures, contents, and social targets. For example, unscrambling sentences related to hostility leads to perceiving a vaguely aggressive behavior as more aggressive than unscrambling sentences unrelated to hostility (Wyer & Srull, 1989). We do not list all the variations in

content but, rather, attempt to mention only some extensions of the original paradigm that we believe are of particular theoretical or practical importance.

Priming of Stereotypes

A classic study by Devine (1989) showed that subliminally priming associations of the category of African Americans (e.g., Negroes, Blacks, and lazy) influenced hostility ratings of an ambiguously aggressive target. Importantly, individual differences in racism, as measured by the modern racism scale (McConahay, Hardee, & Batts, 1981), did not mediate the effect, thus suggesting that stereotype activation influences judgments independent from stereotype endorsement (i.e., regardless of participants' prejudice level). Recent research challenged this notion (see Kunda, 2000). Whereas some criticized the lack of sensitivity of the scale (Fazio, Jackson, Dunton, & Williams, 1995), others (Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997) noted that the words Devine used were mostly negative and might have activated negative associates of the category of African Americans, which would not be activated by neutral primes. Using the Modern Racism scale, Lepore and Brown (1997) found no difference in stereotype activation for negative primes of the African American stereotype (rude, dirty, drugs) between people high and low in prejudice. For neutral primes (e.g., African and West Indians), however, only highly prejudiced individuals experienced automatic activation of the negative stereotype (see also Wittenbrink et al., 1997). This suggests that people low in prejudice do not automatically associate negative concepts with the stereotype of Blacks. However, if the negative aspects are primed, participants automatically activate the entire, culturally shared negative stereotype, regardless of prejudice level.

Contextual Priming

Primes may be hidden within the context of the target stimulus and affect its interpretation. Gilovich (1981) asked participants to render judgments about the wisdom of American intervention in a hypothetical international conflict that was described in a given scenario. Two familiar and evocative historical events were primed within the context of the scenario: One was the American intervention during World War II, and the other was America's involvement in the Vietnam War. For example, the World War II script was primed by describing a hypothetical, modern-day invasion in terms of a "Blitzkrieg invasion" that resulted in refugees fleeting to a neutral country in boxcars. The Vietnam script was invoked by describing the same invasion in terms of a "quick strike invasion" resulting in refugees' flight to a neutral country in small boats. Participants were also shown a map with contours and labels invoking associations either to Europe or to Southeast Asia. The results showed that participants primed with World War II were more in favor of direct U.S. intervention. Notably, the primes did not change anything substantial in the described situation. Presumably, if asked to justify their choice, none of

the participants would invoke these as the reasons behind their recommendations regarding U.S. foreign policy.

Labels and Expectancies

Social labels may serve as primes and direct subsequent perceptions. In a classic experiment by Darley and Gross (1983), participants watched a videotaped interaction of a child in an examination, performing ambiguously (her answers were sometimes correct and sometimes incorrect). Prior to watching the videotape, some participants were informed that the child comes from an upper-class family, whereas some participants were informed that she comes from a lower social class. Participants were asked to evaluate how well she performed in the videotaped situation. Social class biased the perceptions, such that performance was judged to be higher for the upper-class child than for the lower-class child.

We should note that many of the effects of labels and expectancies are not pure effects of knowledge accessibility because they might be justified in logical, propositional terms. For example, in the aforementioned study, the researchers could have asked participants to predict the child's grade point average 10 years later. In that case, it could be reasonable to predict higher performance based on social economic status, in light of the well-established relation between these variables. Notably, in Darley and Gross's (1983) study, the question concerned the videotaped performance rather than future performance. In addition, Darley and Gross found that labels created the bias only when they were presented prior to watching the videotape, but had no biasing effect if presented between the video and the judgment task.[3] Logical effects should be relatively insensitive to the order of presentation of information, unlike accessibility effects, which occur, by definition, only if the prime precedes rather than follows the target (Fiske & Taylor, 1991).

Expectancies can be enduring and can affect information processing. A recent study found a relation between alcohol cues and expectancies concerning sexual arousal for male participants, which in turn influence arousal-related judgments (Friedman, McCarthy, Förster, & Denzler 2005). More specifically, it was found that unconscious exposure to alcohol-related cues, relative to nonalcohol cues, increased the tendency to judge women as sexually attractive but had no effect on judgments of their intelligence. This, however, was true only for those who held a belief in the aphrodisiac qualities of alcohol. Those who did not expect alcohol to increase their sexual desire were unaffected by the alcohol-related primes.

Priming of Dimensions

Sherman, Mackie, and Driscoll (1990) demonstrated priming of dimensions of evaluation. They asked participants to choose between two candidates for an election who were described on dimensions of foreign policy and economic matters. One candidate was described as having positive features relevant to foreign policy and a neg-

ative feature relevant to economic matters, whereas the other candidate was described in the opposite way. Both candidates were also described by neutral information. After having been primed with either the dimension of foreign policy or the dimension of economic matters, participants based their choice of the candidates on the primed dimensions. Recall for information from the dimensions was also measured, and it served as a proxy of attention. Attention mediated choice in that people attended more to the primed dimension and used it in their decision.

Priming of Opposites

Closely related to the notion of priming dimensions is the idea that constructs may be associated with their bipolar opposites. Park, Yoon, Kim, and Wyer (2001) showed that "healthy" may prime "unhealthy." If the implication of the target information were closer to the latter, it would be used for the judgment. It was found that priming "good taste" and "bad taste" before evaluating artificially flavored milk resulted in an assimilation effect: Participants evaluated the milk more favorably after "good taste" primes compared with "bad taste" primes. However, when primed with "healthy" or "unhealthy," both conditions rated the milk more unfavorably than a control, no priming condition. The authors reason that because "healthy" is associated with "unhealthy," both are made accessible by the prime and can be used for the judgment. However, when the value of the target is unambiguous, such as the case in which artificial flavor (which has a bad reputation related to health) is added to milk, people are primed only by the bipolar opposite when primed with "healthy." Importantly, the study shows that when activating a concept increases the attention to information that exemplifies its bipolar opposite, it could appear as an automatic contrast effect (see below). Notably, it is possible to conceptualize the foregoing results as priming of a dimension. Namely, it is possible that the entire dimension of health was primed, and the milk was judged as less healthy simply because this end of the dimension was more applicable than the "healthy" end.

Assimilation and Contrast Effects in Knowledge Activation

The effects of accessibility do not always directly reflect the primed construct. Before describing these effects, however, we would like to clarify some terms. If a measure reflects the implication of the prime, it is called an *assimilation effect*, whereas if it reflects the opposite implications of the prime, it is called a *contrast effect*. Several mechanisms of contrast have been distinguished in the literature (see DeCoster & Claypool, 2004). Consciously or unconsciously attempting to avoid using the prime or to correct its influence is called a *correction effect*. When used as a standard, contrasting away from the prime is called an *anchoring effect*. If knowledge is made inaccessible below the baseline of activation it is called *inhibition*. In addition, correction often (but not always) results in a

contrast effect, although such contrast should be distinguished from that resulting from anchoring. The main difference is that correction is by definition a second-stage, metacognitive process that operates on a presumed previous influence of a prime and correction may or may not occur, whereas anchoring is a primary cognitive process that may happen by default.

The distinction between *effects* and *processes* is of central importance for understanding contrast effects. Suppose that after having been primed with "reckless," people judge a target person to be *less* reckless than an unprimed control group. Different processes could have produced the effect. For example, the participants might have been suspicious about the prime and decided not to use it or attempt to counteract its effect (judged the prime as unusable). Notably, this conscious process of correction may lead not only to contrast but also to assimilation (see Lombardi, Higgins, & Bargh, 1987) (e.g., if the person has an incorrect theory about the direction of the influence or if the correction was not strong enough) (Strack, 1992). Contrast may also occur if a prime is used as a standard of comparison, or an anchor. For example, if a person thinks that compared to the prime the target does not seem reckless, the person might judge the target to be "careful." It is also possible that the prime information is simply made inaccessible because, for example, people feel that the priming task has been completed and that anything relevant has already been "cleared" from their mental system, or, in other words, inhibited. Note that correction takes place after knowledge activation and before knowledge use (the efferent response stage), whereas inhibition takes place immediately after the activation state, regardless of whether or not a new target is encountered (the afferent perceptual stage). It appears that anchoring effects occur at both the afferent and the efferent stage, although more research is needed to clarify this issue.

In social psychological literature, assimilation and contrast are sometimes discussed at the level of effects (i.e., if researchers are interested in the end result of the priming manipulation) and at other times at the level of processes (e.g., if researchers are interested in the process of comparison). We discuss both.

Probably the most general framework that accounts for these effects is the inclusion–exclusion model by Schwarz and Bless (1992). It suggests that assimilation occurs when the target is included in the prime, whereas contrast occurs when the target is excluded from the prime. One can visualize the underlying metaphor with the prime and the target as two blots and the question of assimilation versus contrast as whether they are seen as one whole blot or two distinct blots. Fuzzier boundaries of either of the blots and a shorter distance between them would then promote assimilation. Moreover, a detailed, close look that concentrates on the differences may foster a perception of two blots (i.e., exclusion and contrast), whereas a more general, distal perspective and concentration on similarities may make them seem as one entity (i.e., inclusion and assimilation). Zooming in on the two blots is likely to produce a contrast, whereas placing both in a much wider context is more likely to re-

sult in viewing them as closer to each other and thus produces assimilation. Also, inclusion and exclusion are affected by motivation and habit—if one is motivated to assimilate the two stimuli, or if one habitually does so, assimilation would ensue. We believe that much research on assimilation and contrast can be viewed as an explanation for moderators that derive from the inclusion–exclusion model. We turn now to review some of this research. We focus here on the primary processes of anchoring and inhibition and later we discuss the secondary processes of conscious or motivated correction.

TRAITS, CATEGORIES, AND EXEMPLARS

In a series of experiments, Stapel, Koomen, and Velthuijsen (1998) found the typical assimilation effect after priming participants with sentences related to dependency: They judged an unrelated person to be more dependent than the nonprimed control group. Other participants, however, were primed with an exemplar, "Linda," who was described as a dependent person. Those participants rated the target to be less dependent than the nonprimed control group.

Exemplars typically have more distinct boundaries than concepts, and thus, according to the inclusion versus exclusion view, are expected to lead to contrast more often (see Stapel, Koomen, & van der Pligt, 1996). It is possible, though, that this tendency could be reversed when exemplars are vague and the concepts they stand for are clear. Förster, Kuschel, and Liberman (2006) found participants' extent of knowledge about an exemplar to be positively correlated with contrast. For example, when primed with a person who constituted a high standard for alcohol consumption (the British pop singer Robbie Williams), participants were more likely to assimilate estimates of their own alcohol consumption (how many days per months do you consume alcohol?) the less they knew about him. Because all participants knew that the exemplar is a pop singer, it appears that sometimes when exposed to an exemplar, participants may construe it in terms of the category, making assimilation more likely.

CLOSENESS AND SIMILARITY

Stapel and Koomen (2000, 2001; Stapel & Suls, 2004) examined variables related to the distance between the prime and the target. In an extensive research program, they showed that close, highly relevant, similar and indistinct standards lead to assimilation rather than contrast. Furthermore, affective assimilation (e.g., feeling happy), when comparing oneself to an upward standard, was shown to occur with attainable and controllable standards (for a review, see also Markman & McMullen, 2003), which may be assumed to be more psychologically proximal than unattainable standards. In a similar vein, Mussweiler and Bodenhausen (2002) asked male participants to compare themselves to either a very tidy, clean male or female person. Accessibility was measured with a lexical decision task and it was found that an ingroup, male target led to enhanced accessibility of standard consistent knowledge, whereas a female, outgroup target enhanced accessibility of standard-inconsistent knowledge.

Thus, close targets are more likely to produce inclusion and thereby lead to assimilation than psychologically distant targets.

EXTREMITY OF PRIMES

Using a modified version of the priming paradigm by Higgins and colleagues (1977), Herr, Sherman, and Fazio (1983) primed participants with moderate and extreme exemplars of animals in terms of ferocity and size and then asked them to judge a fictitious animal. The study found assimilation after moderate exemplar priming and contrast after extreme exemplar priming. Two studies conceptually replicated these results with hostility priming and judgment of an ambiguously aggressive target (Herr, 1986; see also Mussweiler, Rüter, & Epstude, 2004b). Furthermore, Dijksterhuis and colleagues (1998) argued that extreme and distinct exemplars, as opposed to mild and less distinct exemplars, give rise to a process in which one compares oneself to the target, a process that naturally produces a contrast (see also Stapel et al., 1996, 1997).

Extreme primes are, by definition, more remote from the target and thus are more likely to be excluded and lead to contrast than moderate primes. Indeed, if prime extremity does not accompany an increased distance to the target, it does not enhance the likelihood of contrast. Skowronski, Carlston, and Isham (1993) found assimilation effects for extreme exemplars such as "idiot" before judging a target that was labeled "mentally retarded."

SUBJECTIVE VERSUS OBJECTIVE SCALES

Mussweiler and Strack (2000) demonstrated that priming people with moderately high and low standards of drug consumption (Frank Zappa vs. Steffi Graf) led to assimilation on an objective scale (e.g., the number of times per month that one consumes drugs) and to a tendency of contrast on a subjective scale (e.g., what is the extent of drug consumption). This is also consistent with the inclusion–exclusion view: A subjective scale is adjusted according to the range of the judged set of stimuli (Parducci, 1965; Parducci, Perrell, & Marsh, 1969) and is therefore more likely to lead to zooming in on the stimuli, leading to contrast. An objective scale, on the other hand, is more likely to be wider than the distance between the prime and the stimulus and thus places them closer to each other than zooming in on only these two stimuli. We would predict that an objective scale that zooms in on the prime and the target would produce contrast much like subjective scales (see also Stapel & Suls, 2004).

INTERPRETATION VERSUS COMPARISON

Assimilation and contrast may ensue from different processing goals. In their interpretation–comparison model (ICM) of social comparison, Stapel and colleagues (Stapel, in press; Stapel & Koomen, 2000, 2001; Stapel & Suls, 2004; see also Blanton, 2001, for a related view) proposed a distinction between an *interpretation mindset*, in which people try to make sense of a target, and a *compari-*

son mindset, in which people compare a target to a standard. They suggested that the former produces assimilation, whereas the latter produces contrast, given that the standard is sufficiently extreme to be used as a comparison standard. Obviously, comparison, more than interpretation, involves distinguishing the target from the prime and thus exclusion. In their study, Stapel and Koomen (2001) primed participants with mindsets of interpretation (e.g., priming "comprehend" or "interpret") or comparison (e.g., priming "compare" or "differ") and demonstrated assimilation after interpretation priming and contrast after comparison priming.

TESTING HYPOTHESES ABOUT SIMILARITY VERSUS DISSIMILARITY

Processing goals may create assimilation versus contrast also by directing the search of information toward similarities versus differences between the target and the standard. This view was advanced by Mussweiler and Strack's selective activation model (SAM) (Mussweiler, 2003; Mussweiler & Bodenhausen, 2002; Mussweiler & Strack, 1999, 2000a; Strack & Mussweiler, 1997), which focuses on the comparison process itself. This model suggests that the process of comparative evaluation (such as "Am I athletic?") consists of selecting a standard (e.g., a friend or a neighbor), comparing oneself to the standard, and integrating knowledge. During the comparison phase, people search for and focus on hypothesis-consistent information. For example, comparing her athletic abilities with an extremely athletic standard, a person might try to examine the hypothesis "I am not athletic" and thus search her memory for confirming information (e.g., a situation in which she was lazy). Using lexical decision tasks, Mussweiler and Strack (2000a) found activation of selective information when people were exposed to moderate or extreme standards. For example, after comparison with a moderately low standard with respect to athletic ability, participants' lexical decisions were faster for words associated with "athletic" than with words associated with "unathletic," whereas the reverse pattern was found for a moderately high standard.

This selective accessibility mechanism from confirmatory hypothesis testing also applies to classic anchoring paradigms (Tversky & Kahneman, 1974). Mussweiler and Strack (2000b; see also Strack & Mussweiler, 1997) demonstrated that numeric anchors selectively activate information that is consistent with the high or low anchor. For example, in one study participants were asked whether the annual mean temperature in Germany is higher or lower than 20° C (high anchor) or 5° C (low anchor) and then asked to estimate the average temperature in Germany. The estimates reflected the typical assimilation effect, being higher after the high anchor than after the low anchor. In addition, a subsequent lexical decision task with summer-related and winter-related words showed faster lexical decisions for summer words after high anchors than after low anchors and the reverse for winter words. These results show that participants engaged in an anchor-consistent search that selectively activated semantic information used for the judgment.

In a recent demonstration of the search process, Mussweiler and colleagues (Mussweiler, 2001; Mussweiler et al., 2004a) primed a similarity versus dissimilarity search by introducing a first task in which participants searched for either similarities or dissimilarities. They showed that if participants searched for similarities in an unrelated first phase of the experiment, they engaged in a similarity search when exposed to moderate standards yielding assimilation. However, a dissimilarity search before exposure to the same standard yielded contrast.

GLOBAL VERSUS LOCAL PROCESSING

Including information and the search for similarities between the prime and the target requires global processing, whereas excluding information and the search for differences between the prime and the target requires local processing (Navon, 1977). For example, to include an exemplar into a category (e.g., Is a camel a vehicle?) a broadening of conceptual scope and more abstract representations might be necessary (Isen & Daubman, 1984). On the contrary, exclusion might involve focusing on concrete details that makes the exemplar distinct (e.g., deciding that a camel is not a vehicle involves noticing its distinct features that do not fit the category; see Friedman & Förster, 2002). As a result, focusing one's attention on abstract features (i.e., global processing) might lead to assimilation, whereas a focus on concrete information (i.e., local processing) might lead to contrast.

Förster and colleagues (2006) showed that global versus local processing (manipulated in a task used by Macrae & Lewis, 2002, a variant of the Navon, 1977, task, in which participants attend to either the big letter or the small letter when presented with a big letter that is made of small letters) may engender assimilation versus contrast. To give one example, in a social comparison task global processing led to assimilation even when the primes were extreme and even when the scale was subjective, and local processing led to contrast even when the primes were moderate or when the scales were objective. Moreover, in one experiment, selectively activated knowledge as measured with a lexical decision task was shown to mediate the assimilation effect on the judgment after global processing and the contrast effect after local processing: Consistent knowledge with the standard was activated when global processing was primed, whereas inconsistent knowledge with the standard was activated when local processing was primed. Further experiments showed that global processing also leads to assimilation in the classic Donald paradigm and in anchoring tasks, whereas local priming leads to contrast. Thus, it seems that global versus local processing styles predict inclusion versus exclusion and thus produce assimilation or contrast.

CONVERSATIONAL NORMS OF NONREDUNDANCY

Exclusion and inclusion may be motivated also by conversational norms. According to Strack, Schwarz, and colleagues, conversational norms of nonredundancy (Grice, 1975) can eliminate and even reverse assimilative

effects of priming (Schwarz, Strack, & Mai, 1991; Strack, Martin, & Schwarz, 1988). For example, a question about one's happiness with his or her marriage may prime knowledge that can be used in a subsequent question on "happiness with life in general," producing an assimilation effect. If, however, the questions are presented in close proximity to each other, thus emphasizing the relevance of the norm of nonredundancy, assimilation is eliminated.

COMBINATION OF THE MODERATORS OF ASSIMILATION VERSUS CONTRAST

Examining how the various factors mentioned earlier combine with each other is a matter of evolving research and a current debate in social psychology (e.g., Do implicit comparisons with extreme primes lead to assimilation or contrast?). Recently, Markman and McMullen (2003) proposed an integrative account of assimilation versus contrast effects in self-perception. Specifically, their reflection and evaluation model (REM) proposes a sequence of stages in which exclusion versus inclusion concerns initiate comparison or interpretation processes by specifying which information will be included in the construal of the target and which information will be excluded from it. Comparison (evaluation) and reflection (interpretation) then selectively activate knowledge, thereby making consistent or inconsistent knowledge accessible. *Reflective* processes involve experiential modes of thinking related to simulating situations (e.g., imagine myself being as smart as Albert Einstein), whereas *evaluative* modes of thinking involve using comparison standards as reference points to judge and evaluate reality (for a related view, see Epstein, Lipson, Holstein, & Huh 1992; Strack & Deutsch, 2004). In this model, reflective and evaluative processes are not mutually exclusive, and might operate simultaneously or consecutively. As noted before, the relative strength of these processes is determined by the extent to which various factors encourage a person to think about him- or herself and the standard as single units (e.g., inclusion) or as distinct entities (e.g., exclusion).

Behavior, Goals, and Accessibility

In this section, we examine the effects of priming on behavior. Some studies used standard priming procedures and examined their effects on subsequent behavior. Research in a different direction examined the effects of goals on the accessibility of goal-related constructs. The findings of this research uncovered special properties of accessibility from active goals, which are distinct from effects of simple priming on accessibility.

Priming of Behavior

Behavior has been shown to be affected by priming of various constructs. In what were perhaps the first studies using behavior as a dependent variable, Straumann and Higgins (1987) primed ideal versus ought (i.e., responsibilities, duties) discrepancies. Positive attributes that

were discrepant with ideal selves slowed down talking whereas positive attributes that were discrepant with ought selves speeded up talking. It was theorized that this was because discrepancies from ideals produce a dejection-related affect, which is known to slow down level of activity, whereas discrepancies from oughts produce agitation that tends to have the opposite effect.

Priming stereotypes appear to be of considerable interest for social psychology. In a now classic example, Bargh, Chen, and Burrows (1996) subliminally primed participants with either African American faces or Caucasian faces. The former condition was intended to prime aggression, as many Americans have been shown to associate African Americans with aggression (e.g., Devine, 1989). After priming, participants were filmed as they completed a boring computer task during which the computer unexpectedly crashed. Participants' behavior was more aggressive after having been exposed to African American faces than to Caucasian faces.

Researchers have found assimilation effects with other behaviors as well, including slower walking after priming of the elderly (Bargh et al., 1996), enhanced helping behavior after semantic priming (Walther, Müller, & Schott, 2001), and conformity after semantic priming (Epley & Gilovich, 1999). Recently, Shah (2003a, 2003b) demonstrated that priming participants with significant others activated goals associated with those people (see also Susan Andersen's work, reported later). For example, priming one's father, by subliminally presenting the words "father" and "dad" before each anagram, enhanced performance, commitment, and persistence on the task that was presented as a test of analytical reasoning. However, this facilitative effect occurred only for participants who were close to their father and whose fathers valued analytic reasoning.

Automatic mimicry represents another interesting case of priming of behavior. Chartrand and Bargh (1999) found that people tend to mimic the behavior and expressions of their interaction partners without being aware they are doing so. In this case, unlike in other studies, the prime is not a semantic stimulus but, rather, an observed behavior in an interaction, which produces similar behavior in the observer. Chartrand and her colleagues found that automatic mimicry increases liking between conversation partners and smoothens conversation, and argued, accordingly, that automatic mimicry serves a social function (for a review, see Chartrand, Maddux, & Lakin, in press).

CONTROLLING BEHAVIOR BY PRIMING: MYTH OR REALITY?

From a theoretical social cognitive perspective, the notion that behavior may be affected by priming may seem just a small leap from the established finding that priming affects cognition. From a different perspective, however, the idea that priming may change a person's behavior is far reaching and even unsettling. In the realm of consumer behavior and political voting, the possibility of altering such behavior using subliminal priming is frightening. But are such effects actually possible? The current answer is a cautious "yes" (Hassin, Uleman, & Bargh,

2005; Wyer, in press). Although such influence is possible, there are limitations on their effects.

Bargh, Gollwitzer, Lee-Chai, Barndollar, and Troetschel (2001) demonstrated that participants primed with cooperation behaved in a more cooperative way in a resource-dilemma task than did nonprimed participants. They also found that priming achievement via a word-search puzzle enhanced the tendency to persist in an achievement-related activity, in spite of the presence of an obstacle created as an instruction to stop working on the task. They concluded that primes may activate behavior if prior experience established a strong link from the primed semantic concept (e.g., cooperation) to a particular behavior (behaving in a considerate way in common-resource situations).

Other research demonstrated, more directly, that effects of priming on behavior depend on appropriate associations. Mussweiler and Förster (2000) found that sex-primed male and female participants showed enhanced perceptual readiness (i.e., faster lexical decisions) for aggression-related words. However, men and women differed in their *behavioral* responses to those primes as well as in the way these primes affected disambiguation. Specifically, behavioral effects of enhanced aggression were evident in men but not in women, whereas disambiguation effects were evident for women but not for men (sex-primed women rated an ambiguous target as more aggressive). The authors explain this pattern in terms of gender differences in experiences with sex, proposing that because women are usually the victims of sexual aggression and men the perpetrators (Koss et al., 1987), women might have learned to be sensitive to detecting aggressive behavior in males in those situations, while men might have learned to act more aggressively. Notably, even though both males and females seem to relate sex to aggression as measured by the lexical decision task, the implications of this perceptual readiness for behavior and judgment may be different for men and women. More theoretically, the study shows dissociation between the effect of priming on perceptual readiness, disambiguation, and behavior.

Other studies have found that priming affects behavior only if it is consistent with an already-existing motivation. Strahan, Spencer, and Zanna (2002) showed that priming people with thirst made them consume more beverages only if they were thirsty. Thus, a preexisting motive to pursue the goal seems to be necessary for priming to have an effect on behavior.

A similar demonstration of the need for a behavior-consistent motivation was shown by Aarts, Gollwitzer, and Hassin (2004), who used short paragraphs on a male's wooing behavior to prime males with the goal of having casual sex. They found that primed participants, compared to nonprimed participants, were more helpful toward a female confederate but not more helpful toward a male confederate (helping a woman had been shown to reflect gallantry). In addition, this study demonstrated that perceived goal appropriateness moderated the effect of priming on behavior. In one of the studies the paragraph used for priming was about a flirting protagonist who was engaged in a serious relationship. It was assumed that in this case casual sex would be morally unacceptable to the participants. Results indicated no priming effect on helping behavior in this condition, although participants gave no reports of being aware of either the effect of priming or the blocking effect of goal inappropriateness. These studies demonstrate that appropriateness limits the effects of goal priming on behavior.

Alternative focal goals may also block the effects of priming on behavior. Macrae and Johnston (1998) found that participants primed with helpfulness did not stop to help a confederate to pick up his pens (scattered on the floor) when they were in a hurry to the next experimental session. Concurrent activation of incompatible goals might also explain the fact that people do *not* get up in the movie theater to buy soft drinks when they are subliminally primed with these drinks during the film: They have the focal goal of watching the movie and not buying a soft drink (see Bargh et al., 1996).

Furthermore it is worth mentioning that behaviors are activated by a configuration of stimuli, some of which may be related to the situation at hand and others that may be activated by factors that occurred before the situation (i.e., primed concepts, or chronically accessible concepts). The same primed concepts can influence different behaviors, depending on the situation in which the behavior is relevant. For example, priming Black faces may elicit expressions of aggression in response to a frustrating situation but high motivation to perform well in a music contest (see Wyer, 2004).

In sum, it appears that priming (e.g., Coke) may affect behavior (e.g., buying Coke) under certain conditions, including (1) if there is a previously established conceptual link between the prime and the behavior in question (e.g., if people have the habit of drinking coke when thirsty); (2) if the general motive underlying the behavior is active (e.g., if people are at least somewhat thirsty); (3) if the behavior is appropriate (e.g., if there is no stigma attached to buying drinks in movie theaters); and (4) if there are no competing goals (e.g., to watch the movie instead of buying a soft drink). It is beyond our expertise to conclude whether, in light of these limitations, the effect of priming on behavior is worth commercial pursuit. We believe, however, that such effects have vast theoretical importance.

There are important yet unanswered questions in the domain of goal priming, including the differences between priming general motives (e.g., competitiveness) and specific goals or behaviors (e.g., working on a puzzle), the mechanism that underlies priming of goals and motives, and what determines how these motives will be channeled into specific situation-appropriate behavior. Future research will have to explain how we can know, in advance, whether a prime (e.g., glasses) would elicit semantic associates (e.g., lenses) or a related goal (e.g., I have to find my glasses). Future research has also to clarify what kind of behavior is activated by goal priming. If food is primed, what would determine whether people start salivating, making donations for a third-world country, or prepare a shopping list? Some research has begun to uncover the answers to some of the questions about goal priming.

Assimilation and Contrast

As with disambiguation, contrast effects have also been found with priming of exemplars as opposed to categories. Dijksterhuis and colleagues (1998) replicated a study (Bargh et al., 1996) showing that participants primed with concepts related to the elderly stereotype (e.g., Florida, old, and lonely) walked more slowly down a hallway compared to those primed with words unrelated to age (e.g., thirsty, clean, and private). In addition, some of the participants were primed with an old exemplar (the then 89-year-old Dutch princess Juliana). These latter participants showed a contrast effect in that they walked faster than a control, neutral prime group.

Taking a more functional perspective of priming, Cesario, Plaks, and Higgins (2006) argued that the activation of social categories is used for preparing interactions with a certain category member. In one study, they found that if undergraduates liked the elderly, they walked more slowly after the "elderly" prime, but if they disliked the elderly, they walked more quickly after the "elderly" prime. Thus, assimilation effects were only found for people who had a motivation to approach the elderly, whereas for those who wanted to avoid them, faster walking was observed, serving their motivation. Obviously, liking may be viewed as promoting inclusion within the liked category whereas dislike promoted exclusion.

Automatic Action–Situation Links: The Case of Habits

Aarts and Dijksterhuis (2000) argued that habits constitute strong goal–action links that cause an automatic activation of the action upon goal activation. They suggested that habits may evolve from controlled actions by means of repetition of the goal–action compound. In a series of studies, participants were primed with a reason to go to a certain location (e.g., shopping) and were then exposed to 40 different locations, for which they had to decide as fast as possible between different travel modes. Only some of the locations were possible to get to by bicycle. The frequency with which the participants used their bikes to go to the specified places (e.g., the university or the mall) was used as a measurement of participants' habit strength. In the experiment, faster responses were found for habitual bike users after goal priming only for those places visited by bike on a regular basis; no effect was found for the group not primed with the goal. The latter group allowed for elimination of the possibility that the effect was due to a simple associative link between cycling and some locations.

Effects of Priming on Performance

A study by Dijksterhuis and colleagues (1998) first demonstrated effects of priming on performance on intelligence tests. Participants wrote for 5 minutes about the typical activities, lifestyle, and appearance of either a professor or a secretary. A third control group was not primed. Participants then performed a general knowledge test. Priming of "professor" enhanced performance compared to the other two groups. This effect may seem controversial but only, according to the authors, if one thinks of ability-related performance as fixed and the measurement as error-free. The effects of priming, they argue, may stem from situational variations in motivation, concentration, and levels of aspiration, all of which could be affected by priming.

Priming also affects memory performance. Dijksterhuis, Aarts, Bargh, and van Knippenberg (2000) had two groups of young participants perform a lexical decision task either with subliminal primes of the elderly (e.g., the words *old*, *gray*, and *bingo*) presented before the target words in the task or with no priming. They found that participants primed with "elderly" who had frequent contact with elderly people performed worse in a surprise recall test of the target words than did participants who were not primed and participants who were primed but had infrequent contact. The authors interpreted this as an automatic assimilation of performance to the stereotypical elderly trait of forgetfulness (for similar results, see Dijksterhuis, Bargh, & Miedema, 2000; Dijksterhuis, Spears, & Lépinasse, 2001, study 2).

However, contrast effects on performance have also been reported. In a study by Levy (1996) old and young participants were subliminally primed with either positive elderly-stereotypical words (e.g., wisdom), or negative elderly-stereotypical words (e.g., senility). Memory performance was assessed both before and after priming. Assimilation effects occurred for elderly participants (i.e., reduced performance after negative primes and enhanced performance after positive primes), whereas young participants exhibited a slight contrast on two of the memory indices. It seems that an enhanced distance of the target (i.e., the self) from the prime (i.e., elderly) increased the tendency toward contrast rather than assimilation, as would be predicted by the inclusion–exclusion hypothesis.[4]

Flexibility of performance can be primed as well. A recent study (Hassin & Bargh, 2005) primed participants with semantic associates of flexibility (e.g., elastic and bend), and found enhanced performance on the Wisconsin Card Sorting Task (WCST; Berg, 1948), which assesses whether participants can detect changing rules.

Furthermore, effects of priming on creativity have been demonstrated by Förster, Friedman, Butterbach, and Sassenberg (2005) with priming concepts indirectly related to creativity. They primed participants with stereotypes of a punk or an engineer, words that were pretested as being associated with high versus low deviancy. It was shown that the punk prime increased creativity and impaired performance on analytical tasks, whereas the reverse was true for the engineer prime. Notably, in the pretests punks and engineers did not differ on the dimension of creativity. Thus, unlike Hassin and Bargh (2005) who primed the concept relevant to the task (flexibility) directly, the effects in this study cannot be explained by different activation of the creativity dimension but rather by a semantic link between deviancy and creativity.

It is entirely possible that contrasts in creativity could also be obtained via priming. For example, a highly cre-

ative exemplar prime, such as Leonardo da Vinci, might remind oneself of one's relative lack of creativity, thereby decreasing motivation, producing intimidation, or eliciting careful behavior, all of which hinder creative thinking (Friedman & Förster, 2001). This prediction awaits future testing.

Overall, effects of priming on performance have been found with intelligence tests, memory performance, flexibility, and creativity. In some of these areas, both assimilation and contrast were demonstrated, whereas in other areas, the possibility of a contrast waits for future examination. These effects demonstrate situational variations in ability and could be mediated by different goals, processing styles, motivation, or level of aspiration elicited by the primes.

The Effects of Goals on Accessibility: Differences from Semantic Priming

In the previous section we looked at the effects of priming on behavior. We now turn to examine the effect of goal activation on construct accessibility, especially how these effects differ with respect to their underlying processes from the effects of semantic priming.

PERSISTENCE OF ACCESSIBILITY

Whereas the effects of simple semantic priming decrease fairly rapidly (e.g., Srull & Wyer, 1979), the effects of goals on accessibility seem to remain as long as the goal is active. Goschke and Kuhl (1993) demonstrated this principle by having participants rehearse a series of actions and then informing them that they would either perform the actions (i.e., a goal) or observe another person performing them (i.e., a nongoal). Using a recognition test, they found a higher accessibility of the actions in the goal condition than in the no-goal condition, even when rehearsal of the actions was not possible in the intervening time. A similar persistence of the effect of goals was demonstrated by Bargh and Barndollar (1996), who activated achievement and found that delay increased, rather than decreased, the effects of priming.

POSTFULFILLMENT INHIBITION

Another difference between conceptual priming and goal priming is postfulfillment inhibition—a tendency for goal-related constructs to be inhibited after the goal is fulfilled. This effect was first described by Zeigarnik (1938), who demonstrated that people remembered interrupted tasks better than completed tasks (see Butterfield, 1964; Heckhausen, 1991). More recently, Marsh, Hicks, and Binks (1998; Marsh, Hicks, & Bryan, 1999), in an extended version of the paradigm by Goschke and Kuhl (1993) described earlier, used a lexical decision task to examine the accessibility of action-related constructs both before and after completion of the action. Replicating Goschke and Kuhl, before completion, accessibility of intended actions was enhanced relative to nonintended (i.e., to-be observed) actions. Moreover, after completing the action, accessibility of action-related constructs dropped below the level of the control, no-goal group, reflecting postfulfillment inhibition.

Recently, Fiedler, Schenck, Watling, and Menges (2005) conceptualized a state of an "open set" (or, in our terms, an unfulfilled goal) as a situation in which people try to fit an inappropriate label to a social interaction scene (e.g., the verb "to follow" to a scene that describes a person attacking another person), and a completed goal as a situation in which an appropriate verb is fit to a situation. They demonstrated inhibition after goal fulfillment and enhanced accessibility after lack of fulfillment.

It can be contended that positive feedback signifies a completed goal, whereas negative feedback signifies an incomplete goal. Consistent with this idea, Rothermund (2003) found inhibition of goal-related constructs after success feedback and increased accessibility after failure feedback. Altogether, postfulfillment inhibition is a general principle that can be observed when goals, as opposed to semantic concepts, are primed.

POSTFULFILLMENT INHIBITION VERSUS CONTRAST

Unlike contrast or correction, inhibition happens at the afferent stage of the priming process, before encountering the target. Therefore, postfulfillment inhibition is best demonstrated with measures of perceptual readiness (e.g., lexical decision) or perceptual interference (e.g., Stroop), because those measures preclude conscious suppression or contrast. It is interesting to examine whether some of the effects that were previously attributed to suppression or contrast due to comparison processes may instead reflect postfulfillment inhibition. One such attempt was recently advanced by Liberman, Förster, and Higgins (2005) who showed that unfulfilled priming tasks lead to enhanced accessibility, whereas a completed priming task leads to inhibition. Using a paradigm from Martin (1986), Liberman and colleagues (2005) asked participants to rate either 12 or 8 sentences on a certain trait dimension. Participants who were told to rate 12 sentences were interrupted after 8 sentences. Thus, all participants rated 8 sentences, but only for those planning to rate 12 sentences was stopping after 8 sentences experienced as an interruption. Accessibility of prime-related constructs was measured with both a lexical decision task and a disambiguation task. On both measures, they found enhanced accessibility after the interrupted priming task and reduced accessibility after completed priming. Thus, it is possible that some contrast effects may in fact reflect inhibition at the afferent stage rather than correction due to judged uselessness.

PROPORTIONALITY TO MOTIVATION STRENGTH

Förster, Liberman, and Higgins (2005) found that goal-related accessibility and postfulfillment inhibition were proportional to the motivation to achieve a goal. In their experiments, participants were instructed to search through a presented series of pictures for a target combination (glasses followed by scissors) and report it to the experimenter. The results reflected an increase in acces-

sibility of goal-related words before the combination was found and inhibition after the combination was found, when compared to a group of participants who watched the pictures without having a goal to find the target combination. Moreover, the authors manipulated the expectancy of achieving the goal (by telling participants that the target combination was present in 90% of the cases vs. only in 5% of the cases), the value of the goal (by telling participants that they would receive £1.00 vs. only £0.05 for finding the combination), or both expectancy and value. The experiments also included a no-goal, control group that merely saw the pictures without being given the goal of finding the target combination. As predicted, accessibility of goal-related words prior to fulfillment and inhibition after fulfillment were found in the high-expectancy and high-value conditions but not in the low-expectancy and low-value conditions. The combination of high value and high expectancy enhanced the effects. Therefore, the effects of expectancy and value on goal-related accessibility and on postfulfillment inhibition were interactive (i.e., multiplicative) and similar to the effects of expectancy and value on motivation (e.g., Atkinson, 1964; Fishbein & Ajzen, 1974; Vroom, 1966). It therefore appears that goal-related accessibility and postfulfillment inhibition are motivational phenomena that are proportional to the strength of the motivation.

MUTUAL PRIMING OF GOALS AND MEANS

According to goal systems theory, goals form a network that associates them with superordinate goals, subordinate goals (i.e., means), and competing goals (Kruglanski et al., 2002; Shah, Friedman, & Kruglanski, 2002). These networks may or may not be constructs that constitute common semantic associates of the goal. For example, the goal of buying groceries may be subordinated to a higher level goal of preparing a party and superordinate to the means of driving a car to the supermarket, but it is not necessary that groceries, party, and car would be associated with each other. Hence, another way to distinguish between the effects of goals and concept priming on accessibility is to look at the activation of superordinate goals, means, and competing goals. Indeed, Kruglanski and colleagues (2002) have shown using subliminal priming that lexical decision times for goal-related means were faster after priming goals, and lexical decision for goals were faster after were primed. Moreover, increased accessibility of the goal due to priming of means was associated with increased task persistence and better performance (Shah & Kruglanski, 2003).

Sometimes, an individual may have multiple means of achieving the same goal (the case of equifinality). It is possible to get to work by taxi, bus, or bike. It is also possible that the same means serve more than one goal: One can ride a bike for transportation, exercise, or entertainment (the case of multifinality). Shah and colleagues (2002) found that a goal's association with a means of fulfilling it is weakened by equifinality (by the presence of other available means; a fan effect; see Anderson, 1974, 1983). They showed that when primed with a goal, participants who listed fewer means toward that goal showed faster lexical decision for words related to these means. A similar effect was found with multifinality (i.e., the same means fulfilling more than one goal); the fact that some means (e.g., reading) were associated with many goals (e.g., entertainment, learning, and understanding) reduced the facilitative effect of primed means on perception of goals (see also Shah, Kruglanski, & Thompson, 1998).

ACTIVATION AND INHIBITION OF COMPETING GOALS

Goal shielding occurs when alternative goals are inhibited by pursuit of a focal goal. It has been demonstrated that priming chronic or situation-specific goals inhibited the accessibility of conflicting goal constructs compared to goal-unrelated concepts (Shah et al., 2002). In some of their experiments, goals were activated without the participants knowing, thus suggesting that self-regulatory tactics might be automatically engaged prior to conscious awareness. Moreover, a variety of motivational variables mediated the effect. They reflect the fact that inhibition increases with commitment to the goal and with the extent to which its attainment could substitute for the attainment of the alternatives. Inhibition also increased with need for closure and level of anxiety and decreased with level of depression. Presumably, these variables are related to a motivation to achieve closure on active goals.

Shah and Kruglanski (2002) also found that priming of alternative goals impedes performance on focal goals if the two goals are unrelated but facilitates performance on the focal goal if the two goals are related (see also Shah & Kruglanski, 2003). In a series of studies, participants completed anagrams after which they expected to perform a "functional thinking" task in which they were asked to list as many uses as possible for a box. While performing the anagram task, participants were subliminally primed with the other task by a brief presentation of the words "box use" before each anagram (participants in the control condition were primed with "view it"). It has been found that for participants who perceived the two tasks to be related, priming facilitated performance (as measured by both persistence and number of solutions), whereas for participants who perceived the tasks to be unrelated, priming interfered with performance. These studies demonstrate that priming of nonfocal goals may interfere with performance if the primed goal is unrelated to the focal goal but facilitate it if the two goals are perceived as related.

ACTIVATION OF SUPERORDINATE GOALS IN SELF-CONTROL SITUATIONS

When trying to resist temptations (e.g., drugs and sex), making higher-order goals (e.g., laws and religion) accessible can aid self-control. Recent research by Fishbach, Friedman, and Kruglanski (2003) showed that temptations activate higher-level goals that presumably help in self-regulation (i.e., to resist the temptation). Specifically, the authors subliminally primed temptation-related words (e.g., drugs, sex, and premarital) and found that

participants were much faster in making lexical decisions about words related to religious goals (e.g., prayer, bible, and god) than a control group that was primed with words unrelated to temptations (e.g., friends and talk). This effect also occurred under high cognitive load, when participants had to rehearse a nine-digit number during the task. Thus, it seems that people can learn to connect temptations directly with higher-order goals, in this case religious values, and that the accessibility of those higher goals might help self-control.

From this review, it seems possible to distinguish the effects of active goals on accessibility from the effects of concept priming on accessibility. Förster and Denzler (in press; Förster, Liberman, & Friedman, in press; see also Kruglanski, 1996) summarized some of the hallmarks of goal-related accessibility: (1) there is an increase in accessibility of goal-related constructs until the goal is fulfilled or abandoned; (2) inhibition of goal-related constructs occurs after goal fulfillment; (3) goal-related accessibility and postfulfillment inhibition are proportional to the motivation to pursue the goal; (4) goals can activate superordinate goals and/or subordinate means; and (5) goals can shield against competing goals.

Procedural Priming

Procedural priming refers to priming of procedures—at the priming phase people perform an action, and then (facilitating) carryover effects on performing other actions are examined (see, e.g., Schooler's [2002] notion of processing shifts). For example, solving addition problems facilitates solving other such problems. Notably, experimental demonstration of procedural priming does not involve learning in the sense that the action performed at the priming stage is overlearned to such an extent that a few more instances of practice make virtually no difference (e.g., adding numbers and inferring traits from behavior). The distinction between semantic priming and procedural priming is best exemplified in a study by Smith and Branscombe (1987), who replicated a classic study on the effect of priming on disambiguation (Srull & Wyer, 1979). In addition to the original priming group, in which participants unscrambled sentences describing hostile behaviors (e.g., leg her break he), they added a procedural priming condition, in which participants were presented with the same sentences in an unscrambled form and chose the matching trait (e.g., hostile). Note that inferring traits from behavior is, in fact, the procedure people have to do when judging the vaguely aggressive target person in the second phase. After either 15 seconds or 15 minutes participants rated the aggressiveness of a description of a vaguely aggressive behavior. Priming traits (i.e., unscrambling sentences) had an assimilative effect only after a short delay, but procedural priming continued to have this effect after a longer delay. It seems that participants practiced the procedure of extracting traits from behavior, and that making this procedural knowledge accessible facilitated making such inferences at the test stage. Moreover, it appears that the effect of procedural priming is more enduring than the typical accessibility effect from priming con-

cepts. In addition to these content-specific effects, there seem to be smaller, but reliable, *general* effects of procedural priming, in that after matching traits to behaviors, people are faster not only with these same traits but also with matching other behaviors to *different* traits (Smith, 1989; Smith, Branscombe, & Borman, 1988).

In a different domain, Higgins and Chaires (1980) demonstrated that priming may make procedures accessible that would enhance performance on the Duncker candle problem (Duncker, 1945), in which participants must recognize that a box filled with tacks can be used as a platform for a candle rather than just as a container for the tacks. The authors found an enhanced rate of solving the task after participants were primed with differentiated linguistic constructions (e.g., tray and tomatoes), as compared to undifferentiated linguistic constructions (e.g., tray of tomatoes). It seems that priming a differentiation procedure carried over to the following task that profited from performing this mental action.

In a related vein, Friedman, Fishbach, Förster, and Werth (2003) showed procedural priming effects on creativity. More specifically, participants were asked to complete visual tasks that forced them to focus perceptual attention on either a broad or narrow visual area. In an ostensibly unrelated task, participants performed a creative generation task (e.g., find as many creative ways to use a brick as possible). Participants primed with a broad focus produced more creative solutions than participants primed with a narrow focus. It was reasoned that the procedural priming expanded (or constricted) the focus of perceptual attention, a procedure that was carried over to the semantic network, and thereby improved (or diminished) creativity because creative generation profits from a broad conceptual scope (Förster, Friedman, Özelsel, & Denzler, 2006; Isen & Daubman, 1984).

Closely related to procedural priming is Schooler's (2002; Schooler, Fiore, & Brandimonte, 1997) theory of "processing shift," which assumes that a cognitive procedure that is activated in the course of engaging in one task stays remains active and is transferred to subsequent tasks. Transfer-appropriate processing shifts are said to result when the residually activated procedures are beneficial for subsequent processing, whereas transfer-inappropriate shifts are said to result when the procedures at hand impair subsequent processing. For example, Macrae and Lewis (2002) found that performance in a face recognition task was enhanced following a global processing task and was impeded following a local processing task.

Moreover, in social psychology, many researchers have used manipulations of "mindsets" to examine similar carryover effects. For example, action phase theory (Gollwitzer, 1990; Heckhausen, 1991) distinguishes between a predecisional stage, characterized by a deliberative mind state in which one contemplates alternatives and a postdecisional state, characterized by an implemental mindset in which attention is focused on doing rather than contemplating. Gollwitzer, Heckhausen, and Steller (1990) induced a deliberative mindset by making people contemplate personal change or an implemental mindset by making participants plan the ex-

ecution of a personal plan in detail. Participants were then asked to complete a fairytale that began with a description of a character with a decision conflict. Analysis revealed that deliberative mindset participants ascribed more deliberative and less implementational efforts to the protagonist of the fairytale than implemental mindset participants (for additional effects of implemental and deliberative mindsets, see Taylor & Gollwitzer, 1995). Thus, thinking in an implemental or deliberative way carried over to an unrelated task.

In construal-level theory (see Liberman, Trope, & Stephan, Chapter 15, this volume), high-level versus low-level reasoning may be manipulated through a mindset. For example, Fujita, Trope, and Liberman (2005) had participants think of *why* they would like to maintain social contacts (a high-level construal mindset) or *how* they would maintain social contacts (a low-level construal mindset) and then examined performance on an unrelated self-control task that involved squeezing a handgrip in order to obtain self-diagnostic information. It has been suggested that self-control would be facilitated by a high-level construal of the situation and thus would be enhanced by procedurally priming participants with thinking in high-level terms. Consistent with this prediction, it has been found that a high-level construal mindset enhanced self-control relative to the low-level construal mindset.

How can one distinguish between concept and procedural priming? For one thing, procedural priming seems to persist longer than concept priming. However, it seems that activation of semantic and procedural knowledge may occur simultaneously and might have independent effects on information processing. Smith (1990) argued that in some priming tasks, the process in the priming phase is both conceptually and procedurally similar to the test phase. For example, if people translate behavioral information into traits in the first phase and then have to judge a target's aggressiveness (a trait) on the basis of her described behavior, the priming task might confound procedural and concept priming. Further research is needed to distinguish between the two kinds of priming.

Finally, we would like to briefly address the question of contrast effects with procedural priming. It seems that procedural priming tasks do not tend to engender correction processes. This may be the case for several reasons, including lack of awareness of the effects of procedural priming or lack of motivation to counteract them. For example, if a person has been primed with a wider perception, and if this facilitates creative thinking, she might not be aware of its effect on creative problem solving or might not be interested in correcting for this influence. It is also possible that people would not know how to counteract or undo effects of procedural priming, as that would, typically, involve more complicated processes than adjusting one's response on a scale (e.g., the person must know which processes facilitate creative thinking or which mechanisms can be used to switch from one thinking style to another). Of course, it is sometimes possible to stop a priming procedure that we suspect of having undesirable effects on us (e.g., one may stop watching a detective movie if she thinks that it makes her too paranoid). Moreover, stopping an already elicited procedure may be easier than actually counteracting it. Notably, it is entirely possible that future research would uncover correction of procedural priming. However, it seems to us that such correction is less likely than correction of disambiguation.

Affective and Evaluative Priming

There is ample evidence showing that feelings can influence affect, behavior, and cognitions in both assimilative and contrastive ways. Our definition of accessibility is broad enough to include mood as a source of accessibility. These effects, however, are being discussed in detail in by Schwarz and Clore (Chapter 16, this volume) and are not addressed here.

Affective Priming

In a classic affective priming (or evaluative priming) paradigm, Fazio and colleagues (1986; for a review, see Fazio, 2000) showed that deciding on the valence of a strongly valenced word (e.g., love and delightful) is better facilitated when it is preceded by an evaluative-consistent prime word (e.g., sunshine) than by an inconsistent prime word (e.g., death; for a review on affective priming, see Klauer & Musch, 2003).

It has been argued that evaluative priming demonstrates response facilitation, such that a positive prime prepares participants to select the response "positive," which is accelerated if the target word is positive as well. When the task is changed into a pronunciation task (i.e., participants have to pronounce the target word rather than classify it as positive vs. negative), the effect is reduced considerably and sometimes even disappears (see Klauer & Musch, 2003). Nevertheless, the existence of evaluative priming with a pronunciation task (Bargh, Chaiken, Raymond, & Hymes, 1996) indicates that evaluation is elicited spontaneously and constitutes a necessary part of processing the stimulus.

Affective priming has also been demonstrated using a lexical decision task, in which faster decisions are expected after evaluative-congruent primes compared to evaluative-incongruent primes (Wentura, 1998, 2000). This task, however, has been criticized as well. Klauer and Musch (2003) argued that evaluative consistency (e.g., between the words "love" and "sunshine") elicits a feeling of plausibility, which facilitates affirmative responses, whereas evaluative inconsistency (e.g., "love" and "war") elicits a feeling of inconsistency, which facilitates negative responses. A critical test for this model was a study by Wentura (1998, 2000) in which the assignment of "yes" and "no" responses to words and nonwords in a lexical decision task was reversed. Consistent with the model, words preceded by evaluative-inconsistent primes yielded faster "no" responses than evaluative-consistent primes. The findings cast doubt on the conclusion that evaluative dimensions are processed automatically and facilitate the processing of subsequent stimuli of equal valence.

From a cognitive perspective, it might seem strange that an overlap of a single semantic feature (i.e., valence) would produce semantic priming (Huber et al., 2001; Huber, Shiffrin, Lyle, & Quach, 2002). However, it is possible that affective priming is not equivalent to semantic priming. Namely, primes may not activate all the evaluative-consistent constructs. One possible alternative is that the affective prime facilitates a response, such as approach or affirmation, and, therefore, similar responses to evaluative-consistent targets are facilitated. This would also explain why affective priming produces weaker effects with pronunciation tasks than with evaluative tasks, and why affective priming with a lexical decision task is sensitive to the type of response. Later we discuss the specifics of affective priming and its potential functionality.

Chronic Accessibility

So far, we have discussed situational manipulations of accessibility produced by priming. However, there are additional, more permanent sources of accessibility—chronic accessibility and high accessibility of certain constructs, which occur without priming. In a study by Higgins, King, and Mavin (1982), participants listed the traits of a person fitting the following categories: liked, disliked, frequently encountered, whose company they sought, and whose company they avoided. The descriptions were used to extract each person's chronically accessible traits by using the first trait listed as chronic (Higgins et al., 1982, study 2). One week later, participants were invited for another session in which they read a description of a target person that contained 12 individually tailored behaviors, 6 that moderately exemplified the chronic constructs of each participant and 6 that moderately exemplified nonchronic traits, which were yoked to be chronic traits of another participant. After evaluation of their responses, it was found that participants' impressions and memories of the target person were related to their chronic constructs rather than to the nonchronic constructs.

Chronic Accessibility as a Continuous Variable

Higgins and Brendl (1995) used a measure of chronicity (for the content "conceited"; Higgins et al., 1982) as a continuous variable, defined as the frequency of mentioning the term and its synonyms in descriptions of people. Participants were measured, and a few weeks later, they were called back to the lab to judge a vague social target on conceitedness. The authors found a strong relation between chronicity scores and conceitedness ratings. Lau (1989) found similar chronicity effects for political constructs and evaluating political candidates.

Bargh and Pratto (1986) used a color-naming Stroop task to measure perceptual readiness of chronic concepts, using the same criteria for chronicity as mentioned earlier (Higgins et al., 1982). They found more interference (i.e., higher accessibility) for chronic concepts than for nonchronic concepts. In addition, Bargh and Thein (1985) found that impressions and recall of a target person's behavior were better when they were related to a participant's chronic dimension than when they were unrelated to a chronic dimension.

Other aspects of efficiency of processing information related to chronic concerns have been demonstrated as well. Markus (1977) found that independent schematics (i.e., people who responded at the high end of both a self-descriptiveness and an importance scale for at least two of three independence-related traits) processed independent-related stimulus information faster and more consistently than independent aschematics (i.e., people who responded in the middle of the scale). Using the same chronic constructs and a dichotic listening task, Bargh (1982) showed that processing of chronically accessible information used fewer resources in the attended channel and more resources in the ignored channel.

Other research has focused on qualitatively different chronic concerns. In her work, Andersen (for a review, see Andersen & Chen, 2002) argued that representations of significant others are chronically accessible and thus have a special readiness to be activated by minimal contextual cues. In her social cognitive model of transference (Andersen & Glassman, 1996), evidence has been obtained showing that in encounters with a new person, representations of significant others may be activated, leading the perceiver to interpret the target person in ways that are derived from the activated representation. Research also showed that people respond emotionally, motivationally, and behaviorally to a target person in ways that reflect his or her relationship with the retrieved significant other. In their studies, Andersen and colleagues ask participants to describe significant others, and adjectives that were related to those people were extracted from these descriptions. These adjectives were then used in the second phase, a few weeks later, to describe a new target person. The authors showed that participants were more likely to interpret the behavior and to draw inferences from the behavior of the new target person in accord with the significant other to whom the adjectives were ascribed in the first phase of the experiment, although in themselves the adjectives did not afford such inferences. Moreover, the effects were reliably stronger than priming effects for stereotypes and other contents (see Chen & Andersen, 1999).

Put briefly, chronic concerns or beliefs can yield similar effects to frequent priming.

How Chronic Accessibility Develops

It has been suggested that chronic accessibility develops from repeated use of a construct (Higgins et al., 1982). However, it is not known whether goals and concerns play a crucial role in creating chronic accessibility. For example, while searching for a car to buy, one becomes sensitized to "for sale" ads on cars. However, it is unclear whether chronic accessibility would develop to the same extent if one were exposed to the same number of stimuli without having a relevant goal. The question of whether or not chronic accessibility reflects only extensive exposure or personal concerns still awaits empirical examina-

tion. Clearly, however, both personal concerns (i.e., goals, wishes, and preoccupations) and more extensive construct use create priming effects that last longer. Both constitute viable explanations for how chronic accessibility develops.

PRINCIPLES

We now discuss a number of general principles and theoretical questions that have been examined in the literature on priming. We refer to decay after priming, extent of priming, and to the question of how accessibility from different sources combines. We also discuss the compensatory relation between accessibility and applicability and the context-dependent effects of priming.

Decay after Priming

We have already looked at some findings on decay rates of different types of priming. It has been demonstrated that more extensive *semantic priming* causes the effects to decay slower. Srull and Wyer (1979) asked participants to judge an ambiguously aggressive target 5 minutes, 1 hour, or 1 day after the priming task. After 1 hour, when the number of primes was low (6 of 30 scrambled sentences were related to hostility), there was little or no priming effect, but when the number of primes was high, priming effects were observed after 1 hour and were even visible after 24 hours. Thus, it appears that more extensive priming slows the rate of decay of the priming effect. However, when other constructs compete with the primed construct in terms of accessibility, effects of priming on judgments disappear as quickly as a few minutes (Higgins, Bargh, & Lombardi, 1985; Higgins & Brendl, 1995). Similarly, in an experiment by Smith and Branscombe (1987), effects of semantic priming were found after 15 seconds but not after 3 minutes.

Other studies have shown that the effects of *procedural priming* can last as long as 1 week (see Smith, Stewart, & Buttram, 1992). Thus, procedural priming can be distinguished from semantic priming in that the former has a slower rate of decay.

In contrast to procedural priming, the effect of *affective priming* on evaluation appears to decay even faster than semantic priming. Klauer, Rossnagel, and Musch (1997) conducted a study based on the original affective priming study by Fazio and Williams (1986) and added six levels of delay between prime and target exposure (-100, 0, 100, 200, 600, and 1,200 msec). They found affective priming effects only for the 0 and 100 msec delays. Likewise, Hermans, De Houwer, and Eelen (2001) found affective priming effects after delays of 0 and 150 msec but not for 300 and 450 msec. Using a pronunciation task, the authors found a similar pattern for 150 msec delays, which so far is the longest delay for which affective priming effects were obtained.

As previously mentioned, active goals enhance accessibility of goal-related terms and produce a relatively slow decay (and may even produce an increase in accessibility over time after the onset of the goal) and is dependent on goal fulfillment (Liberman & Förster, 2005). For example, Bargh and colleagues (2001) showed that activation of an achievement goal by priming constructs related to achievement increased performance on an anagram task and that this effect increased over a delay (see also Goschke & Kuhl, 1993).

By definition, *chronic accessibility* has a slow decay rate, as it refers to accessibility that is not created by a recent exposure to an external prime. Because chronic accessibility is the result of extensive construct use or of enduring personal goals or concerns, it is possible to view the slow decay characteristic of it as similar to the slower decay rates produced by extensive exposure to primes and by goals and concerns.

In sum, the effects of affective priming are relatively short-lived. Effects of semantic priming last slightly longer, while procedural priming, goal activation, and chronic accessibility decay slowly. Goal accessibility may even increase in activation over time.

Extent of Priming

Priming seems to follow principles of learning in that it obeys Thorndike's (1898) *law of effect:* The more recently and frequently a stimulus–response bond is activated, the more effectively it is engrained. Thus, it is logical that frequent priming should produce stronger effects than less frequent priming. To demonstrate, Srull and Wyer (1979) manipulated frequency of priming by presenting people with 48, 24, 12, or 6 scrambled sentences related to hostility. Assimilation effects increased as the frequency of priming increased, especially after short delays (see also Srull & Wyer, 1980). Subsequent studies replicated the prime frequency effect with subliminal priming (Bargh & Pietromonaco, 1982; see also Devine, 1989). The effects of priming on performance are also influenced by frequency of priming. For example, Dijksterhuis and van Knippenberg (1998) found that extended priming of "professors" led to better performance on a general knowledge quiz.

In procedural priming, the number of priming episodes may be thought of as practice. It should not be surprising, then, that more extensive priming leads to stronger effects. More habitual actions become overlearned and automatic and are performed more easily than less habitual actions (Gupta & Cohen, 2002; Smith & Lerner, 1986).

Priming may also be manipulated in terms of quality, not only by quantity of priming episodes. Consistent with this notion, it appears that depth of processing of the prime is associated with more robust priming effects (for a review, see DeCoster & Claypool, 2004). For example, reading paragraphs or sentences related to a trait produced stronger priming effects (both assimilation and anchoring) than simply reading trait words or being exposed to exemplars associated with the trait. In cognitive psychology, extent of processing was found to produce stronger and more enduring priming effects (Becker et al., 1997).

All told, it seems that extent of priming is a universal principle.

To the best of our knowledge, there is no research on the extent of goal priming effects on the accessibility of goal-related constructs. Possibly, extended priming may enhance the chance that a goal would be activated; however, once a goal is active, two sources may further enhance accessibility of goal-related constructs: rehearsal, which relates to extent of priming, and motivation to pursue the goal, which is not necessarily related to extent of processing. In a study by Förster, Liberman, and Higgins (2005) accessibility increased with greater proximity to goal fulfillment. This effect may be due to increased rehearsal of the goal over time or may be an independent effect of increased motivation as goal fulfillment is perceived to be closer (i.e., the fact that goals loom larger closer to fulfillment, see Förster, Higgins, & Idson, 1998; Miller, 1944). In another study, Goschke and Kuhl (1993) endowed participants with a goal and then systematically prevented them from rehearsing it by asking them to carry out a visuospatial task requiring spoken responses as participants were approaching the goal. Nevertheless, they found increased accessibility.

From the available research, extent of priming seems to be a moderator that affects semantic, evaluative, behavioral, and procedural priming and the effects of priming on goal activation. The effect of goals on accessibility of goal-related constructs may be produced independently of extent of priming through strength of motivation.

Reinstatement of Priming Effects

In a computer simulation, Smith and DeCoster (1998) found that accessibility increases faster if it has been primed frequently before, an effect they labeled *savings in relearning*. That is, the increase of accessibility for an already-studied construct is greater than the increase of a newly studied construct. Savings in relearning is also known to exist in explicit memory, such as recall and recognition tasks, as well as in implicit memory (Carlston & Skowronski, 1994). For example, it has been found that memorizing paired associates made it easier to memorize the same list even after intervals of several years and even after the learning material can no longer be intentionally recalled (see Burtt, 1941; Ebbinghaus, 1885/1964). To our knowledge, there is no *experimental* evidence showing such effects with accessibility; however, the effect is important and needs further research.

It is interesting to consider the possibility that savings in relearning could also occur with affective, procedural, or goal priming. Is the reoccurrence of emotional distress facilitated by previous instances of such events, even if they occurred in the distant past? Or, would a goal to interrupt a colleague because of his sexist remarks, which you forgot over time, be easier to reinstate by priming?

An interesting effect of reinstatement of priming has been demonstrated by Kunda, Davies, Adams, and Spencer (2002), who had participants watch a videotaped interview of a Black person for 12 minutes. One might expect that due to the long exposure, the category of Blacks would be extensively primed and produce an especially high level of accessibility. However, contrary to the findings of previous studies on priming of stereotypes, stereotype accessibility (measured as a speed of lexical decision) was evident only if the participant realized that the Black person expressed an attitude that deviated from the participant's own attitude. The authors explain these results by suggesting that without the counterattitudinal remark, the extended exposure to the Black person created an individual perception of him that replaced the categorical perception. However, if the person in the interview behaved in a negative way, categorical information became accessible again. Thus, it seems that motivational factors can trigger reinstatement of accessibility.

Combined Effects of Recency and Frequency of Priming

Higgins and colleagues (1985) examined the relation between recent and frequent priming by priming two alternative constructs that were equally applicable to an ambiguous behavior (see also Bargh, Bond, Lombardi, & Tota, 1986). One construct was primed four times, while the other was primed only once at the very end of the priming task. Before the ambiguous target was introduced, there was either a 15- or 120-second delay. Results indicated that recent priming tended to affect the judgment after a short delay, whereas frequent priming tended to affect the judgment after a longer delay (see also Lombardi et al., 1987). These results are consistent with slower decay of frequently primed constructs.

A similar relation was found between chronic accessibility and recent priming. Using a procedure similar to Higgins and colleagues (1982), Bargh, Lombardi, and Higgins (1988) selected participants who had chronic accessibility for the traits "inconsiderate" or "outgoing," participants who had chronic accessibility for both traits, and participants who did not have chronic accessibility for either trait. A few weeks after this assessment, participants were called back to the lab and primed with the nonchronic trait. The delay between priming and presentation of an ambiguous target was also varied. The main finding was a pattern similar to the studies examining recent versus frequent priming, in that participants used the primed construct slightly more than the chronic construct after the short delay (15 seconds); the reverse was true after a longer delay (120 seconds). Recently, Wyer (2004) suggested that frequent and recent activation are qualitatively different. Whereas recent priming is transitory and reflects the resonance of knowledge representation with a probe cue, frequent activation lowers the activation threshold of knowledge representation. The former is said to resonate for awhile but later decays, whereas the latter is said to persist over time. More research is needed to confirm these distinctions.

Combining Accessibility from Different Sources

Extent of priming also may be enhanced by combining priming of different types. We discuss these effects in the following subsections.

CHRONIC ACCESSIBILITY AND SEMANTIC PRIMING

Bargh and colleagues (1986) selected chronic and non-chronic participants with respect to "kindness" and asked them to judge an ambiguously kind person weeks after the assessment. Some participants were subliminally primed with "kind" and some not. Construct-related impressions of the target person were stronger as a function of both chronic and situation-induced accessibility (see also Higgins & Brendl, 1995; see Smith & DeCoster, 1998, for a computer simulation).

PROCEDURAL PRIMING AND SEMANTIC PRIMING

Cumulative effects of procedural and semantic priming have been discussed in the ACT* model by Anderson and colleagues (Anderson, 1983; Pirolli & Anderson, 1985). According to this model, the effects of data strengthening (i.e., practice of a specific target item) and procedural strengthening (i.e., practice of the general process) on perceptual readiness are multiplicative. As noted before, some of the priming procedures may have combined procedural and semantic priming, such as when both priming and target perception involve inferring traits from written materials (Smith & Branscombe, 1987). Thus, it is possible that in these cases accessibility increased because both semantic and procedural priming were combined (see also Hannover & Kühnen, 2005).

Applicability and Accessibility Compensate Each Other

It has been shown that high accessibility of a construct may cause people to apply it to targets of low applicability, targets that would not elicit the construct in other situations. Higgins and Brendl (1995) used a vague (as opposed to an ambiguous) description of a person's behavior, one that was related to conceitedness no more than a no-information baseline. However, this vague description was spontaneously judged as conceited if accessibility was extremely high (e.g., chronics on the dimension that were, in addition, primed with conceitedness and tested very soon after priming). This finding reflects the fact that high accessibility may compensate for low applicability. Beyond its theoretical plausibility, this finding may shed light on the nature of interpersonal discrepancies in the construal of reality. People may tend to see and experience reality in a way that is unique and dependent on their highly accessible constructs, a way that may seem strange and unjustified to others without this view.

It would be interesting to examine the same principle with other types of accessibility. For example, would stronger motivation mean that a wider group of targets would be judged as goal related and increase in accessibility? The literature on learning seems to provide a positive answer to the first question—stronger reinforcements (i.e., a higher motivation) tend to produce a wider generalization over stimuli (Miller, 1944). This principle has far-reaching implications, suggesting the possibility that stronger goals may be substituted with less relevant stimuli. For example, would increasing the intensity of an aggressive motive (e.g., by thwarting more important goals of an individual) make the actor more likely to aggress toward weakly associated targets? It would be worthwhile also to examine whether high procedural accessibility produces transfer of the procedure to more remote, less relevant situations. These principles await further empirical examination.

METACOGNITIVE PROCESSES IN KNOWLEDGE ACTIVATION AND USE

In many cases, people do not know that their judgments and perceptions have been influenced by a prime. Sometimes, however, when they do learn about the possibility of such an influence, they might try to correct it, because they may regard the influence as illegitimate or wrong (i.e., if they believe that the accessible knowledge is unusable). For such a process, motivation and capacity to perform the correction are necessary. Correction often (but not always) results in a contrast effect (see "Effects of Accessibility"), but such contrast should be distinguished from a contrast that results from anchoring. To emphasize, the crux of the difference is that correction, by definition, is a second-stage, metacognitive process, one that operates on a presumed previous influence of a prime, whereas anchoring is a primary cognitive process, one that may happen by default (Decoster & Claypool, 2004). In this section we discuss secondary, correction processes but not anchoring processes (which were discussed earlier). We first review the conditions under which correction occurs and then discuss more general principles of metacognitive correction processes.

Awareness of the Influence of Priming

Awareness of a prime has been shown to result in a contrast effect, most likely because known primes might cause the prime to be seen as irrelevant (Lombardi et al., 1987). Likewise, blatant priming is more likely to create a contrast than subtle priming, possibly for the same reason (Martin, 1986; Newman & Uleman, 1990).

Source monitoring explanations have been advanced to explain effects of awareness on correction. Specifically, Martin (1986) argued that when priming is subtle, people may be unaware of the fact that the primed construct has been activated, and when later confronted with the target of judgment, they may commit a source monitoring error in assuming that the activated construct is their genuine reaction to the target (see Shrum, Wyer, & O'Guinn, 1998). By contrast, when blatantly primed, people are likely to remember the priming episode and be aware of the source of the activation. In this situation, people might judge it to be an inappropriate influence on their perception of the target and try to correct for this influence. Similar models have been proposed for the influence of affect (Schwarz & Clore, 1983) and body feedback (Förster, 2004) on judgments. Interestingly, Huber and colleagues (2001) recently proposed a somewhat similar source monitoring model to account for facilitation and inhibition effects of semantic priming in cognitive psychology.

Like awareness, memory of the prime also enhances the likelihood of a contrast effect. To test this hypothesis, Lombardi and colleagues (1987) divided participants into two groups—those who remembered and those who did not remember the priming events. They found assimilation for those who did not remember the primes and contrast for those who did remember. More directly, Strack, Schwarz, Bless, Kübler, and Wänke (1993) reminded some participants of the priming episode but did not remind others, and again found that the former produced contrast whereas the latter produced assimilation effects.

However, awareness of the prime is not necessary for a correction process to occur. This should come as no surprise, for repeatedly performed correction, as any overlearned process, may become automatic. For example, it has been shown that participants low on racism may automatically perform correction of stereotype-related primes in making judgments about stereotyped targets (for a review, see Devine & Monteith, 1999). Presumably, correction becomes automatic as a result of extensive practice of it in similar situations.

Motivation to Correct Judgments

Awareness of the prime is not sufficient for correction to occur, as people also need to be motivated to correct. Kruglanski (1989) proposed that people who chronically or temporarily avoid closure might be more motivated to inhibit highly accessible knowledge before responding. Indeed, research demonstrated that assimilation effects of priming are weaker for participants high in the need to avoid closure (Ford & Kruglanski, 1995; Thompson, Roman, Moskowitz, Chaiken, & Bargh, 1994). Similar effects were found with interpersonal differences in need for cognition: People high in need for cognition were theorized to possess an enhanced motivation to examine their judgments and feel more accountable for them. In line with this prediction, judgments were found to be less prone to assimilative effects of priming, presumably from an enhanced tendency to correct effects of priming (Martin, Seta, & Crelia, 1990, study 3). A high need to avoid closure and a high need for cognition may represent a general motivation to correct for effects of priming or currently accessible knowledge. We should note that such motivation does not need to be conscious in order to produce the effects.

In addition, there may be more specific motivations for people to correct suspected effects of priming. Correction may occur if the person finds the suspected effect of priming irrelevant for the judgment at hand, morally wrong, culturally objectionable, or socially undesirable, or if he or she thinks that such influence should be suppressed due to conversational norms. Regardless of the reason for the judged uselessness, the result is an attempt to correct for any influence.

For example, Devine (1989) showed that weakly prejudiced and highly prejudiced people have equal knowledge about the African American stereotype but nevertheless respond differentially to a stereotype-related prime if made aware of its activation and if resources to control for its assumed influences are available. If unaware of the prime, however, both groups show enhanced accessibility of stereotype-related constructs. Thus, it seems that people with low levels of prejudice know the stereotypes but refrain from using them when possible (see also Biernat & Vescio, 1993; Higgins & King, 1981, study 1). In that way, contemporary models of prejudice distinguish possession of stereotypical knowledge from endorsement of those stereotypes (see, e.g., Lepore & Brown, 1997; Plant & Devine, 1998).

Cognitive Capacity

Even if a person suspects an effect of priming and thinks that that effect should be corrected, he or she might not have enough cognitive capacity to do so. Martin, Seta, and Crelia (1990) conducted a priming experiment in which they manipulated motivation to correct the priming effect and cognitive load. They showed that assimilation occurred when either motivation or cognitive resources were low, but contrast occurred when both were high (see DeCoster & Claypool, 2004, for a meta-analysis on cognitive capacity as a moderator).

Furthermore, correction processes may differ in cognitive complexity. Strack and Hannover (1996; see Strack, 1992) distinguished between *recomputation* and *adjustment*. In the former case, a person goes back to the information or seeks new information to arrive at her judgment (e.g., "I gave the female student a C in math. I am probably prejudiced, so let me look at the exam again."). In the latter case, a person simply adjusts the judgment (e.g., "let me raise the student's grade to a C+."). Because recomputation is more effortful, it is more dependent on cognitive resources. Thus, people may resort to the simpler process of adjustment when resources are constrained. Yet, for adjusting a judgment (not recomputation), one needs to know both the direction and strength of the contaminating effect. Because such knowledge is often unavailable, adjustment might be problematic, as discussed next.

Knowing How to Correct

Even if people are aware of the prime, want to correct for its influence, and have the capacity to do so, they still may be incorrect about how to correct or not know *how* to do it. Cognitive processes, including the effects of priming, are not accessible to introspection, meaning that people cannot know via introspection what happened in their mind any more than they could know about what happened in their lungs (see Nisbett & Willson, 1977). Certainly, sometimes we are aware of how our mental system operates, but the source of this knowledge is not direct experience with the system at work but rather theories, supplemented with vague and often partial experiences. Similarly, we might have theories about our lungs and might try to imagine how they work if we inhale smoke and then cough. This, however, is not knowledge from direct experience but rather a theory that may not directly reflect reality. We shall refer to people's theories of how their minds work as metacognitive theories—

theories about cognition (see also Strack & Förster, 1998).

The source of metacognitive theories might be culturally shared truisms (e.g., "Commercials are trying to convince you to buy products."), ideologies (e.g., "It is wrong to discriminate by race."), or personal knowledge (e.g., "My best decisions are made on the fly."). Often, these are general theories that specify tendencies, allow for exceptions, and can be applied to specific cases. For example, a person might experience a sudden urge to buy a coke, wonder why that might be, and try to think whether she could have been influenced by a commercial recently seen on TV. She might invoke the reasoning that commercials influence people's behavior and therefore try to resist the urge to buy a coke. It is possible to apply a correct theory to specific instances in a wrong way. For example, in a specific case, the urge to buy a coke could have been there even without the commercial. In addition, a person might endorse a number of metacognitive theories, some of which might have opposite implications in a particular situation (e.g., commercials influence people's behaviors, but I have stable preferences for soft drinks that will not be affected).

Strack (1992) emphasized that in order for correction to yield accurate judgments, a person must know both the direction *and* the magnitude of the influence of the prime, which are also part of one's metatheories about the influence of primes. If a person (erroneously) thinks he was influenced in the opposite direction than the actual effect, an assimilation effect might occur even though he was aware and tried to correct for the influence. In a series of studies, Petty and Wegener (1993) demonstrated that the direction of correction depended on people's beliefs about the putative influence.[5] Thus, an assimilation effect can be produced by both correction and unconscious influences of accessible knowledge.

WHAT IS THE FUNCTION OF ACCESSIBILITY EFFECTS?

Functional Aspects of Semantic and Procedural Priming

Enhanced accessibility after using a construct is most likely functional for facilitating perception of related constructs, which would be beneficial if redundancy in the environment is expected (Neely, 1991). For example, it is functional that processing the word "nurse" would enhance the accessibility of the construct "doctor" because in many contexts these two stimuli tend to co-occur. Manipulation of the expectation of co-occurrence has yet to be investigated, but it may be that semantic priming is reduced when target descriptions contain disconnected sentences or word lists rather than more coherent texts.

In addition, decay rates may also be affected by functional considerations, for example, if faster decay occurs in contexts in which the expectation of redundancy means a more frequent occurrence of a related target. To illustrate, if we see an ant, another one might be within a small radius; however, if we see a mouse, another one might be around, but not within such a small radius. One would expect steeper decay of ants than of mice. Likewise, procedural priming may be functional if an expectation of repetition is assumed. If actions require repetition, it would be useful to retain the knowledge of how to perform an action after completing it. Savings in relearning may be interpreted in a similar way, as resting on the assumption that a once-learned action is likely to be needed in future situations.

Functional Aspects of Accessibility from Affect

As discussed earlier, affective priming can be thought of as having two components: (1) response facilitation of congruent stimuli (approach or affirmation and positive or avoidance or negation and negative), and (2) semantic priming, thereby facilitating processing, of other evaluative-congruent stimuli. Perhaps it is useful to consider the functionality of these components separately.

Response facilitation may be justified if stimuli of similar valence tend to appear in spatial and temporal proximity to each other; environments are somewhat valence redundant (e.g., "lake" signifies the presence of plants and of fish, which all have positive valence; "night" signifies dark, cold, and danger, which all have negative valence). If all positive stimuli elicit approach and all negative stimuli elicit avoidance, and if stimuli of similar valence tend to co-occur, then it is functional to preserve the response from a focal stimulus to other spatially and temporally proximal stimuli.

It is more difficult to justify semantic affective priming on logical grounds: Why would stimuli facilitate the processing of all other stimuli of equal valence? It appears that this would create activation of too many constructs. However, adopting a self-regulatory perspective, selective attention to valence-compatible cues makes sense. Similar to cognitive tuning models (e.g., Schwarz & Bless, 1991), one may assume that negative cues signal danger that needs to be eliminated. Thus, attending selectively to negative cues might facilitate goal attainment. Positive cues, on the other hand, might signal a benign environment, a state that needs to be maintained. Thus, in this case selective encoding of positive cues facilitates goal attainment.

Functional Aspects of Accessibility from Active Goals

It is reasonable to assume that a heightened accessibility of goal-related constructs helps in detection of stimuli in the environment that are instrumental for efficient goal pursuit, which contributes to the likelihood of goal attainment (see Ach, 1935; Gollwitzer, 1999; Kuhl, 1983; Liberman & Förster, 2005). For example, while searching for something (e.g., keys), it would be useful to activate concepts related to the target, such as its typical or likely locations (e.g., purse and bedroom) or typical activities for which the target is used (e.g., access to home and work), because the search target is likely to be in proximity to the associated ob-

jects or activities (e.g., the keys are likely to be in your purse on your desk at work). As long as the related objects and activities are activated, one is more likely to think of places where the keys are likely to be. It is therefore functional to retain a state of heightened activation of goal-related constructs until the goal is fulfilled. What is more, it is easy to see the functionality of suppressing competing goals (e.g., sending a memo to one's boss) for the successful pursuit of the focal goal (e.g., finding the keys). Without such suppression, people would be easily sidetracked by competing goals and would lose sight of a more pressing goal.

Furthermore, inhibition of the search target upon goal fulfillment is also functional because goal-related constructs can interfere with other tasks. Clearing them from the system allows other goals to be more easily pursued after fulfillment. In cybernetic models of goal hierarchies (Carver & Scheier, 1999; Vallacher & Wegner, 1987), goal attainment (e.g., watering plants) is followed by a reinstatement of a higher-order, superordinate goal (e.g., arranging the room for a romantic dinner) that was the reason for pursuing the original, subordinate goal. In this view, constructs related to the fulfilled goal (e.g., water, watering can, and fertilizer) are irrelevant and could potentially interfere with successful performance of the next task at hand (e.g., cooking or cleaning). "Clearing up" the mental system from such constructs thus may be highly instrumental (see also Gollwitzer, 1999; Kuhl, 1983; Liberman & Förster, 2005; Mayr & Keele, 2000).

Recently, Hedberg and Higgins (2005) showed that decay curves of accessibility after goal fulfillment are moderated by chronic accessibility of promotion (accomplishments, growth, ideal) versus prevention (security, responsibility, oughts) concerns. For predominant promotion-focused participants, decay was fast, whereas for predominant prevention-focused participants, accessibility increased after goal fulfillment. This was presumably the case because in a promotion focus, individuals pursue goals by ensuring engagement with means of advancement and ensuring against missing opportunities. After goal completion, a decreased accessibility of goal-related concepts frees up cognitive resources for new, upcoming pursuits, thereby supporting promotion goal pursuit strategies. In a prevention focus, individuals pursue goals by ensuring rejections of mistaken paths to the goal and against committing such mistakes. After goal completion, and before any necessity to pursue another goal, a maintained high accessibility of goal-related concepts preserves cognitive resources for the activity that the person has carried out successfully so far, and thus supports prevention goal pursuit strategies. In general, these studies reflect a dynamic mental system that is able to prepare action without people being aware of the processes.

It is interesting to think from a functional perspective about the proportional relationship between motivational strength and goal-related accessibility and post-fulfillment inhibition. We believe that a high motivation to fulfill a goal means that the individual is ready to invest energy in goal pursuit, giving it a high priority over other goals. Enhancing the accessibility of goal-related con-

structs could help in fulfilling the focal goal, but at the same time, it could interfere with achieving other goals (Shah et al., 2002). For example, enhanced accessibility of "keys" and its associates could help in finding the keys, but it would probably detract attention from simultaneously pursuing other goals that require attention (e.g., engaging in a conversation). Thus, because of its potential costs, enhanced accessibility of goal-related constructs should be proportional to the emphasis one is ready to assign to the pursuit of this particular goal (i.e., to motivation).

It is also reasonable to assume that highly motivating goals are more likely than less motivating goals to suppress (or put "on hold") other important alternative goals. For example, a high motivation to respond to sexist remarks in a conversation with a colleague may suppress the important goal of self-presentation, but a weak motivation would not suffice to produce such suppression. Therefore, as a rule, turning to the next goal pursued after completion of a previous goal is likely to be more important if the previously fulfilled goal was highly motivating. Attending to this important goal is then facilitated by strong inhibition of the constructs related to the previous goal. Thus, the more motivating the initial goal was, the more functional it is to inhibit its related constructs after goal fulfillment in order to turn back to the important alternative goal that was suppressed.

In sum, it is easy to see how a system that incorporates priming effects of different types has evolved from a functional need. It is also possible, however, that functionality is not only an evolutionary explanation of how these effects came to exist but also an important factor that influences the effects of priming in humans today. Thus, we believe that the notion that priming effects are functional is not only interesting in its own right but also serves to derive interesting predictions regarding possible situational moderators of priming effects.

APPLICATIONS OF PRINCIPLES OF ACCESSIBILITY IN SOCIAL PSYCHOLOGY

Lessons learned from priming research were applied to many content domains in social psychology, including study of cross-cultural differences (Hannover & Kühnen, 2005), consumer psychology (Wyer, in press), and research on basic motives, such as terror management (Pyszczynski, Greenberg, & Solomon, 1999). In the present chapter, however, we chose to focus on three other domains: measurement of stereotypes, measurement of goals and concerns, and uncovering the processes involved in thought suppression.

Measurement of Attitudes: Using Affective Priming

Because affective priming facilitates processing of evaluative-congruent targets, Fazio and Williams (1986) suggested that the extent of such priming should be stronger with primes for which strong attitudes exist. If so, the extent to which a stimulus primes positive versus

negative targets should reveal the strength of the individual's positive (vs. negative) attitude toward the prime. This notion has been put to use in modern social psychology. For example, Dunton and Fazio (1997) primed participants with African American versus Caucasian faces and examined the speed with which participants could classify adjectives as positive versus negative. They found that the extent to which African American faces primed negative adjectives was indicative of their attitudes toward African Americans. Subsequent studies found that this measure also correlated with behavior. For example, it predicted students' anticipated comfort in interacting with African American students (Towles-Schwen & Fazio, 2003) and maintenance of eye contact with an African American interviewer (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). Importantly, this method, often termed an implicit measure of attitudes, does not measure explicitly expressed attitudes, which often correlate only with controllable behaviors. For example, explicit attitudes, as measured by the modern racism scale, predicted explicitly stated evaluations of the African American interviewer but did not predict less controlled behaviors, such as maintaining eye contact with him or her. The reverse was true for implicit attitudes, which were measured as the extent of affective priming (Dovidio et al., 1997; for a review, see Olson & Fazio, 2002). Affective priming has been used to measure attitudes toward social groups as well as toward other attitude objects, including smoking (Sherman, Rose, Koch, Presson, & Chassin, 2003) and TV programs (Frings & Wentura, 2003).

There are other ways to measure attitudes implicitly, primarily with the Implicit Associations Test (IAT; Greenwald, McGee, & Schwartz, 1998; see Olson & Fazio, 2002, for a review of the differences between the measures). Importantly, implicit measures of attitudes are not invariant; they are sensitive to many factors, including context, motivation, and practice. This should not be surprising in view of the fact that all priming effects are sensitive to these factors.

It is beyond the scope of this chapter to review all the ways in which implicit measures of attitudes differ from explicit measures. One aspect particularly relevant to the principles of knowledge accessibility is that explicit attitudes are more susceptible to control and correction processes than implicit attitudes. Therefore, discrepancies between implicitly and explicitly measured attitudes should exist when correction processes are likely to operate. In other words, when people suspect they were biased, are motivated to correct the bias, have the capacity to perform the correction, and posses the knowledge of how to do it, differences between implicit and explicit attitudes should be apparent. The case of racial attitudes is a prime example (Olson & Fazio, 2002), but other examples include attitudes toward objects of temptations (e.g., smoking and partying) and ambivalence. For example, if one considers visiting museums important but boring, the concept of "museums" will be ambivalent for the person and would likely elicit discrepancies between implicit and explicit attitude measures.

Measuring Goals and Concerns: Using Goal Accessibility and Chronic Accessibility

We have already shown in this chapter that active goals and concerns enhance the accessibility of related constructs. This principle implies that accessibility may be used to uncover goals and concerns. Within regulatory focus theory (Higgins, 1987, 1997), this idea was used to measure the strength of dispositional promotion focus (the strength of one's hopes and aspirations) and dispositional prevention focus (the strength of one's duties and obligations). Specifically, Higgins, Shah, and Friedman (1997) developed a computerized measure that examines the time it takes participants to list "ideals" and "oughts," using this information to evaluate one's standing on them (see Higgins et al., 1997; Shah & Higgins, 1997). A faster overall response time for ideals indicates a stronger concern for them and indicates a promotion focus, while a faster time for "oughts" indicates a prevention focus. This personality measure has been used widely in research (e.g., Förster et al., 1998; Förster, Higgins, & Strack, 2000; Higgins et al., 1997; Shah & Higgins, 1997, to name a few).

Accessibility may be used to assess other dispositional concerns as well. To give some examples, accessibility of depressive contents can be used as a measure of depression or accessibility of achievement-related contents as a measure of achievement concerns. In fact, traditional projective tests, such as the Thematic Apperception Test (TAT; Murray, 1938), may be viewed as tests of accessibility (Sorrentino & Higgins, 1986). It may be contended that these tests operationalize accessibility as ease of retrieval from memory, or perceptual readiness and high accessibility as indicative of a high concern with the contents in question. For example, a high concern with achievement is revealed in the TAT as a high frequency of achievement-related themes in one's interpretations of scenes in pictures (Atkinson, 1954; McClelland, 1953). We should note that from the perspective of contemporary social and cognitive psychology, what often appear to be problematic in projective tests are not the assumptions about the connection between accessibility and people's concerns. Rather, it is the assumption that such concerns are necessarily reflective of a stable personality characteristic and the lack of objectivity in coding the contents of the participants' responses.

Suppression of Thoughts: Using Principles of Goal-Related Accessibility to Uncover Motivations

It is hard, if not impossible, to suppress thoughts. Moreover, suppressed thoughts rebound after suppression and become more frequent than without suppression (Wegner, 1994). Macrae, Bodenhausen, Milne, and Jetten (1994) suggested that this rebound effect is due to high accessibility caused by priming of the suppressed construct during the suppression stage. In their study, they measured accessibility of skinhead stereotypes using a lexical decision task after asking participants to sup-

press skinhead stereotypes. Those asked to suppress showed higher accessibility after suppression when compared to those who did not engage in suppression. More recently, Liberman and Förster (2000) conceptually replicated the effects and added a condition in which participants expressed the stereotype after suppression. For example, in one of the studies they asked participants, who first suppressed stereotypes of African Americans in a story they wrote, to take the perspective of a racist and write a second story using racial stereotypes. It was found that participants who first suppressed stereotypes and then expressed them, inhibited racist constructs. The authors suggested the motivational inference model (MIMO) of postsuppressional rebound, which claims that suppression instructions introduce or enhance a motivation to think about the suppressed thought. Expression after suppression fulfills this motivation and is followed by a postfulfillment inhibition. Without prior motivation to use the construct, construct use would have enhanced the accessibility of the construct (see Förster & Liberman, 2004, in press). The finding that expression of suppressed thoughts reduced, instead of enhanced, accessibility speaks for the activation of goals rather than semantic priming, which would increase from two sources of accessibility (suppression and expression; for further evidence and additional implications of the model, see Förster & Liberman, 2001). These studies allowed for clarification of the nature of suppression. Previous models viewed enhanced accessibility of a suppressed construct as accessibility due to construct use, whereas MIMO explains it as goal-related accessibility; empirical evidence has shown support for the latter view.

CONCLUSION

There is abundant research on the effects of knowledge accessibility on affect, cognition, and behavior. Theories about these effects are also slowly emerging. We believe that there is much to gain in considering the similarities between different effects and by applying general, overarching principles to explain them. This chapter has tried to examine some of the findings on knowledge accessibility and principles that these findings instantiate. We also discussed how these phenomena and principles may be used to examine, extend, and test models in social psychology. Future research will have to develop integrative models to explain the effects and, in particular, include a discussion of functional aspects of accessibility.

ACKNOWLEDGMENTS

NOTES

1. This definition diverges from Wilson and Brekke's definition of mental contamination as undesirable or subjectively wrong influences on processing of targets. In our definition of priming, priming effects occur whether or not a person considers such effects desirable, as they are simply independent of such judgment. In other words, a desirable effect of priming (e.g., when a person does want to yield to the effects of prior mental processes on subsequent processing of a target) would not be considered a mental contamination, but it would, nonetheless, remain an effect of priming. Of course, undesirable demonstrations of priming effects are especially strong because they rule out conscious, strategic, and deliberate processes.
2. This does not mean that there won't be any difference between negated or affirmed primes, or between priming by a credible versus discredited source. These differences, however, may be related to secondary, metacognitive correction processes, rather than to the primary effects of accessibility, as discussed in the fourth part of this chapter.
3. This effect is similar to Srull and Wyer's (1980) finding in that priming "hostility" influenced a target's person behavior only if the concepts were primed before the behavioral information and had no effect when they were primed afterward.
4. One may wonder why Dijksterhuis et al. (2000) and Levy (1996) obtained different results after subliminally priming young participants with the elderly stereotype. Seibt (2003) suggests that one clue to the apparent contradiction are the primes used in the two studies—Dijksterhuis et al.'s primes were neutral (e.g., old and gray), whereas Levy's primes were either negative and related to forgetfulness (e.g., Alzheimer's, senile, dementia, and diseased) or positive and irrelevant to memory (e.g., guidance, wise, and alert). It is possible that in Levy's studies, the younger participants were either repelled by the negative valence of the primes and contrasted themselves away from them (in the negative condition) or did not associate the primes with memory performance (in the positive condition).
5. Strack, Förster, and Werth (2005) were able to change people's memory judgments by suggesting opposite theories of contaminating influences.

REFERENCES

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology, 78,* 53–63.

Aarts, H., Gollwitzer, P. M., & Hassin, R. R. (2004). Goal contagion: perceiving is for pursuing. *Journal of Personality and Social Psychology, 87,* 23–37.

Ach, N. (1935). Analyse des Willens. In E. Abderhalden (Ed.), *Handbuch der biologischen Arbeitsmethoden, Bd. VI.* Berlin: Urban & Schwarzenberg.

Andersen, S. M., & Chen, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review, 109,* 619–645.

Andersen, S. M., & Glassman, N. S. (1996). Responding to significant others when they are not there: Effects on interpersonal inference, motivation, and affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 3. The interpersonal context* (Vol. 3, pp. 262–321). New York: Guilford Press.

Anderson, J. R. (1974). Retrieval of propositional information in long-term memory. *Cognitive Psychology, 6,* 451–474.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language*. San Diego, CA: Academic Press.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237–313). New York: Academic Press.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*, 68–100.

Atkinson, J. W. (1954). Exploration using imaginative thought to assess the strength of human motives. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (pp. 56–112). Lincoln: University of Nebraska Press.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*, 272–281.

Bargh, J. A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality and Social Psychology*, *43*, 425–436.

Bargh, J. A. (1994a). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (pp. 1–61). Mahwah, NJ: Erlbaum.

Bargh, J. A. (1994b). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., & Barndollar, K. (1996). Automaticity in action: The unconscious as repository of chronic goals and motives. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 457–481). New York: Guilford Press.

Bargh, J. A., Bond, R. N., Lombardi, W. J., & Tota, M. E. (1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology*, *50*, 869–878.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 109–128.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Experimental Social Psychology*, *71*, 230–244.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Troetschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*, 1014–1027.

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically accessible constructs in person x situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology*, *55*, 599–605.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437–449.

Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology*, *22*, 293–311.

Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology*, *49*, 1129–1146.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1059–1082.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, *39*, 15–22.

Biernat, M., & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology*, *29*, 166–202.

Blanton, H. (2001). Evaluating the self in the context of another: Assimilation and contrast effects in social comparison. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium and the Legacy and Future of Social Cognition* (pp. 75–88). Mahwah, NJ: Erlbaum.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*, 123–152.

Burtt, H. E. (1941). An experimental study of early childhood memory: Final report. *Journal of Genetic Psychology*, *58*, 435–439.

Butterfield, E. C. (1964). The interruption of tasks: Methodological, factual and theoretical issues. *Psychological Bulletin*, *62*, 309–322.

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*, 840–856.

Carver, C. S., & Scheier, M. F. (1999). Themes and issues in the self-regulation of behavior. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 12, pp. 1–106). Mahwah, NJ: Erlbaum.

Cesario, J., Plaks, J. E., & Higgins, E. T. (2006). Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, *90*, 893–910.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Chartrand, T. L., Maddux, W., & Lakin, J. (in press). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *Unintended thought II: The new unconscious*. New York: Oxford University Press.

Chen, S., & Andersen, S. M. (1999). Relationships from the past in the present: Significant-other representations and transference in interpersonal life. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 31, pp. 123–190). San Diego, CA: Academic Press.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

DeCoster, J., & Claypool, H. M. (2004). A meta-analysis of priming effects on impression formation supporting a general model of informational biases. *Personality and Social Psychology Review*, *8*, 2–27.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Devine, P. G., & Monteith, M. J. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 339–360). New York: Guilford Press.

Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, *36*, 531–544.

Dijksterhuis, A., Bargh, J. A., & Miedema, J. (2000). Of men and mackerels: Attention, subjective experience, and automatic social behavior. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 37–51). Philadelphia: Psychology Press.

Dijksterhuis, A., Spears, R., & Lepinasse, V. (2001). Reflecting and deflecting stereotypes: Assimilation and contrast in impression formation and automatic behavior. *Journal of Experimental Social Psychology*, *37*, 286–299.

Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D. A., Koomen, W., Van Knippenberg, A., et al. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology*, *75*, 862–871.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology, 74*, 865–877.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*(5, Whole No. 270).

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*, 316–326.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York: Dover. (Original work published 1885)

Epley, N., & Gilovich, T. (1999). Just going along: Nonconscious priming and conformity to social pressure. *Journal of Experimental Social Psychology, 35*, 578–589.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62*, 328–339.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 204–243). New York: Guilford Press.

Fazio, R. H. (2000). Accessible attitudes as tools for object appraisal: Their costs and benefits. In G. Maio & J. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 1–36). Mahwah, NJ: Erlbaum.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.

Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology, 51*, 505–514.

Fiedler, K. (2003). The hidden vicissitudes of the priming paradigm in evaluative judgment research. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (p. 109–137). Mahwah, NJ: Erlbaum.

Fiedler, K., & Schenck, W. (2001). Spontaneous inferences from pictorially presented behavior. *Personality and Social Psychology Bulletin, 27*, 1533–1546.

Fiedler, K., Schenck, W., Watling, M., & Menges, J. I. (2005). Priming trait inferences through pictures and moving pictures: The impact of open and closed mindsets. *Journal of Personality and Social Psychology, 88*, 229–244.

Fishbach, A., Friedman, R. S., & Kruglanski, A. W. (2003). Leading us not into temptation: Momentary allurements elicit overriding goal activation. *Journal of Personality and Social Psychology, 84*, 296–309.

Fishbein, M., & Ajzen, I. (1974). Attitudes toward objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*, 59–74.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Ford, T. E., & Kruglanski, A. W. (1995). Effects of epistemic motivations on the use of accessible constructs in social judgments. *Personality and Social Psychology Bulletin, 21*, 950–962.

Förster, J. (2004). How body feedback influences consumer's evaluation of products. *Journal of Consumer Psychology, 14*, 415–425.

Förster, J., & Denzler, M. (in press). Selbstregulation [Self-regulation]. In D. Frey (Ed.), *Handbuch der Sozialpsychologie*. Berlin: Hogrefe.

Förster, J., Friedman, R., Butterbach, E. M., & Sassenberg, K. (2005). Automatic effects of deviancy cues on creative cognition. *European Journal of Social Psychology, 35*, 345–360.

Förster, J., Friedman, R., Özelsel, A., & Denzler, M. (2006). Enactment of approach and avoidance behavior influences the scope of perceptual and conceptual attention. *Journal of Experimental Social Psychology, 42*, 133–146.

Förster, J., & Higgins, E. T. (2005). How global vs. local processing fits regulatory focus. *Psychological Science, 16*, 631–636.

Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the "goal looms larger" effect. *Journal of Personality and Social Psychology, 75*, 1115–1131.

Förster, J., Higgins, E. T., & Strack, F. (2000). When stereotype disconfirmation is a personal threat: How prejudice and prevention focus moderate incongruency effects. *Social Cognition, 18*, 178–197.

Förster, J., Kuschel, S., & Liberman, N. (2005). *How processing styles affect assimilation and contrast in social judgment.* Unpublished data set, International University Bremen.

Förster, J., Kuschel, S., & Liberman, N. (2006). *The effect of global vs. local processing styles on assimilation vs. contrast in social judgment*. Manuscript submitted for publication.

Förster, J., & Liberman, N. (2001). The role of attribution of motivation in producing post-suppressional rebound. *Journal of Personality and Social Psychology, 81*, 377–390.

Förster, J., & Liberman, N. (2004). How motivational inferences influence post-suppressional rebound. In S. Shohov (Hrsg.), *Advances in psychology research* (Vol. 34, pp. 63–88). New York: Nova Science.

Förster, J., & Liberman, N. (2005). A motivational model of post-suppressional rebound. *European Review of Social Psychology, 15*, 1–32.

Förster, J., & Liberman, N. (in press). Inhibition processes in comparisons. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology*. New York: Psychology Press.

Förster, J., Liberman, N., & Friedman, R. (in press). *The seven principles of self regulation*. In J. Bargh & P. Gollwitzer (Ed.), *The psychology of action and goals*. New York: Guilford Press.

Förster, J., Liberman, N., & Higgins, E. T. (2005). Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology, 41*, 220–239.

Friedman, R., Fishbach, A., Förster, J., & Werth, L. (2003). Attentional priming effects on creativity. *Creativity Research Journal, 15*, 277–286.

Friedman, R. S., & Förster, J. (2001). The effects of promotion and prevention cues on creativity. *Journal of Personality and Social Psychology, 81*, 1001–1013.

Friedman, R. S., & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology, 38*, 41–55.

Friedman, R., McCarthy, D. M., Förster, J., & Denzler, M. (2005). Automatic effects of alcohol cues on sexual attraction. *Addiction, 100*, 672–681.

Frings, C., & Wentura, D. (2003). Who is watching Big Brother? TV consumption predicted by masked affective priming. *European Journal of Social Psychology, 3*, 779–791.

Fujita, K., Trope, Y., & Liberman, N. (2005). *Construal levels and self-control*. Manuscript submitted for publication.

Gilovich, T. (1981). Seeing the past in the present: The effect of associations of familiar events on judgment and decision. *Journal of Personality and Social Psychology, 40*, 797–808.

Gollwitzer, P. M. (1990). Action phases and mindsets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Volume 2. Foundations of social behavior* (pp. 53–92). New York: Guilford Press.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493–503.

Gollwitzer, P. M., Heckhausen, H., & Steller, B. (1990). Deliberative and implemental mindsets: Cognitive tuning toward congruous thoughts and information. *Journal of Personality and Social Psychology, 59*, 1119–1127.

Goschke, T., & Kuhl, J. (1993). Representation of intentions: Persisting activation in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1211–1226.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit

association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Grice, H. P. (1975). Logic and conservation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics. 3: Speech acts* (pp. 225–242). New York: Academic Press.

Gupta, P., & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review*, *109*, 401–448.

Hannover, B., & Kühnen, U. (2005). Culture, context, and cognition: The semantic procedural interface model of the self. *European Review of Social Psychology*, *14*, 297–333.

Hassin, R. R., & Bargh, J. A. (2005). *Flexibility and rigidity in automatic goal pursuit*. Manuscript submitted for publication.

Hassin, R. R., Uleman, J.S., & Bargh, J. A. (Eds.). (2005). *The new unconscious*. New York: Oxford University Press.

Hebb, D. O. (1948). *Organization of behavior*. New York: Wiley.

Heckhausen, H. (1991). *Motivation and action*. Berlin: Springer.

Hedberg, P., & Higgins, E. T. (2005). *What remains in your mind after you are done?* Unpublished data set, Columbia University.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion*, *15*, 143–165.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94, 319–340.*

Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought: The limits of awareness, intention and control* (pp. 75–123). New York: Guilford Press.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T., Bargh, J. A., & Lombardi, W. (1985). The nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 59–69.

Higgins, E. T., & Brendl, C. M. (1995). Accessibility and applicability: Some "activation rules" influencing judgment. *Journal of Experimental Social Psychology*, *31*, 218–243.

Higgins, E. T., & Chaires, W. M. (1980). Accessibility of interrelational constructs: Implications for stimulus encoding and creativity. *Journal of Experimental Social Psychology*, *16*, 348–361.

Higgins, E. T., & King, G. (1981). Accessibility of social constructs: Information processing consequences of individual and contextual variability. In N. Cantor & J. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 69–121). Hillsdale, NJ: Erlbaum.

Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, *43*, 35–47.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Higgins, E. T., Shah, J., & Friedman, R. S. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, *72*, 515–525.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Quach, R. (2002). Mechanisms of source confusion and discounting in short-term priming. 2: Effects of prime similarity and target duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 120–1136.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review*, *108*, 149–182.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology*, *47*, 1206–1217.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 371–391). Mahwah, NJ: Erlbaum.

Klauer, K. C., Rossnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 246–255.

Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology*, *55*(2), 162–170.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.

Kruglanski, A. W. (1996). Motivated social cognition. Principles of the interface. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 331–378). San Diego, CA: Academic Press.

Kuhl, J. (1983). *Motivation, Konflikt und Handlungskontrolle*. New York: Springer.

Kunda, Z. (2000). *Social cognition*. Cambridge, UK: MIT Press.

Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology*, *82*, 283–299.

Lau, R. R. (1989). Construct accessibility and electoral choice. *Political Behavior*, *11*, 5–32.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*, 275–287.

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, *71*, 1092–1107.

Liberman, N., & Förster, J. (2000). Expression after suppression: A motivational explanation of post-suppressional rebound. *Journal of Personality and Social Psychology*, *79*, 190–203.

Liberman, N., & Förster, J. (2005). Motivation and construct accessibility. In J. P. Forgas, K. D. Kipling, & S. M. Laham (Eds.), *Social motivation: Conscious and unconscious processes* (pp. 228–248). Cambridge, UK: Cambridge University Press.

Liberman, N., Förster, J., & Higgins, E. T. (in press). Set/reset or inhibition after goal fulfillment?: A fair test between two mechanisms producing assimilation and contrast. *Journal of Experimental Social Psychology*.

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization. *Personality and Social Psychology Bulletin*, *13*, 411–429.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, *69*, 397–407.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*, 808–817.

Macrae, C. N., & Johnston, L. (1998). Help, I need somebody: Automatic action and inaction. *Social Cognition*, *16*, 400–417.

Macrae, N. C., & Lewis, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science*, *13*, 194–196.

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality and Social Psychology Review*, *7*, 244–267.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, *35*, 63–78.

Marsh, R. L., Hicks, J. L., & Bink, M. L. (1998). Activation of com-

pleted, uncompleted and partially completed intentions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 350–361.

Marsh, R. L., Hicks, J. L., & Bryan, E. S. (1999). The activation of unrelated and canceled intentions. *Memory and Cognition*, *27*, 320–327.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, *59*, 27–37.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" versus "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*, 433–449.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action in role of backward inhibition. *Journal of Experimental Psychology: General*, *129*, 4–26.

McClelland, D. C. (1953) The measurement of human motivation: an experimental approach. *Proceedings of the Conference on Testing Problems, Educational Testing Service*, pp. 41–51.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined? It depends upon who's asking and what is asked. *Journal of Conflict Resolution*, *25*, 563–579.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.

Miller, N. E. (1944). Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 431–465). New York: Ronald Press.

Murray, H. A. (1938). *Exploration in personality*. New York: Oxford University Press.

Mussweiler, T. (2001). "Seek and ye shall find": Antecedents of assimilation and contrast in social comparison. *European Journal of Social Psychology*, *31*, 499–509.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.

Mussweiler, T., & Bodenhausen, G. (2002). I know you are but what am I?: Self-evaluative consequences of judging ingroup and outgroup members. *Journal of Personality and Social Psychology*, *82*, 19–32.

Mussweiler, T., & Förster, J. (2000). The sex–aggression link: A perception-behavior dissociation. *Journal of Personality and Social Psychology*, *79*, 507–520.

Mussweiler, T., Rüter, K., & Epstude, K. (2004a). The man who wasn't there—Subliminal social standards influence self-evaluation. *Journal of Experimental Social Psychology*, *40*, 689–696.

Mussweiler, T., Rüter, K., & Epstude, K. (2004b). The ups and downs of social comparison: Mechanisms of assimilation and contrast. *Journal of Personality and Social Psychology*, *87*, 832–844.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, *35*, 136–164.

Mussweiler, T., & Strack, F. (2000a). The "relative self": Informational and judgmental consequences of comparative self-evaluation. *Journal of Personality and Social Psychology*, *79*, 23–38.

Mussweiler, T., & Strack, F. (2000b). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology*, *78*, 1038–1052.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibition less spreading activation and limited-capacity attention. *Journal of Experimental Social Psychology: General*, *106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Newman, L. S., & Uleman, J. S. (1990). Assimilation and contrast effects in spontaneous trait inference. *Personality and Social Psychology Bulletin*, *16*, 224–240.

Nickerson, R. S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory and Cognition*, *12*, 531–552.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84, 231–259.*

Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition*, *20*, 89–104.

Parducci, A. (1965). Category judgment: A range–frequency model. *Psychological Review*, *72*, 407–418.

Parducci, A., Perrett, D. S., & Marsh, H. W. (1969). Assimilation and contrast as range-frequency effects of anchors. *Journal of Experimental Psychology*, *81*, 281–288.

Park, J., Yoon, S., Kim, K., & Wyer, R. S., Jr. (2001). Effects of priming a bipolar attribute concept on dimension versus concept-specific accessibility of semantic memory. *Journal of Personality and Social Psychology*, *81*, 405–420.

Petty, R. E., & Wegner, D. T. (1993). Flexible correction processes in social judgment: Correcting for context induced contrast. *Journal of Experimental Social Psychology*, *29*, 137–165.

Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 136–153.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Pyszczynski, T., Greenberg, J., & Solomon, S. (1999, October). A dual-process model of defense against conscious and unconscious death-related thoughts: An extension of terror management theory. *Psychological Review*, *106*, 835–845.

Ratcliff, R., & McCoon, G. (1996). Bias effects in implicit memory tasks. *Journal of Experimental Psychology: General*, *125*, 403–421.

Roediger, H. L., & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology*, *36*, 213–242.

Rothermund, K. (2003). Automatic vigilance for task-relevant information: Perseverance after failure and inhibition after success. *Memory and Cognition*, *31*, 343–352.

Rundus, D. (1973). Negative effects of using list items as retrieval cues. *Journal of Verbal Learning and Verbal Behavior*, *12*, 43–50.

Santa, J., & Lamwers, L. L. (1974). Encoding specificity: Fact or artifact. *Journal of Verbal Learning and Verbal Behavior*, *13*, 412–423.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, *16*, 989–997.

Schooler, J. W., Fiore, S. M., & Brandimonte, M. A. (1997). At a loss from words: verbal overshadowing of perceptual memories. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 293–334). Hillsdale, NJ: Academic Press.

Schul, Y., Mayo, R., & Burnstein, E. (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology*, *86*, 668–679.

Schwarz, N., & Bless, H. (1991). Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning. In J. P. Forgas (Ed.), *Emotion and social judgments. International series in experimental social psychology* (pp. 55–71). Elmsford, NY: Pergamon Press.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Schwarz, N., Strack, F., & Mai, H. P. (1991). Assimilation and contrast effects in part–whole question sequences: A conversational-logic analysis. *Public Opinion Quarterly*, *55*, 3–23.

Seibt, B. (2003). *Risky and careful processing under stereotype threat: How performance is influenced by activated self-stereotypes*. Unpublished dissertation, Universität Würzburg.

Shah, J. (2003a). Automatic for the people: How representations of significant others implicitly affect goal pursuit. *Journal of Personality and Social Psychology*, *84*, 661–681.

Shah, J. (2003b). The motivational looking glass: How significant others implicitly affect goal appraisals. *Journal of Personality and Social Psychology*, *85*, 424–439.

Shah, J. Y., Friedman, R., & Kruglanski, A. W. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, *83*, 1261–1280.

Shah, J., & Higgins, E. T. (1997). Expectancy × value effects: Regulatory focus as determinant of magnitude and direction. *Journal of Personality and Social Psychology*, *73*, 447–458.

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How goal pursuit is affected by accessible alternatives. *Journal of Experimental Social Psychology*, *38*, 368–383.

Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: Bottom-up priming of goals by means and its effects on self-regulation. *Journal of Personality and Social Psychology*, *84*, 1109–1122.

Shah, J. Y., Kruglanski, A. W., & Thompson, E. P. (1998). Membership has its (epistemic) rewards: Need for closure effects on in-group bias. *Journal of Personality and Social Psychology*, *75*, 383–393.

Sherman, S. J., Mackie, D. M., & Driscoll, D. M. (1990). Priming and the differential use of dimensions in evaluation. *Personality and Social Psychology Bulletin*, *16*, 405–418.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, *22*, 13–39.

Shimamura, A. P., & Squire, L. R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General*, *113*, 556–570.

Shrum, L. J., Wyer, R. S., & O'Guinn, T. (1998).The effects of watching television on perceptions of social reality. *Journal of Consumer Research*, *24*, 447–458.

Skowronski, J. J., Carlston, D. E., & Isham, J. T. (1993). Implicit versus explicit impression formation: The differing effects of overt labeling and covert priming on memory and impressions. *Journal of Experimental Social Psychology*, *29*, 17–41.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *76*, 504–513.

Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology*, *25*, 500–523.

Smith, E. R. (1990). Content and process specificity in the effects of prior experiences. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 3, pp. 1–60). Hillsdale, NJ: Erlbaum.

Smith, E. R., & Branscombe, N. R. (1987). Procedurally mediated social inferences: The case of category accessibility effects. *Journal of Experimental Social Psychology*, *23*, 361–382.

Smith, E. R., Branscombe, N. R., & Bormann, C. (1988). Generality of the effects of practice on social judgement tasks. *Journal of Personality and Social Psychology*, *54*, 385–395.

Smith, E. R., & DeCoster, J. (1998). Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology*, *74*, 21–35.

Smith, E. R., & Lerner, M. (1986). Development of automatism of social judgments. *Journal of Personality and Social Psychology*, *50*, 246–259.

Smith, E. R., Stewart, T. L., & Buttram, R. T. (1992). Inferring a trait from a behavior has long-term, highly specific effects. *Journal of Personality and Social Psychology*, *62*, 753–759.

Sorrentino, R. M., & Higgins, E. T. (Eds.). (1986). *The handbook of motivation and cognition: Volume 1. Foundations of social behavior*. New York: Guilford Press.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, *37*, 1660–1672.

Srull, T. K., & Wyer, R. S., Jr. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology*, *38*, 841–856.

Stapel, D. A. (in press). In the mind of the beholder: The interpretation comparison model of accessibility effects. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology*. New York: Psychological Press.

Stapel, D. A., & Koomen, W. (2000). Distinctness of others and malleability of selves: Their impact on social comparison effects. *Journal of Personality and Social Psychology*, *79*, 1068–1087.

Stapel, D. A., & Koomen, W. (2001). I, we, and the effects of others on me: How self-construal moderates social comparison effects. *Journal of Personality and Social Psychology*, *80*, 766–781.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1996). The referents of trait inferences: The impact of trait concepts versus actor–trait links on subsequent judgments. *Journal of Personality and Social Psychology*, *70*, 437–450.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1997). Categories of category accessibility: The impact of trait versus exemplar priming on person judgments. *Journal of Experimental Social Psychology*, *33*, 44–76.

Stapel, D. A., Koomen, W., & Velthuijsen, A. S. (1998). Assimilation or contrast?: Comparison relevance, distinctness, and the impact of accessible information on consumer judgments. *Journal of Consumer Psychology*, *7*, 1–24.

Stapel, D., & Suls, J. (2004). Method matters: Effects of explicit versus implicit social comparisons on activation, behavior, and self-views. *Journal of Personality and Social Psychology*, *87*, 860–875.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, *64*, 211–220.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.

Strack, F., & Förster, J. (1998). Self-reflection and recognition: The role of metacognitive knowledge in the attribution of recollective experience. *Personality and Social Psychology Review*, *2*, 111–123.

Strack, F., Förster, J., & Werth, L. (2005). "Know thyself!": Idiosyncratic self-knowledge may influence recognition. *Journal of Memory and Language*, *52*, 628–638.

Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 579–596). New York: Guilford Press.

Strack, F., Martin L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology*, *18*, 429–442.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*, 437–446.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Strahan, E. J., Spencer, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: Striking while the iron is hot. *Journal of Experimental Social Psychology*, *38*, 556–568.

Strauman, T. J., & Higgins, E. T. (1987). Automatic activation of

self-discrepancies and emotional syndromes: When cognitive structures influence affect. *Journal of Personality and Social Psychology*, *53*, 1004–1014.

Swinney, D., Prather, P., & Love, T. (2000). The time-course of lexical access and the role of context: Converging evidence from normal and aphasic processing. In Y. Grodzinsky, L. P. Shapiro, & D. A. Swinney (Eds.), *Language and the brain: Representation and processing* (pp. 273–294). New York: Academic Press.

Taylor, S. E., & Gollwitzer, P. M. (1995). The effects of mindsets on positive illusions. *Journal of Personality and Social Psychology*, *69*, 213–226.

Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S., & Bargh, J. A. (1994). Accuracy motivation attenuates covert priming effects: The systematic reprocessing of social information. *Journal of Personality and Social Psychology*, *66*, 474–489.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review*, *5*.

Todorov, A. (2000). Context effects in national health surveys: Effects of preceding questions on reporting serious difficulty seeing and legal blindness. *Public Opinion Quarterly*, *64*, 65–76.

Towles-Schwen, T., & Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Personality and Social Psychology Bulletin*, *29*, 170–182.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *85*, 1124–1131.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3–15.

Vroom, V. H. (1966). Organizational choice: A study of pre- and post-decision processes. *Organizational Behavior and Human Performance*, *1*, 212–225.

Walther, E., Müller, D., & Schott, O. (2001). Automatisches soziales Verhalten. Wie wirkt sich die Aktivierung der Konzepte Altruismus und Egoismus auf Hilfeleistung aus? *Zeitschrift für Experimentelle Psychologie*, *48*, 248–257.

Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, *19*, 175–196.

Watkins, M. J. (1975). Inhibition in recall with extralist "cues." *Journal of Verbal Learning and Verbal Behavior*, *14*, 294–303.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Wentura, D. (1998). Affective priming in the lexical decision task: Evidence for post-lexical judgmental tendencies. *Sprache and Kognition*, *17*, 125–137.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 456–469.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

Wyer, R. S., Jr. (2004). *Social comprehension and judgment: The role of situation models, narratives, and implicit theories.* Mahwah, NJ: Erlbaum.

Wyer, S. R., Jr. (in press). The role of knowledge accessibility in cognition and behavior: Implications for consumer information processing. In C. Haugvedt, F. R. Kardes, & P. M. Herr (Eds.), *Handbook of consumer research.*

Wyer, R. S., Jr., Adaval, R., & Colcombe, S. J. (2002). Narrative-based representations of social knowledge: Their construction and use in comprehension, memory, and judgment. *Advances in Experimental Social Psychology*, *34*, 131–197.

Wyer, R. S., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychologcial Review*, *106*, 89–118.

Wyer, R. S., & Srull, T. K. (1989). *Memory and cognition in its social context.* Hillsdale, NJ: Erlbaum.

Zeigarnik, B. (1938). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of gestalt psychology* (pp. 300–314). New York: Harcourt, Brace, & World.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168–171.

Causal Explanation

From Social Perception to Knowledge-Based Attribution

DENIS HILTON

Being able to infer causes and communicate reasons provides both humans and other primates with considerable adaptive and organizational advantages. For example, some Diana monkeys in the wild are able to infer, on hearing a chimpanzee's alarm call for *Leopard!*, that a leopard must be present, and utter the Diana-ese warning call for *Leopard!* to their own conspecifics, thus giving them a "reason" to hide (Zauberbühler, 2000a, 2000b). More generally, the automatic recognition of emotional reactions in others can enable animals to make inferences about what caused those reactions and thus learn about the nature of their environment. For example, Blair (2003) argues that emotional expressions serve as unconditioned stimuli that enable a conspecific to evaluate the eliciting stimulus as positive or negative; for example, if another animal reacts negatively to a eating a food, then a conspecific can learn to avoid this food "because" it is disgusting.

These findings in animals suggest that they possess basic causal inference processes that they share with humans, and whose conclusions are communicated—whether automatically or intentionally—to others. However, while humans have both innate action and emotion recognition systems (see below) as well as general-purpose associative learning algorithms for detecting covariations that they share with other animals (e.g., Shanks & Dickinson, 1988), they also uniquely possess other faculties relevant to causal explanation, such as a "social cognition" module, mental simulation of causality, and conversation. In this chapter, I examine these faculties before moving on to examine more traditional concerns of attribution theory such as the covariational analysis of causality, the dispositional inference process, and the question as to whether this last process is accompanied by a tendency to underestimate the influence of situational causes on behavior (e.g., Gilbert, 1998; Trope & Gaunt, 2003).

FRAMEWORKS FOR STUDYING CAUSAL ATTRIBUTION AND EXPLANATION: KNOWLEDGE STRUCTURES, COVARIATIONAL ANALYSIS, AND CONVERSATIONAL EXPLANATION

Heider's (1958) book *The Psychology of Interpersonal Relations* has influenced social psychological thinking about attribution through two routes—one indirect, and the other a direct one. The indirect route was taken through artificial intelligence and cognitive science and came about through Heider's influence on Schank and Abelson's (1977) knowledge-structure approach to text understanding and question answering. Schank and Abelson's approach extends Heider's conceptual analysis of key concepts (e.g., "can" and "ought") and personal causality in the naive analysis of action, through explicating automatic and unconscious inference processes that enable goal-based explanations of action sequences described in narrative texts (Carbonell, 1981; Wilensky, 1981, 1983). The knowledge-structure approach has

spawned social psychological analyses of attribution processes (Abelson & Lalljee, 1988; Hilton & Slugoski, 1986; Lalljee & Abelson, 1983; Read, 1987), and we shall see that its emphasis on goal-based explanation is consistent with recent work in developmental psychology and neuroscience on "theory of mind" in children (Malle, 2004). The approach also bears important formal similarities with recent work in artificial intelligence and cognitive psychology that has used structural modeling to understand the role of counterfactuals in causal reasoning (Pearl, 2000; Sloman & Lagnado, 2005).

As is widely known, Heider's book also had a direct impact on attribution theory through the interpretations given it by Jones and Davis (1965) and Kelley (1967). In contrast to the knowledge-structure approach, this "lay scientist" perspective on causal attribution focuses on more deliberative causal reasoning processes, such as covariation analysis and discounting. It has several central assumptions, namely, (1) the partition into internal (dispositional) and external (situational) causes; (2) the analysis of covariation information in establishing causality; (3) a hydraulic model of causation that assumes that causes are polar opposites; that is, a cause can be internal or external but not both; and (4) the comparison between lay and scientific models of causal inference. The "lay scientist" perspective led to a concern with the issue of bias and error in lay causal attribution as compared to normative models of judgment, which was quite natural given social psychology's concerns with accuracy in impression formation and the concurrent growth of interest in heuristic and biases in reasoning and judgment (Nisbett & Ross, 1980).

However, both the knowledge-structure and "lay scientist" perspectives share an important limitation. They both see causal attribution as an essentially cognitive process by which an isolated individual forms a mental representation of the world. This is consistent with Heider's early work on the philosophy and psychology of visual perception, but it fails to do justice to the social aspects of ordinary causal explanation. In Heider's use of the term, "attribution" is a cognitive process whereby one can trace a product to its origins, such as when one attributes a reflection in a window to the light from the sun, or a painting to an Old Master. However, as (Hilton, 1990, 1991) notes, *explanation* is a form of social interaction, constituting a three-place predicate whereby *someone* explains *something* to *someone else*. For example, I can attribute the 9/11 attacks to someone, but that is not the same as explaining the attacks to him. Good explanations should therefore obey conversational principles such as relevance to another's informational needs (Grice, 1975; Slugoski, Lalljee, Lamb, & Ginsburg, 1993) and may also obey societal imperatives such as social control (Tetlock, 2002). The knowledge-structure and covariational approaches therefore need to complemented with question-answering systems (cf. Lehnert, 1978). This allows us to see, for example, that people may change their explanations simply in order to take another's person's perspective into account (causal backgrounding), not because they have changed their own mind about the causes of an event (causal discounting).

Three Ways to Explain the Verb Effect on Causal Explanation?: Covariation, Knowledge Structures, and Conversational Principles

The three approaches described above are not incompatible and can deliver complementary insights. For example, they can illustrate diverse aspects of the verb effect on causal explanation (Brown & Fish, 1983; McArthur, 1972), whereby verbs describing emotion (e.g., *Ted admires Paul*) are typically attributed to the stimulus, whereas verbs describing actions (e.g., *Ted cheats Paul*) are more likely to be attributed to the person. In addition, the English language (like many others) has a rich vocabulary of adjectives that describe variability in the disposition of those stimuli to elicit emotions (*likable, loathsome*, etc.) but few derived from state verbs to describe variability in the enduring disposition of people to experience emotions (e.g., *likeful* and *loathive*). In contrast, natural languages have far more adjectives derived from action verbs for describing variability in the disposition of people to produce actions (*charming, deceitful*, etc.), than they have to describe variability in the disposition of people to be acted upon in certain ways (*charmable, deceivable*, etc.).

Thus, the "lay scientist" approach would explain verb effects on causal attribution in terms of causal schemata (i.e., assumed patterns of covariation–Kelley, 1973; see also Rudolph & Försterling, 1997). In general people assume that actions tend to covary with the actor (*few others do it to y*; *x does it to many others*) whereas emotions tend to covary with the object (*many others feel the same way about y, x feels the same way about few others*). Attribution of actions to actors and emotions to stimuli would thus follow from the covariation rule. In particular, Brown and Fish (1983) conjecture that the hardwired nature of emotional recognition systems will make them consensually shared by conspecifics and thus very diagnostic about the nature of the eliciting stimuli, whereas actions that are under the control of higher-order cortical structures will be more diagnostic about agents' personalities.

However, knowledge structures can refine and even inverse these basic attributional patterns. For example, emotional reactions can enable us to make inferences about the person experiencing them when they are "disproportionate" to their object (Smith, 1789/2002). Thus an adult who expresses wild delight at receiving an ice cream might seem "childish," whereas someone who carries on regardless after receiving news of the death of a loved one would seem "unfeeling." World knowledge, here in the form of higher-order knowledge about general patterns of behavior, tells us that such reactions are abnormal and therefore need explaining. Research using facial expressions (Trope, 1986) and emotional responses presented in verbal scenarios (Brown & van Kleeck, 1989) shows that people do indeed make these kinds of dispositional inferences about the person when they display inappropriate emotional reactions. Such attributions serve the important function of creating "cognitive balance" (Heider, 1958) in a way that is very consistent with current connectionist models of how cognitive systems achieve equilibrium (Read, Vanman, & Miller, 1997).

Finally, conversational constraints such as the requirement to be informative to others (Grice, 1975) can influence the explanations that are generated. When asked to explain a normal event such as *The policeman protects Paul*, people tend to focus on something unusual about Paul (e.g., he is a celebrity), as it would be very uninformative to mention the policeman as a cause, even though protection is strongly associated with a policeman's role. Indeed, it is the socially shared nature of common knowledge about policemen's roles that makes this redundant to mention in an explanation to an adult (including to an experimenter). However, as would be predicted by Gricean principles, people will instead refer to the policeman's role when asked to explain the event to a 4-year-old, as young children may be expected not to know what it is that policemen generally do (Hilton & Slugoski, 2001).

In sum, the covariational, knowledge structure, and conversational perspectives may best be thought of as highlighting different aspects of the attribution and explanation process. Through innate expectations or through observing covariations we may develop knowledge structures of varying levels of abstraction, such as *people are sad on hearing bad news* and *policemen protect people*. But if we observe something that does not fit our covariation-based expectations (e.g., someone does not look sad upon hearing bad news), then attribution processes are likely to be triggered. However, if someone else asks us to explain an event, I focus on something that is informative from their point of view, not from ours.

PLAN OF THIS CHAPTER

In the first part of this chapter I begin by tracing the origins of the attribution process in the automatic causal perception of motion. We will see that social perception appears to be "Kantian" in that action and emotion perception are based in large part on innate and hardwired specialized modules that are directly stimulated by certain kinds of cues. These innate social perception modules form the basis of a "naive psychology" that yields knowledge structures that enable explanation of sequences of human behavior.

In the second part I review evidence that these knowledge structures support causal explanation through counterfactual reasoning using mental simulation. These simulations may yield complex causal chains that produce an event to be explained, whereas in ordinary explanation we usually talk about only one or at most two factors as "the" cause. In the third part I show that people skillfully and intuitively use conversational intuitions to select "the" factor from these complex cognitive representations that is most relevant to mention in a conversation about causes. These conversational inference and explanation processes mostly seem to occur automatically and outside conscious awareness in adults.

In the fourth part I show how the aforementioned knowledge structure and conversational perspectives can illuminate issues and resolve problems in the analysis of variance (ANOVA) framework of the classic lay scientist approach (Kelley, 1967). In particular, attention to the role of prior knowledge in the inference process, the different kinds of causal inference (e.g., causal explanation of events vs. dispositional attribution to entities), and the conversational pragmatics of experimental conversations about causes suggests that people may be rather more rational and less biased in their causal attributions than early studies have suggested (McArthur, 1972). In the fifth part, I extend the knowledge-based approach to the analysis of discounting processes (Jones & Davis, 1965) and again show that evidence for bias is less strong and less general than might be supposed from well-known studies in this paradigm (e.g., Gilbert, Pelham, & Krull, 1988; Jones & Harris, 1967).

THE DEVELOPMENT OF FOLK PSYCHOLOGY: ACQUISITION OF KNOWLEDGE STRUCTURES FOR CAUSAL EXPLANATION

In this section I first examine evidence from neuroscience and developmental psychology to establish what humans inherit in their "social cognition" module and then consider how, in combination with learned knowledge acquired through experience, this can mature into a folk theory of psychology that is used to interpret and explain human behavior. In particular, neuroscience research has painted a picture of the human brain as containing a highly specialized set of modules for various aspects of perception, memory, and inference, which differs from the British empiricist approach of Hume and Mill to causal inference that has influenced social psychological research, with its focus on general-purpose covariation analysis.

Is There an Innate Social Cognition Module?

Social Perception through Observation of Movement

People have a natural disposition to see self-propelled movement as expressing intentional agency. Heider and Simmel (1944) demonstrated this disposition by creating a cartoon in which colored geometric figures (like a green square, a red triangle, and a blue square) moved around a black frame in a quasi-random fashion, before the black frame disintegrated at the end of the sequence. A remarkable 49 of 50 observers described the dots' movements in the language of description for human behavior: The dots were described as *chasing*, *following*, *avoiding* each other, and so on. Oatley and Yuill (1984) showed a version of the cartoon to participants with titles that expressed higher-order themes such as "jealous lover." For example, the observers had no difficulty in attributing emotions such as anger to the figures that "explained" why they then went on to destroy the "house" (the frame that disintegrated).

Neuroscience research suggests that perception of motion is integrated with the attribution of intention and causality in the human brain, in particular in brain circuits that recruit the superior temporal sulcus (STS),

which has been found to be "sensitive to biological motion, but more generally to stimuli that signify intentions or intentional activity" (Gallagher et al., 2000, p. 19; see also Allison, Puce, & McCarthy, 2000). For example, Castelli, Happe, Frith, and Frith (2000) used cartoons inspired by Heider and Simmel to induce perception of dot motion either in terms of intentional behavior or as random movements They found that observers perceiving the dot movements as intentional selectively activated the temporal parietal junction (which is very close to the STS), the basal temporal gyrus, the extrastriate cortex, and the medial prefrontal cortex, whereas those perceiving the dots as moving randomly selectively activated the occipital cortex. Grossman and Blake (2002) found that the posterior STS was selectively activated by 12 moving-point displays that give the impression of a human walking, but not by scrambled dot movements. On the other hand, other areas involved in the recognition of objects and bodies, such as the extrastriate body area (EBA) and the lateral occipital cortex (LOC) did not show differential activity in response to scrambled and nonscrambled "walking dot" displays.

Perception versus Knowledge of Causality

Impressions of causality based on visual perception of motion seem to be impervious to top-down knowledge, as when we still "see" dots that collide with each other screen as causing each other to move even when we "know" that their coordinated movement is simply the result of careful computer programming. They are thus a prime candidate for "System 1" thinking based on automatic pattern recognition rather than more deliberate, inferential "System 2" thinking (Sloman, 2002).

Blakemore and colleagues (2001) used patterns inspired by Michotte (1963) that either gave an impression of causality, as in the "launching effect" when a blue ball rolls horizontally across a screen and collides with a stationary red ball, which then moves off screen, compared to noncausal sequences where the red ball rolls across the screen but moves round the stationary red ball by passing under it, or visual transient events where the blue ball rolls across a screen and then changes color to red. Bilateral activation of the STS and activity in the left intraparietal sulcus was greater for causal than for noncausal sequences, regardless of whether participants had been told to make a judgment about the direction of motion of the ball or a judgment about whether there was a causal relationship.

The attribution of intention to self-propelled objects seems to be hardwired in human beings and universal across cultures. Thus, Morris, Nisbett, and Peng (1995) report no difference between Americans and Chinese in perceiving internal force, external force, and animacy of moving dots in Michotte-style entraining and launching movements. Cultural differences did, however, emerge when they presented similar displays using fish (either alone or in groups). Americans tended to perceive that the fish moved more under internal influence whereas the Chinese were more likely to perceive that movement was due to external influence. Although *perceptions* of

causality in Michotte-style launching and entraining movements are not different between Americans and Chinese, some differences emerge in verbal explanations they give (Peng & Knowles, 2004): Americans were more likely to attribute the dot's movements to "dispositional explanations" (e.g., weight and mass, composition, inertia, and shape), whereas Chinese were more likely to produce "contextual explanations" (e.g., the other object, gravity, friction, and wind). These findings suggest that while visual perception of causality is hardwired, universal, and "direct," verbal explanations of these movements may appeal to culturally specific "naive theories" of physics.

The Perception of Equifinality in Infants and Children

The principle of equifinality, which Heider (1958) characterized as "the invariance of the end and variability of the means" (p. 101), is a core assumption of his analysis of commonsense explanation of human action. In Heider's theory of personal causality, intentions make good causal explanations because they economically focus on the factor that makes the most difference to an outcome: "Attribution to personal causality reduces the necessary conditions essentially to one, the person with intention, who, within a wide range of environmental vicissitudes, has control over the multitude of forces required to create the specific effect" (Heider, 1958, p. 102).

Children appear to be endowed with an "intentional stance" to interpret behavior from an early age. For example, even 6-month-old infants appear have developed expectations about goal-directed action. Using a habituation procedure, Woodward (1998) found that infants of this age expect animate objects (specifically, human hands) to do things such as reach for objects they were just reaching for previously but do not expect inanimate objects that resemble human hands (e.g., a garden tool "claw") to "reach" toward the familiar object in similar circumstances. In a related vein, Tomasello, Carpenter, Call, Behne, and Moll (2005) review evidence showing that children encode movements in terms of intentional actions; for example, they imitate what an adult is *trying* to achieve by a movement and do not imitate the adult's actual physical movements (e.g., when impeded) in a literal-minded way.

In a striking demonstration of infants' understanding of equifinality, Gergely, Nádasdy, Csibra, and Bíró (1995) habituated infants to a small self-propelled dot "jumping" over an obstacle when approaching a large dot. Later, with the obstacle removed, 12-month-old infants showed surprise when the same jumping motion was repeated. Even at this age, children seem to show an innate sense of counterfactual trajectories (as if the child asks herself "With the obstacle gone, why would the small dot still "jump" and not go directly to the large dot as it could (and should) have done?"). Even these very young children therefore appear to understand "equifinality," as they did not show surprise at the small dot now going directly to the large dot (even though this was a new motion).

The Development of a Theory of Mind

Premack and Premack (1995) argue that the development of a "theory of mind" from infancy to early childhood will go through three stages: (1) automatic activation by movement of the perception that an agent is intentional; (2) interpretation of interactions of moving objects as "helping" or "hindering" each other; and (3) at about the age of 4 years, the development of a "theory of mind" similar to that of adults with the predicates, *see, desire and believe*. Social interaction and involvement in conversation may be important here: Harris (1996) suggests that after learning language and engaging in conversation, children are able to learn that others have beliefs different from theirs and thus move from a "desire" psychology (using words such as "want") at the age of 3 years to a "belief–desire" psychology at the age of 4 years (using words such as "think" as well as "want"). This helps them succeed on "false belief" tasks that require a child to take into account another child's incorrect beliefs (Wimmer & Perner, 1983). Consistent with this view, Carpendale and Lewis (2004) review evidence that acquisition of a "belief psychology" is associated with factors associated with quality and quantity of social interactions. For example, children tend to be later in success on the false-belief task if they have fewer siblings, experience less family talk about mental states, and suffer from deafness or attachment anxiety. Lack of language and communication may also help explain why deaf children with hearing parents who cannot communicate with them in sign language, like children younger than 4 years of age, have difficulty on "false-belief" tasks (Woolfe, Want, & Siegal, 2002; see also Tomasello, Call, & Hare, 2003, for related evidence on chimpanzees).

Mentalizing: Development of Higher-Order Attributional Inference

Although direct perception of agency is available through the visual system, causal judgments of the kind described by attribution theories in social psychology seem to require higher-order inferential processes (cf. Lieberman, Gaunt, Gilbert, & Trope, 2002). Evidence that attribution of beliefs is a higher-order process may be surmised from work suggesting that genetic disorders cause problems in the development of a mature "theory of mind." Thus autistic children appear to have difficulties with tasks involving attribution of mental states that require inference rules similar to the noncommon effects principle of Jones and Davis (1965). Although autistic children are able to perceive the direction of gaze, unlike normal or mentally retarded children they are unable to infer that if a cartoon figure is looking at one kind of chocolate bar rather than another, it is because he or she prefers that bar (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995).

Autists are thought to have problems in "mentalizing," by which Frith and Frith (2003) mean the activity necessary for understanding narratives (as opposed to incoherent sequences of text) that require attributions of beliefs and intentions to story characters (Schank &

Abelson, 1977). Mentalizing recruits the temporal poles and the medial prefrontal cortex (MFPC) as well as the STS involved in visual perception of motion and causality. The fact that mentalizing involves centers such as the MPFC, which has direct links to the STS and the temporal poles, suggests that adult attributional inference in understanding stories integrates brain circuits involved in the visual perception of causality with the MPFC, that is also recruited by higher-order reasoning processes about social rules (Adolphs, Bechara, Tranel, & Damasio, 1995).

Personal Causality in Commonsense Explanation: A Revision of Heider

What does the matured "theory of mind" look like in adults? We may surmise that it has many of the characteristics given by Heider (1958) in his analysis of personal causality. However, Malle (1999) suggests that Heider's formulation suffered from some incoherences that led his theory of personal causality to be conflated with the person–situation distinction that came to dominate later attribution research. Malle proposed a modified version of Heider's "naive psychology" that incorporates the notion of belief, desire, intention, skill, and awareness, part of which is reproduced in Figure 10.1. First, Malle makes a distinction between intentional and unintentional actions. While unintentional actions can be partitioned reliably into internal and external causes, intentional ones cannot. For example, explanations of accomplishments (McArthur, 1972) that refer to internal factors such as "Jennifer got into Oxford because she is intelligent" rather than external ones such as "Jennifer got into Oxford because the exams were easy" seem to say meaningfully different things about Jennifer and about Oxford. However, phrases that on the surface mark person versus situation factors in explaining intentional actions such as "Jane bought the house because she wanted privacy" versus "Jane bought the house because it was secluded" (Ross, 1977) seem to appeal to the same underlying general motive–action–outcome schema that motivates intention-based systems for understanding human action (Abelson, 1975). They both seem to say or imply the same thing about Jane (she likes privacy) and the house (it is secluded), and in this respect they have much the same deep meaning.

Malle (1999) argues that intentional explanations appeal to the subjectivity rule: that is folk explainers cite "only those mental states as reasons that (as far as they know) the agent considered and in light of which he formed his decision to act" (p. 36). Sometimes the reason will be explicitly marked: as in explaining why Anne watered her plants "because she *wants* them to grow faster" or unmarked as in "so they'll grow faster." This factor alone (see Malle, 1999, for others) renders it quite difficult to code intentional actions for internality or externality. Using probes such as "Imagine you invited someone to lunch. Why would you?" Malle (1999, expt. 2) generated 489 reason explanations than were then coded using the separate person–situation categories of traditional attribution theory and the scheme of folk ex-
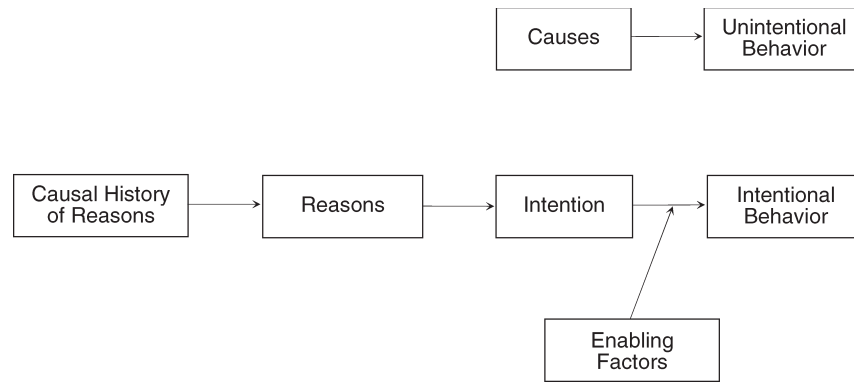
FIGURE 10.1. Malle's depiction of folk theories of behavior.

planation: mental state markers (present or absent); reason content (person or situation); and reason type (belief or desire). The major findings were that 74% of all person attributions referred to reasons with mental state markers (e.g., "I asked somebody out for dinner because I *wanted* to get to know this person better" and "I applauded the musicians because I *enjoyed* their performance") whereas 90% of situation attributions were assigned to unmarked reasons with situation content (e.g., "I watered the plants because *they were dry*" and "I applauded the musicians because *they were good*"). In other words, although person attributions tended to mark mental states and situation attributions did not, they both can express similar underlying ideas about why the action was undertaken (e.g., I *enjoyed* the performance and the musicians *were good*"). As we shall see in the section on conversational pragmatics, the unmarked elements may in fact be presupposed as being true but backgrounded parts of the larger causal explanation (causal backgrounding), which is not the same thing as no longer believing them to have any causal role at all (causal discounting).

Other Sources of "Naive Theories" about Cause–Effect Relationships

Although, the social cognition module equips humans with the kinds of knowledge structures ("theory of mind") needed to understand human action, two other sources of general knowledge about others' behavior that are important for developing a "naive psychology" should briefly be noted.

Emotion Perception

Humans appear to be equipped with brain structures that seem to be tuned to the perception of particular emotional reactions (Keltner, Ekman, Gonzaga, & Beer, 2003); for example, the perception of sad faces activates the left amygdala and the right temporal lobe, whereas perception of angry faces activates the right orbitofrontal cortex and cingulate cortex. Bilateral lesions to the amygdala impair the ability to recognize fearful facial expressions and vocalizations but not the ability to recog-

nize facial expressions of sadness, disgust, or happiness. Recognition of many emotional expressions thus appears to be accomplished without reference to higher-order cortical structures. As was noted in the discussion of the verb effect on causal explanation, emotional expressions will generally be attributed to the presence of environmental stimuli, unless contextual knowledge suggests that they are inappropriate over- or underreactions.

Associative Learning of Causal Relationships

Although I do not have the space to go into detail, it is important to note that humans can *learn* causal relations from experience. They often appear to do so by using simple associative heuristics that are shared by other species (Rescorla & Wagner, 1972; Shanks & Dickinson, 1988; for applications to social attribution tasks, see van Overwalle, 2003; ; van Overwalle & Rooy, 2001a, 2001b; van Overwalle & Timmermans, 2005). These general-purpose "Humean" associative learning heuristics use cues to causality such as spatial and temporal contiguity to build associations that are strengthened by repeated co-occurrence of events (Hume, 1888). They take as inputs multiple observations and output universal causal generalizations (Hilton, 2002).

EXPLAINING THE COURSE OF EVENTS: KNOWLEDGE-BASED COUNTERFACTUAL SIMULATIONS

In the previous section, we have seen how innate "Kantian" structures for interpreting perceptions of actions and emotions, can combine with acquired generalizations to furnish the knowledge structures that can be used to interpret human behavior in terms of "stories"— sequences of emotional reactions that generate intentions that generate plans that in turn generate actions that bring about new emotional reactions, and so on (Schank & Abelson, 1977). In this section, we consider how such *general* knowledge structures drive the counterfactual simulations that can be used to explain *particular* instances of human behavior.

Causal Induction versus Causal Ascription: Acquiring versus Using Knowledge

One problem in comparing the layperson to a scientist is that science is typically concerned with causal induction, whereas lay causal inquiries are typically concerned with causal ascription. Whereas science characteristically answers abstract questions about *types* of event (Why do apples fall? Why do species evolve? How do minds retrieve memories? Why do people show ingroup favoritism?), commonsense causal inquiries are context-specific and typically concerned with explaining why a *particular* event occurred when and how it did. This point is well put by legal philosophers Hart and Honoré (1985), who wrote:

> The lawyer and the historian are both primarily concerned to make causal statements about *particulars*, to establish on some particular occasion some particular occurrence was the effect or consequence of some other particular occurrence. The causal statements characteristic of these disciplines are of the form 'This man's death was caused by this blow.' Their characteristic concern with causation is not to discover connexions between types of events, and so not to *formulate* laws or generalizations, but is often to *apply* generalizations, which are already known or accepted as true and even platitudinous, to particular concrete cases. (pp. 9–10)

As an example, Hart and Honoré consider the case of a man who has died as a result of being shot. The causal generalization that *lack of oxygen in the brain causes death* would not explain why this man dies when and how he did, being true of all deaths. Rather, the pragmatic interests of the lawyer, doctor, and layman would be to identify the *abnormal condition* that distinguishes this man's state, being dead, from the normal state (being alive). The explanation that he died because he was shot identifies a factor that made a difference to the course of events and gives new information that is relevant to the question about the cause of this particular death. On the other hand, the answer that he died because of lack of oxygen cannot explain this particular death, even though it is more likely to be true, and to have a higher covariation with death than being shot.

In ordinary commonsense explanation, we use knowledge of *general* relationships to ascribe causes for *particular* events. Whereas science seeks to *gain* new general knowledge through causal induction (e.g., by testing the causal efficacy of a new drug in controlled experimental trials), commonsense causal explanation *uses* prior knowledge to generate mental simulations that explain particular events that attract attention, for example, if they are surprising (e.g., a known drug such as aspirin fails to have its expected effect). While commonly unspoken, psychologists ignore the role of these axioms in lay causal explanation at their peril. Next, I examine how such knowledge structures are used to support counterfactual simulations in reasoning about causality.

Counterfactual Simulation in Explaining the Course of Events

Recent work in artificial intelligence has suggested that intervention through manipulation—doing things to the world—provides a fundamentally different way of learning about causality to observation of covariation (Pearl, 2000; see also Halpern & Pearl, 2005a, 2005b). Manipulation enables interventions to be conducted on causal chains to see whether an effect still occurs. It also effectively neutralizes the effect of earlier events in the chain that are cut out of the picture (Pearl uses the metaphor of "surgery" to describe this operation) and can thus establish whether the manipulated factor influences the outcome, independent of prior events in the chain. Mental manipulation can also be carried out in counterfactual *simulations* (cf. Kahneman & Miller, 1986), where reasoning based on general knowledge is used to calculate "trajectories" about what should happen if a counterfactual intervention (which did not actually happen) is imagined. For example, in assessing the causes of the crash of the Concorde in 2000, I may mentally "undo" the occurrence of a catastrophic fire in the engines and make a projection that the accident would not have occurred if there had been no fire. I would thus conclude that the fire was on the causal trajectory that led to the accident, whereas if I mentally mutate the speed at Concorde took off (imagine it was traveling faster when it took off) or the kinds of passengers on board (imagine that they were all French rather than German), my mental simulation would find that these factors would make no difference to the outcome. I would therefore conclude that the Concorde's speed at takeoff or the kinds of passengers it was carrying were not on the causal trajectory that led to the accident.

Experimental work has suggested that people do use counterfactual reasoning of this kind to make these kinds of causal judgments (Sloman & Lagnado, 2005) This kind of counterfactual reasoning needs general "covering laws" (Hempel, 1965) to drive inferences about what should happen if a particular event in a causal chain is changed. In interpreting the physical world these may take the form *catastrophic engine fires cause accidents*, or *objects impacting at others at high speed cause damage*, or in political behavior *the strong do what they will, and the weak do what they must* (cf. Tetlock & Lebow, 2001). These inference rules constitute universal causal generalizations that express relationships between *types* of entity that hold across space and time rather than between *particular* entities that are located in space and time (Hilton, 2002; see also Oestermeier & Hesse, 2001). They are modeled as structural equations linking *variables* in Halpern and Pearl's model and are used to calculate the consequences when particular states of the world (called *values*) are altered in a counterfactual simulation (e.g., one imagines what would have happened if Concorde's fuel tank had *not* been impacted by debris at high speed during takeoff).

Some general inference rules about intentional behavior are probably innate, such as the equifinality rule, *If people have a goal, then they will circumvent obstacles to that goal*, whereas other rules will be learned through experience and acculturation, such as scripts like *People always buy something on their visit to supermarkets* or *People sometimes trip up over other people dancing* (Hilton & Slugoski, 1986; Schank & Abelson, 1977), or *If at a party, people tend*

to look happy (Chun, Spiegel, & Kruglanski, 2002; see also Kruglanski, 1988). These general beliefs about type relationships can also be expressed as conditional probabilities (e.g., *If instructed to write a counterattitudinal essay, then people usually produce an unconvincing text*) (cf. Morris & Larrick, 1995). While people in different cultures may acquire different cultural generalizations (Kashima, Siegal, Tanaka, & Kashima, 1992; Morris, Ames, & Knowles, 2001), the fact that these knowledge structures are socially shared means that they will hardly ever be *given* as explanations but will, rather, be presupposed as shared background knowledge in conversations about causes between members of the same group.

General beliefs about what "usually" or "often" happens can the influence the generation of "if only" counterfactuals and responsibility attributions. For example, Catellani, Alberici, and Milesi (2004) showed that participants who held stereotyped expectancies about rapes (as measured by the endorsement of rape myths) generated more "if only" counterfactuals that focused on the victim's rather than the perpetrator's behaviors, and in particular on things that the victim could have done but did not do. In turn, generation of these counterfactuals about what she could have done (but did not) was associated with higher attribution of responsibility for the rape to the victim.

Goal-Based Explanation and the Generation of Counterfactuals

Below I show how Heider's principle of equifinality can generate the inferences necessary to buttress counterfactual reasoning about human action. To illustrate, let us return to the example of the Concorde crash. The Concorde crashed shortly after it hit some debris that had just fallen off a Continental Airlines aircraft that had just taken off from Roissy Airport at Paris. The debris hit the underside of the Concorde's fuel tank and pierced it. This caused fuel to leak out, which caught fire in the atmosphere, leading to a catastrophic fire in the engines, resulting in the crash. Described in this way, many people considered the presence of the debris on the runway to be the cause, as its presence on the runway was abnormal, and counterfactual reasoning tells us that if the Concorde had not hit it, the crash would not have occurred. Although the debris falling off the earlier aircraft is abnormal, people do not intuitively consider this to be a cause, reflecting the principle that causality is not traced through the proximal abnormal condition to the distal abnormal condition (Hart & Honoré, 1985; see Hilton, McClure, & Slugoski, 2005, for relevant empirical results).

Now imagine that the debris had been deliberately placed on the runway by a saboteur hoping to bring about the accident. Here, the presence of the debris on the runway would now seem to be a "means" to the saboteur's "end." In this case, Hart and Honoré (1985) would predict, as Hilton and colleagues (2005) have shown, that causality will be traced through the proximal abnormal condition (e.g., the presence of debris on the runway) to the distal one (e.g., the saboteur's action). This

may be because equifinality-driven counterfactual reasoning now tells us that that even if the debris had not been on the runway, the Concorde would still have crashed (somehow), as the saboteur would have found another way to bring the crash about. In this sense, the debris on the runway is no longer a necessary condition for the Concorde crash, although it is still a necessary part of the causal trajectory that caused the Concorde to crash when and how it did. Note that in contrast the physical distal cause of the accident (the faulty maintenance of the Continental jet that caused the debris to fall off) does not plausibly support a counterfactual of the kind "Even if the Concorde hadn't hit the debris, the faulty maintenance of the Continental jet would still have found a way to make the Concorde crash on another occasion").

Models based both on Schank and Abelson's (1977) framework and Mackie's (1980) theory of counterfactual reasoning have been convergent and successful in predicting what parts of a story people find important and remember. Both give highly convergent predictions, suggesting that they draw on related intuitions, although Mackie's analysis performed slightly better (Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985). A seeming anomaly for these models is that they assume that people create complex underlying representations of the conditions necessary for the occurrence of an event, yet they typically only mention only one or two factors as "the" cause of an event in ordinary causal explanation. A theory of causal selection is therefore needed to understand how people select causes from conditions in conversation, and I turn to this question below.

CONVERSATIONAL PROCESSES IN EXPLANATION

Audience Effects on the Selection of Causes from Conditions

If counterfactual reasoning reveals that there is a plethora of necessary conditions for an event, the problem of causal selection arises, as normally we only mention one or two factors in conversation as "the" cause (Hesslow, 1983, 1988). In particular, we should be sensitive to the questioner's perspective when giving explanations through giving them explanations that are not only true but informative, relevant, and clear to the inquirer (cf. Grice, 1975; Hilton, 1990).

Slugoski and colleagues (1993) demonstrated that people follow these conversational principles in interpersonal explanation. They had participants read a detailed case history about a youth who had committed a crime, which included personality information about the youth, and situational information about the circumstances in which the crime occurred. Participants were then asked to explain the crime either to an interlocutor who knew a lot about the boy's personal history but not the circumstances of the crime or to another interlocutor who knew little about the boy's personal history but a lot about the circumstances of the crime. Slugoski and colleagues found that explainers varied their explanations so as to

complement the questioner's prior knowledge: that is, interlocutors who already disposed of information about the boy's personal history received explanations that emphasized situational factors, and vice versa for the interlocutors who already knew of the situational circumstances surrounding the crime. Slugoski and colleagues' results are consistent with the idea that the explainers are identifying the condition that is abnormal from the interlocutor's point of view (Hilton & Slugoski, 1986). Similar reasoning has been applied by McGill (1989) to explain how changes in presupposed contrast cases can explain actor–observer and success–failure asymmetries in explanation.

Conversational constraints may also govern the formulation of fuller conjunctive explanations for audiences that have little relevant background information. Kashima, McKintyre, and Clifford (1998) found that Australian participants who had to imagine that they had to explain mundane behavior (e.g., going to the Melbourne cricket ground to watch a football match) to other Australians referred primarily to desires (e.g., wanting to watch the match) in their explanations. However, when asked to imagine explaining the event to a tourist, they referred to both desires and beliefs (e.g., "Michael likes football and knows that teams play football at the MCG"), presumably because participants recognized that strangers could not be expected to share the relevant beliefs. Kashima and colleagues obtained this result only with open-ended explanations, not with rating scales, suggesting that only verbally given explanations are sensitive to Gricean constraints (cf. Cheng & Novick, 1991; McGill, 1991). In a similar vein, Malle (in press-a, in press-b) reports a meta-analysis of actor–observer asymmetries in explanation that finds that observers only report more person causes than actors when open-ended explanations are coded, not when rating scales are analyzed. Malle suggests that this may reflect greater use of mental state markers (e.g., referring to beliefs and desires) by observers compared to actors in free explanations, rather than any underlying differences in patterns of attribution.

Abnormality in Commonsense Explanation and Attribution of Responsibility

The general inference rules encapsulated in the knowledge structures used in goal-based explanation will drive inferences about what normally should happen in a given situation. For example, we will probably assume from general world knowledge that food is inexpensive and easily available for an American or European who has stopped in to eat at a restaurant. However, if we learn that an Ethiopian refugee has just enjoyed a meal, we will change our assumptions about what is normal in the situation. Here we would attribute this behavior to the condition that is in fact abnormal in these circumstances, namely, the availability of food (Hart & Honoré, 1985; Hilton & Slugoski, 1986). Consistent with conversational principles of causal explanation, which require an explanation to be informative and relevant (Grice, 1975; Hilton, 1990), McClure and Hilton (1997) indeed found

that goals were perceived as more informative and relevant explanations for routine actions, but preconditions were judged as more informative and relevant explanations for actions when their presence was exceptional (e.g., food is available in Ethiopia).

Types of Causal Question and the Relevance of Explanations

People also have clear intuitions about what kind of causal question is appropriate to a given situation. Experiments on question answering about human behavior in narratives (Graesser, Robertson, & Anderson, 1981) show that "why" questions are typically answered by going up a goal hierarchy from an action (e.g., "Why did the prince get a horse?") to a goal ("Because he wanted to get to the castle"), whereas "how" questions typically involved descending from a goal ("How did he get to the castle") to an action that forms part of a plan to attain the goal ("By getting a horse"). Accordingly, McClure, Hilton, Cowan, Ishida, and Wilson (2001; see also Malle, 1999) found that actions that are normally difficult for the actor to perform, such as passing a difficult exam or a poor man taking a trip round the world, provoke "how come" questions that are most frequently answered by explanations that refer to preconditions (e.g., He's a stats whiz" or "He won the lottery"). On the other hand, "why" questions about routine actions tended to elicit explanations that focused on goals, while questions about "what caused" routine actions focused about equally on goals and preconditions (McClure & Hilton, 1998). Each kind of question set up different expectations about relevance. Thus, while goals were perceived as most informative and relevant to "why" questions, both preconditions and goals were perceived as equally informative and relevant responses to "what caused" questions.

Causal Backgrounding versus Causal Discounting

We have seen previously that factors such as the perceived knowledge and interests of the audience and the kind of causal question posed will change what people focus on and background in a conversationally given explanation. This kind of *backgrounding* of explanations that are still believed to be true but become less relevant to mention in a conversation about causes is not the same thing as *discounting* a causal explanation because one no longer believes it.

In *causal backgrounding*, we drop an explanation because it is no longer informative or relevant to mention in a conversation about causes, not because we no longer believe it to be part of the larger causal story responsible for the occurrence of the target event. For example, context may change what people call a cause by changing what is perceived as normal in the circumstances. Suppose we learn that a hammer strikes a watch, and the watch face breaks. Most people here assume that the watch broke because the hammer hit it. Now suppose that we now learn that this happens as part of a routine testing procedure in a watch factory. In this case, people tend to prefer the explanation that the watch broke be-

cause of a fault in the glass (Einhorn & Hogarth, 1986), as this is the abnormal condition that makes the difference between this watch that breaks and other watches that do not break in the circumstances under consideration. Nevertheless, people still believe that the hammer hit the watch and that this was necessary for the watch to break but now consider the hammer strike to be a less informative and relevant explanation than the fault in the glass (Hilton & Erb, 1996).

In contrast, in *causal discounting*, we drop an explanation because we no longer believe it to be true. For example, suppose a particular individual has contracted cancer. His lawyer may argue that working in an asbestos factory is the cause, thus trying to get his employers to pay compensation. In order to falsify this argument, his employers' lawyers may counter that this worker has smoked for many years, and that this is likely to be the cause of his cancer (Einhorn & Hogarth, 1986). Experimental research does in fact confirm that the activation of the rival asbestos pollution hypothesis does indeed lead people to consider the initial smoking hypothesis to be less probably true (Hilton & Erb, 1996).

Distinguishing causal backgrounding from causal discounting enables us to determine when rival explanations are complementary ways of looking at the same thing or are genuinely competitive (if one is true, the other is probably false). For example, a doctor may explain a victim's lung cancer to his wife by pointing out that he smoked heavily (thus differentiating him from his twin brother who did not smoke and who did not get cancer). But she might explain the same patient's cancer to another doctor by referring to the genetic disposition of members of his family to get cancer, thus explaining why this man got cancer when other heavy smokers did not. But *tuning* her explanation in this way to make it informative to her audience is quite different to *changing* her explanation because she instead now believes an alternative hypothesis (e.g., the man got cancer because he worked in an asbestos factory).

In summary, alternative causal scenarios are genuine rivals in causal discounting (as they are in medical diagnosis) because each particular outcome normally has only one causal history, thus justifying the "hydraulic" assumption made by Morris and Larrick (1995) that events to be explained normally only have one "cause." As we shall see later, knowing when it is appropriate to apply this hydraulic assumption as a normative standard helps determine whether people over- or underdiscount causes.

LOGIC AND KNOWLEDGE IN THE ANOVA FRAMEWORK

As we have seen previously, the knowledge-structure and conversational approaches are quite naturally able to deal with the question of how explanations are selected from sequences of conditionally dependent events, where the occurrence of later events in the causal chain is dependent on the occurrence of earlier ones. But as Kelley (1983) himself recognized, his ANOVA frame-work, with its assumption of conditional independence due to the orthogonal contrast of variables (person, stimulus, occasion), cannot apply to causal chains of events. To overcome this problem, Spellman (1997) proposed a covariational model based on successive reconditionalization (but see Mandel, 2003, for a critique).

In this section, I show that even the relatively simple case of conditional independence presupposed in Kelley's (1967) ANOVA framework came to be loosely formulated, thus sowing confusion about what is and is not a normative causal judgment, giving rise to misconceptions that persist to this day in both theoretical formulations and experimental and questionnaire methodologies. I then show that the knowledge-based, pragmatic approach presented earlier, together with refinements in models of the causal inference process, can illuminate and resolve these problems. I then extend this knowledge-based approach to the discounting principle in the next section.

Implicit Knowledge in the ANOVA Framework

Kelley (1967), following Mill's method of difference, defined a cause as "that condition which is present when the effect is present and which is absent when the effect is absent" (p. 154). Given an event such as *Virginie liked "Helène et les garçons,"* this logic implies that certain "control conditions" need to be examined to test the effect of certain causes. In Kelley's view, the layperson might identify three possible causes: the person (e.g., something about the TV watcher); the stimulus (e.g., something about the TV serial); or the occasion (e.g., something about the day on which the TV serial was seen). To test these three causal hypotheses, the lay scientist should construct and examine three experimental control conditions. The lay scientist should therefore use Mill's method of difference (Jaspars, Hewstone, & Fincham, 1983) to determine which factor made the difference between the occurrence and nonoccurrence of the event and was thus a cause of the target event.

For example, if one wanted to test the hypotheses that something about this particular TV watcher was the cause of her enjoying the program, one would obtain *consensus* information by comparing her reaction to that of other people watching the program at the same time. If it turned out that Virginie liked the program but no one else who watched it did (low consensus), then we might infer that something about Virginie was the cause (or part of the cause) of her enjoyment, as the effect (enjoyment) occurred when Virginie, but not others, was watching. Virginie might be a TV addict, for example. However, if everybody who watched the program enjoyed it (high consensus), then the effect (enjoyment) occurred even when Virginie was not watching and therefore something about her cannot be the cause.

A similar analysis can be applied to determining whether something about the TV serial caused Virginie to enjoy it or not. If it turns out that Virginie liked *Helène et les garçons* but not other programs she has watched (high distinctiveness), then something about this particular TV serial must be the cause (or part of the cause) of

the effect. One might infer that it is a very entertaining TV serial, for example. However, if Virginie enjoys every TV serial that she watches (low distinctiveness), then nothing about *Helène et les garçons* can be the cause of her enjoyment, as she enjoys TV serials regardless of whether the television is showing this particular one or not.

Finally, a similar logic applies to the analysis of whether the occasion plays a role in Virginie's enjoyment of the TV serial. If Virginie enjoys *Helène et les garçons* when she watches it on her birthday but at no other time (low consistency), then something about the occasion may be a cause (or part of the cause) of her enjoying the program. If, on the other hand, she enjoys *Helène et les garçons* every time she watches it, then her enjoyment cannot be explained by the particular occasion, as it occurs both on this particular occasion and on others.

Thus, through use of Mill's (1872/1973) method of difference, people will attribute causality to the factor whose presence makes the difference to the event occurring or not. According to this logic, a configuration of low consensus, low distinctiveness, and high consistency (LLH) should lead to an attribution to something about the person, such as *Virginie is a TV addict*, whereas a configuration of high consensus, high distinctiveness, and high consistency (HHH) such as the one below should lead to stimulus attributions, such as *Helène et les garçons is an entertaining program.* McArthur (1972, 1976) presented verbal vignettes to subjects and found that covariation information of this type did indeed have a considerable effect on subjects' attributions as predicted. On the basis of these results, Orvis, Cunningham, and Kelley (1975) proposed a template-matching model of attributional inference. However, shortcomings in McArthur's experimental methodology and Orvis and colleagues' model led to confusion about what the Kelley ANOVA framework *should* predict for all eight information configurations.

Problems with the Early Formulations of the Kelley ANOVA Model

In retrospect, Kelley's (1967) framework, and the early experiments that tested it (McArthur, 1972, 1976; Orvis et al., 1975) are as notable for what was left out as what was included. First, a fully crossed person × stimulus × occasion (2×2×2) design comprising eight experimental cells is needed to test three hypotheses about why a target behavior occurred (is it due to the person; the stimulus; or the occasion?), Yet Kelley only described information in four cells: the target behavior; consensus information (high or low); distinctiveness information (high or low); and consistency information (high or low). This corresponds to a fractionated cell design (Jaspars et al., 1983; see Figure 10.2).

A second and related omission concerns the response methodology used by McArthur (1972, 1976) and Orvis and colleagues (1975). In addition to the three main-effect causes that are possible (person, stimulus, or occasion) each event should in principle be explicable in terms of four interactions of causes, such as person × stimulus; person × occasion; stimulus and occasion; and person × stimulus × occasion (Jaspars et al., 1983). However, these early studies only presented one interactional response option explicitly to participants, namely, that of attribution to some combination of the person, stimulus, and circumstances, and asked participants to specify which of those factors interacted in a space that was provided for them. This was unfortunate, as the third "main effect" attribution response option was to "something about the particular circumstances," which could also mean a "freak combination of factors" as well as "something about the particular time/occasion." The response format clearly mattered, for when the four interactional attribution options were explicitly specified, the proportion of participants' interactional attributions went up from 35% in McArthur's (1972) experiment to 61% (Jaspars, 1983) and 47% (Hilton & Jaspars, 1987). A final problem was that no "null attribution" response format was presented, indicating that there was no effect to be explained, as would be the case if the target behavior did not deviate from expectations (Hilton & Slugoski, 1986).

World Knowledge and the Missing Dimensions of Covariation

The "empty cells" of the ANOVA framework meant that experimental participants in the McArthur paradigm could "fill" these cells with assumptions from their own general world knowledge. As Hilton and Slugoski (1986) demonstrated, these knowledge-based assumptions could have dramatic effects on the attribution process. Consider first the following example of a high-consensus, low-distinctiveness, and high-consistency (HLH) information configuration:
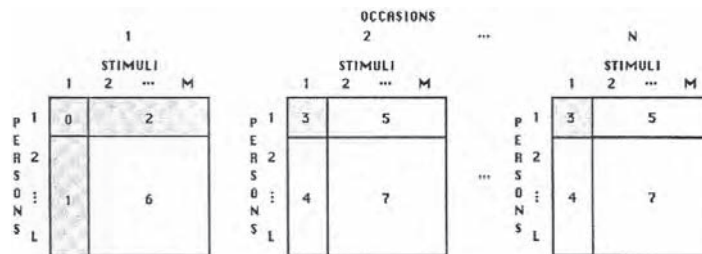
FIGURE 10.2. The missing dimensions of covariation in Kelley's cube (Cheng & Novick, 1990).

Ralph trips up over Joan dancing

Almost everyone who dances with her trips up over Joan dancing

Ralph trips up over almost everyone else he dances with

Ralph almost always trips up over Joan dancing

This configuration produces strong attributions in McArthur's (1972) experiment to "some combination of Ralph and Joan" as the probable cause of the outcome. The high-consensus information is judged as very informative about Joan (Everyone trips up over her) just as the low distinctiveness is informative about Ralph (he trips up over everyone else), leading them both to be judged as clumsy (Hilton & Slugoski, 1986, expt. 1; Hilton, Smith, & Kim, 1995).

However, consider the formally equivalent HLH configuration below for a behavior that is quite normal in its context:

Sally buys something on her visit to the supermarket

Almost everyone buys something on their visit to this supermarket

Sally buys something on her visit to almost every supermarket she visits

Sally almost always buys something on her visit to this supermarket

In this case, when given the option, people tend to shrug their shoulders and conclude that "Nothing special about Sally, the supermarket or the present occasion" caused her to buy something on her visit (Hilton & Slugoski, 1986, expt. 2). Here the HLH covariation information is redundant with socially shared scripts (Schank & Abelson, 1977) about what people normally do, and each of the three items of covariation information is judged as uninformative about Sally, the supermarket, and the present occasion. Indeed, giving such redundant information violates basic rules of conversation that require an interlocutor to be informative (Grice, 1975; Hilton, 1990; Turnbull, 1986; Turnbull & Slugoski, 1988). Indeed, nothing needs to be explained here except why the *speaker* was so obtuse as to pose such a nonquestion in the first place! (Answer: He's an attribution theorist . . .).

Drawing on the analysis of Hart and Honoré (1985), Hilton and Slugoski (1986) argued that the layperson uses knowledge about what is normal to identify the *abnormal conditions* that caused an event to happen, and these are dignified as causes. Thus while the roles of Ralph, Joan, and the present occasion are all necessary conditions for the trip to occur, only Ralph and Joan can figure as explanations, because they are unusually clumsy dancers. In contrast, there is nothing unusual about the present occasion, as Ralph almost always trips up over Joan dancing. In particular, Hilton and Slugoski showed that Ralph and Joan were seen as two sufficient causes in the foregoing example: The preferred response option was "Something about Ralph and something about Joan (even when they are not together)" caused them to trip over each other dancing. Hilton (1988) showed that

these generalizations from world knowledge, called norms, could be used to fill out the missing information in the ANOVA framework. Two examples are shown in Figure 10.3, corresponding respectively to the dancing and the supermarket examples, but collapsing over the time (consistency) dimension for simplicity. They show clearly that in the dancing example, an ANOVA on the high-consensus, low-distinctiveness (HL) data matrix that includes the norm *People sometimes trip up over each other dancing* would yield two "main effect" explanations ("Ralph" and "Joan"), whereas one on the high consensus, low distinctiveness (HL) data matrix that includes the norm *People usually buy something on their visits to supermarkets* would yield no effects at all ("Nothing special about either Sally, the supermarket of the present occasion caused her to buy something on her visit to the supermarket").

In subsequent research that both presented information in all eight cells and used full response formats (Cheng & Novick, 1990; Försterling, 1989, 1992), participants' attributions were found to correspond closely to the patterns that would be expected from a normative ANOVA. But by the time that the twin stable doors of full configurations of covariation information and complete
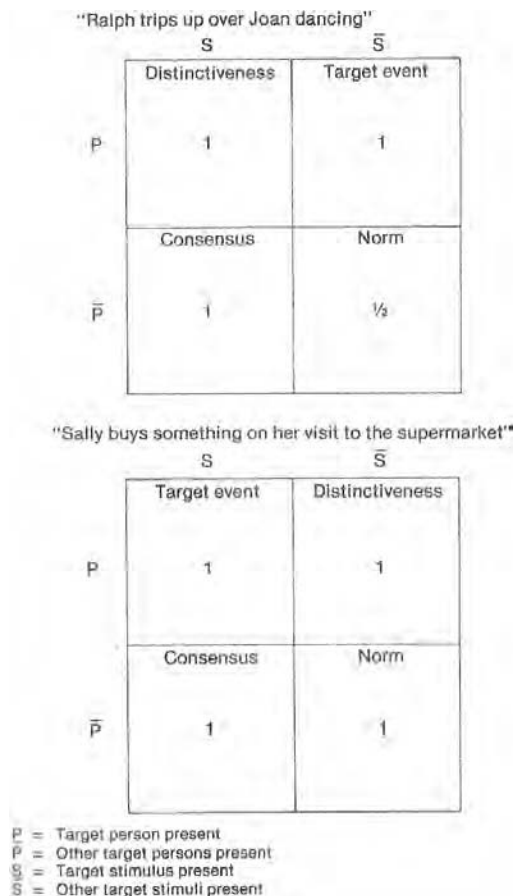
FIGURE 10.3. Covariation matrices completed by implicit knowledge of norms (Hilton, 1988).

response sets had been closed, the horse of causal reason had bolted into the fields of judgmental bias and error. To further muddy these fields, two further problems caused confusion: namely, the influence of experimental methodology on the attributions that were produced and the failure to make a clear distinction between causal explanations and dispositional attributions.

The Use of Covariation Information in Causal Inference

A first issue was that in McArthur's (1972; 1976) experiments with incomplete covariation configurations and response sets, consensus information appeared to have much less effect than distinctiveness information, thereby contributing to the impression that people "underuse" consensus information in causal judgment (Nisbett & Borgida, 1975; Nisbett, Borgida, Crandall, & Reed, 1976), an impression that has had a long shelf life (Fiske & Taylor, 1991; Försterling, 2001). In fact, experiments that present complete response sets show strong effects of consensus information on causal explanations referring to the person when both incomplete (Hewstone & Jaspars, 1987; Hilton & Jaspars, 1987; Hilton et al., 1995; Jaspars, 1983; see Iacobucci & McGill, 1990, for supplementary analyses and a review) and complete covariation configurations are presented (Cheng & Novick, 1990; Försterling, 1989, 1992; van Overwalle, 1997).

In establishing the consensus–person and distinctiveness–stimulus inferential links, the aforementioned experiments also clarified the *direction* of effects of covariation information on causal attributions. Thus, Orvis and colleagues' (1975) "template-matching" model had postulated that high-consensus information should lead to stimulus attributions and low-distinctiveness information to person attributions. However, Orvis and colleagues had based their model on an inductive analysis of McArthur's (1972) results, which were of course biased by her incomplete response methodology (see above). Orvis and colleagues did, however, correctly postulate that low consistency would lead to circumstance (occasion, time) attributions, as would be predicted by Mill's method of difference.

Causal Explanation versus Dispositional Attribution from Covariation Information

However, introducing a distinction between causal explanations and dispositional attributions helps resolve the foregoing inconsistencies. While Orvis and colleagues' (1975) analysis and others (e.g., Anderson, 1978; Medcof, 1990) that proposed consensus–stimulus and distinctiveness–person inference rules have proved to be incorrect concerning *causal explanations*, these inference rules conform to fundamental intuitions concerning *dispositional attribution*, where "dispositions" can be properties of a person or entity to produce a target behavior in a wide range of situations. Whereas causal explanations depend on Mill's method of difference (which also underlies counterfactual reasoning), the induction of general dispositions requires Mill's method of agree-

ment (Hilton et al., 1995; van Overwalle, 1997). Thus if I notice that Sung-Hee is afraid of Fido, Richard is afraid of Fido, and even Mark is afraid of Fido, all these observations of low distinctiveness in fear responses *agree* in one respect: Fido is always present. We can conclude that Fido seems to have a general tendency to produce fear, just as supermarkets have a general tendency to produce shopping behavior.

Of course, to characterize Fido as a *fierce* dog, we have to compare him to the average dog, because all dogs are fierce to some extent. Trait attributions presuppose implicit comparison standards—a rich Sri Lankan is likely to have less money than a rich American, and maybe even less than a poor American, even though she is rich by Sri Lankan standards (Rips & Turnbull, 1980; see also Jones & McGillis, 1976). Consequently, such trait attributions require the application of both the method of difference (showing that Fido is fiercer than the average being in his category, in this case dogs) and the method of agreement (showing that Fido induces fear in a lot of people). Participants do indeed follow these rules of inference for dispositional attribution to the person, the stimulus, or even the occasion (Hilton et al., 1995; van Overwalle, 1997).

The Functions of Covariation Information: When and Why Is It Needed?

Numerous information search studies showed that people consistently use distinctiveness information to answer causal questions about the person and consensus information to answer causal questions about the stimulus (Alicke & Insko, 1984; Bassili & Regan, 1977; Garland, Hardy, & Stephenson, 1975; Hansen, 1980). This of course appears to be inconsistent with the consensus–person and distinctiveness–stimulus inference rules based on the method of difference. However, Hilton, Smith, and Alicke (1988) argued that if participants could already infer relevant consensus information from world knowledge to identify "something special" about the target person, then they would then search distinctiveness information to check the reliability of their judgment. For example, if I learn that Severiano scores 67 on the Marbella Natural Alhambra golf course, and I already know that par (the norm for good players on golf courses round the world) is 72, then I can infer the relevant consensus information that Severiano is a good player because he did better than other players on this golf course. However, to check the reliability of this inference, I need to know whether Severiano usually does well on this course (consistency information), and how he does on other courses (distinctiveness information). Thus, if I do not have the requisite general knowledge about par, I will first need to know how other people did on this course (consensus information) to evaluate whether Severiano did well or not.

Hilton and colleagues (1988) accordingly generated fictitious sporting scenarios where participants were given the equivalent of golfing par for a number of sporting accomplishments (e.g., the average time taken to run a track of a given type). As predicted, when participants

were uninformed of the relevant (track) norms, they were more likely to seek consensus information when asked a causal dispositional question "How good an interstate cross-country runner is Mike Gatting?" but sought distinctiveness information when they had the relevant norms. Prior knowledge of norms affected preference for consensus information only when the question was causal: When a social comparison question was asked that required participants to evaluate the quality of the target's performance, consensus information was always preferred, regardless of whether or not participants had been informed of the norm (see Smith, Hilton, Kim, & Garonzik, 1992, for an extension of this analysis to causal inferences about the stimulus or situation). A final implication of this analysis is that as children acquire world knowledge, the focus of their search in response to causal questions about the person should change from consensus to distinctiveness information. This is exactly the pattern observed by Hortacsu (1987): the information search strategies of children at the age of 9 years corresponded to that of the adults in Hilton and colleagues' "normless" condition, whereas those of children at the age of 17 years corresponded to that of adults in the "normful" condition.

Implications of the ANOVA Framework: From Hydraulics to Rarity

The ANOVA model clearly presupposes that person, stimulus (situation), and occasion causes are manipulated orthogonally, as in an experimental design. Consequently, in this framework there is no reason to expect that the presence of a "person" effect should imply the absence of a "stimulus" or "occasion" effect. This is clearly in contradiction to a "hydraulic assumption" about causality, where internal and external causal attributions should "compete" for causal strength, and strengthened belief in one cause should lead to discounting of the other. However, research has disconfirmed the general validity of the hydraulic assumption: Thus when internal and external attributions are measured separately, they do not correlate (e.g., Fincham, 1985; Lalljee, Watson, & White, 1982; Miller, Smith, & Uleman, 1981; Wimer & Kelley, 1982). Nevertheless, despite its conceptual incompatibility with the ANOVA model and its lack of empirical support, the hydraulic model has become ingrained in attributional methodology, for example, through the use of bipolar scales to measure internal versus external attributions (see McClure, 1998, for a review).

In fact, the ANOVA framework suggests that event normality will determine *how much* explanation is needed. Normal or predicted events do not need explanation (Hastie, 1984; Hilton & Slugoski, 1986; Kanazawa, 1992; Weiner, 1985). Conversely, "extreme" events require more explanation than "moderate" ones (McClure, Jaspars, & Lalljee, 1993). In terms of the ANOVA framework, when covariation information shows that an event is rare (e.g., there is low consensus, high distinctiveness, and low consistency), then participants prefer interaction explanations that refer to a combination of the person, stimulus, and occasion (Cheng & Novick, 1990; Försterling, 1989; Hilton & Jaspars, 1987; Jaspars, 1983). It seems that people may prefer such conjunctive explanations when covariation information suggests that a combination of several unusual factors is necessary to produce an event, and the event to be explained is rare, extreme, or has multiple backgrounds (McGill, 1990).

Conclusions: Implications for Attribution Research

The refinements of the ANOVA model point to several important conclusions. First, normal adult attribution is *knowledge based* in that people use prior world knowledge in the form of universal generalizations stored in scripts, plans, goals, and norms (Hilton & Slugoski, 1986; Schank & Abelson, 1977). Second, once Kelley's (1967) ANOVA analogy had been properly specified and tested, either with complete information sets (Cheng & Novick, 1990; Försterling, 1989) or incomplete sets fleshed out with world knowledge (Hilton & Slugoski, 1986; Novick, Fratianne, & Cheng, 1992), people's causal inferences appear more rational than was previously thought. Third, a distinction between causal explanations and dispositional attributions reveals that these are different causal questions that require different rules of inference and informational premises. Both information search and utilization studies show that people do search and use the appropriate information (Hilton et al., 1988, 1995; Smith et al., 1992; van Overwalle, 1997). Fourth, the ANOVA framework is logically compatible with the coexistence of many factors that combine to produce an effect but incompatible with the monocausal "hydraulic" model of causality.

In conclusion, the early models of the attribution process failed to explicate the difference between causal explanation and dispositional attribution, the role of Mill's methods of difference of causal inference, and the significance of base rate information in the form of presupposed world knowledge. Yet it was the layperson, not the attribution theorist, who found herself accused of being a "poor scientist" who failed to follow normative models, and who underused consensus and base rate information in the attribution process (Fiske & Taylor, 1991; Nisbett & Ross, 1980; Ross, 1977; Trope & Gaunt, 2003). The systematic nature of this misunderstanding and misrepresentation of the nature of lay causal inference is perhaps to be found in the metaphysical bases of current experimental psychology. Experimental psychologists typically manipulate situational factors and seek to eliminate "noise" from participants through procedures such as random assignment of participants to experimental conditions. Consequently, many attribution researchers have attempted to understand attribution processes solely on the basis of experimentally presented information. Yet it is precisely the *interaction* between the prior knowledge that the experienced layperson brings to bear and information available in the judgmental situation that constitutes the mill of the knowledge-based causal attribution process.

The conceptual and methodological problems of the early models and experiments continue to generate am-

biguity in the interpretation of attribution phenomena. For example, Choi and Nisbett (1998) cite a Korean study by Cha and Nam (1985) that found extensive use of consensus information using McArthur's response methodology, which showed "underuse" of consensus information in American participants. However, strong use of consensus information has been found in American subjects when complete response methodologies are used (Hilton et al., 1995). This suggests that cross-cultural research in the area would benefit from closer attention to the more recent conceptualizations and methodology. Similarly, the surprise evoked by the finding that East Asian participants still show correspondence bias in Jones and Harris's (1967) attitude attribution task (Krull et al., 1999) would probably be clarified if researchers adopted the normative model and experimental methodology proposed by Morris and Larrick (1995), which I turn to below.

FIGURE 10.4. Comparison of predicted and actual disobedience rates.

KNOWLEDGE-BASED DISCOUNTING: A NORMATIVE FRAMEWORK

One of the chief reproaches made against the "lay scientist" has been his or her failure to sufficiently discount attributions to the person in the light of situational information. Perhaps the most striking demonstration comes from Milgram's (1974) study of obedience to authority. To everyone's surprise (including Milgram's), in the classic version of this study 26 of 40 participants in the "teacher" condition went up to the full level of shock despite the clear indications that this was dangerous to the "learner." All teachers of social psychology know of the surprise value of this experiment, which is to make the point that even ordinary people can be led to acts of barbarism under appropriate situational pressure. Observers who learned about the experimental paradigm "showed the fundamental attribution error, that is, they consistently and dramatically underestimated the degree to which subjects in general would yield to those situational forces that compelled obedience in Milgram's situation" (Ross & Anderson, 1982, p. 136). Jones (1990) considered the fundamental attribution error to result from "correspondence bias," which he nominated as "a candidate for the most robust and repeatable finding in social psychology" (p. 138).

Correspondence Bias: Inaccurate Beliefs or Faulty Inference Processes?

However, it is important to note that this attribution error is more likely to be due to faulty premises (in the form of inaccurate beliefs about human nature) than to flaws in the attributional reasoning process itself. Given the erroneous subjective expectancies shown in Figure 10.4, which amount to the low-consensus information that *Few people go to the maximum shock level*, it is quite easy to see how observers would quite "logically" infer from Mill's method of difference that the cause of a teacher's going to the limit would be "something about the teacher" (Bierbrauer, 1979). The fundamental attribu-

tion error appears to result from the wrong premises—inaccurate prior beliefs about the power of situational pressure in the Milgram experiment—which lead logical inference processes to the wrong causal conclusion.
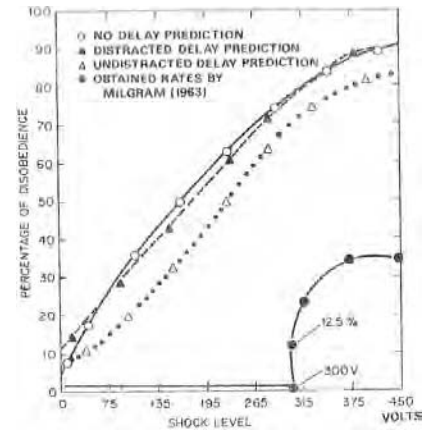
However, apart from the Milgram study and some other exceptions (e.g., Shweder, 1977) there are few actuarial data that enable assessment of whether people's general theories about human nature or other aspects of their environment are correct or not. But we can still assess whether people's attributions are consistent with their situational theories if we have an appropriate normative model to derive causal inferences from subjective beliefs about general cause–effect relations. This will allow us to evaluate whether people's causal inference processes from their subjective beliefs are consistent with those beliefs, even if we have no actuarial data to evaluate the accuracy of those beliefs (in the form of situational theories) per se. In this way we can at least evaluate whether the judge has used faulty inference processes that lead her into self-contradiction, in the form of a conclusion that does not follow from her assumed premises (cf. Henle, 1962).

Rational Discounting in the Attitude Attribution Paradigm

The "attitude attribution" paradigm devised by Jones and Harris (1967) has frequently been adduced to support the existence of a fundamental attribution error (e.g., Ross & Anderson, 1982). In the original version participants read about a group of students who either had a choice or no choice in writing an essay for or against Fidel Castro. In the no-choice condition, participants were told that students had been assigned at random to two experimental conditions, whereby half were obliged to write an essay for and half against Castro. The key finding is that in no-choice conditions, participants are more likely to infer that students who wrote pro-Castro essays are more likely to have pro-Castro attitudes than students who wrote anti-Castro attitudes. Thus the information about situational pressure made little differ-

ence to the tendency to attribute the essay to underlying attitudes, as this tendency was almost as strong as that obtained in choice conditions. But is this an error, or even a bias?

In fact, a striking analysis and experiment by Morris and Larrick (1995) suggests that we can eliminate inferential error as a cause of participants' judgments in Jones and Harris's attitude attribution paradigm. They show that given participants' beliefs about situation–behavior linkages in this experimental paradigm, their dispositional attributions are just what would be expected by a Bayesian analysis of discounting. Morris and Larrick's analysis of discounting takes four factors into account: the prior probability of causes; the sufficiency of the causes for effects in question; the independence of causes; and the number of possible causes for an effect. Of these, the prior probability of the causes (a pro-Castro disposition, and instructions to write a pro-Castro essay) and the perceived sufficiency for writing a pro-Castro essay are directly relevant to the Jones and Harris (1967) discounting paradigm, as random assignment to experimental conditions ensures that the causes are independent (there is no reason to assume that there are more Castro sympathizers in the forced-choice condition compared to the free-choice condition), and only two causes are being considered. Assessments of participants' beliefs showed that the target cause (the underlying attitude) was perceived as relatively rare (people assumed that 23% of the population would be pro-Castro) whereas the situational cause is relatively common (participants knew that 50% of the participants in the no-choice condition had been instructed to write a pro-Castro essay). Morris and Larrick then solicited participants' beliefs about general situation–behavior linkages in the population tested, namely, their expectancies that participants with a pro-Castro attitude would write a pro-Castro essay and that a participant asked to write a pro-Castro attitude would do so. Given participants' beliefs that pro-Castro attitudes are in the minority (23%), and the causes of a pro-Castro essay were strongly but imperfectly sufficient for compliance to occur, Morris and Larrick were able to calculate what a Bayesian inference model would predict. As can be seen from Table 10.1, participants' judgments correspond very closely indeed to the "rational baseline" (Jones & McGillis, 1976) defined by the normative Bayesian model.

In a second experiment, Morris and Larrick (1995) demonstrated that the "hydraulic assumption" that motivates discounting will be appropriately used in cases in which the co-occurrence of causes in the world is known to be inversely related. Thus they manipulated association between causes by informing participants that (1) causes were independent, as assignment of experimental participants to instruction conditions was randomized; (2) causes were positively associated, as 80% of the participants assigned to write pro-Castro essays already held pro-Castro attitudes; and (3) causes were negatively associated, as 80% of those associated to write anti-Castro essays held pro-Castro views. In line with a normative Bayesian model of causal inference, participants showed no evidence of discounting when the causes were in-

TABLE 10.1. Prior Beliefs, Beliefs in the Sufficiency of Causes, and Dispositional Attributions in Morris and Larrick (1995, expt. 1)

Prior beliefs		
p (Pro-Castro)	.23	
p (Instructed Pro-Castro)	.50	
Perceived sufficiency		
p (Essay/Attitude + Instruction)	.81	
p (Essay/Instruction + Attitude)	.85	
		Dispositional attributions
Rational benchmark		
p (Attitude/Essay)	.35	.36
p (Attitude/Essay + Instructions)	.30	.26

dependent (in line with an ANOVA-style orthogonal manipulation of causes), a "hydraulic" pattern of discounting when the causes were negatively associated (in line with the hydraulic assumption), and an inverse pattern of "overdiscounting" when the causes were positively associated. Morris and Larrick's normative approach thus suggests that, given the participants' premises, there is no inferential error.

Morris and Larrick's results suggest that if there is a fundamental attribution error (FAE), it is due to incorrect situational theories, not to faulty inference processes from those theories. In addition to incorrect situational theories, Gilbert and Malone (1995) argue that the FAE has three other major causes: lack of awareness of situational constraints; incomplete corrections of dispositional inferences; and inflated categorizations of behavior. For reasons of space I only review the first two of these and refer the interested reader to relevant reviews (Gawronski, 2004; Pelham & Krull, 2000). Apart from the attitude attribution paradigm, evidence for the FAE has been chiefly drawn from three basic experimental paradigms: the silent interview paradigm (Snyder & Frankel, 1976); the quizmaster paradigm (Ross, Amabile, & Steinmetz, 1977); and attributions for immoral behavior (Bierbrauer, 1979).

The Silent Interview Paradigm: The Effect of Question Focus and Cognitive Resources

A research program by Krull and colleagues suggests that attributional processes can be biased by question focus to either the person or situation, but these tendencies can be corrected if the judge has sufficient cognitive resources. They found that when participants who had a dispositional attribution task watched a silent videotape of a woman purportedly being interviewed about a topic, they made stronger dispositional attributions for the target's emotional expressions and behavior when concurrent memory load was high rather than low (Gilbert, Pelham, & Krull, 1988). Conversely, when only situational questions were asked, the reverse pattern of results was obtained: participants made stronger attributions to the

situation under high rather than low memory load (Krull, 1993; Krull & Erickson, 1995).

Results obtained from the "silent interview" paradigm therefore do not overall support the idea of there being a "dispositional set" in the sense of their being an automatic tendency to make dispositional attributions *per se*, the interpretation initially favored by Gilbert and colleagues (1988). Rather, these results are consistent with general information integration models. Thus if the dispositional question is posed first, cognitive load will prevent integration of situational information into the judgment process, whereas if the situational question is posed first, cognitive load will prevent integration of dispositional information (Trope & Alfieri, 1997; Trope & Gaunt, 1999, 2000). In line with this view, participants high in need for closure may truncate these later information integration processes even when they are not under cognitive load, whether the initial question focus is on the person or the situation (Webster, 1993).

Lack of Awareness of Situational Constraints: The "Questioner Superiority" Effect?

Ross and colleagues (1977) assigned participants arbitrarily to the role of a quizmaster, contestant, or observer in a mock game show. The quizmaster is then instructed to think up 10 challenging general-knowledge questions and pose them to the contestant, who is usually unable to answer more than four questions correctly. The most interesting finding is that participants tend to perceive the quizmaster as having more general knowledge than the contestant, even though it is the quizmaster who can choose which questions to ask. Gilbert and Malone (1995) argue that observers fail to see "the invisible jail" in which the questioners were trapped, and so "concluded that the quizmasters were genuinely brighter than the contestants" (p. 25). However, Gawronski (2003) varied the difficulty of the questions asked and found while quizmasters were overall judged to have more general knowledge than contestants (thus replicating the questioner superiority effect), this effect was considerably reduced when the questions were difficult, thus suggesting that the observers were well aware of the "jail" in which contestants had been placed.

Motive in Attributions about Morality

Work on motive attributions shows that certain kinds of situational cause will have different effects on dispositional attribution. For example, Reeder, Kumar, Hesson-McInnis, and Trafimow (2002) show that aggression in response to a provocation (e.g., an insult) leads to less attributions of immorality than aggression that was motivated by a financial incentive. Motive attributions also lead to discounting in the Jones and Harris (1967) attitude inference paradigm; thus Fein (1996) reports data showing that people do not commit the correspondence bias in when they are led to suspect that the actor wrote the essay for personal gain. In a like vein, Reeder, Vonk, Ronk, Ham, and Lawrence (2004) show that learning that a student had a ulterior motive for helping a profes-

sor (she knows this will improve her chances of winning a monetary award) leads to discounting (people consider her to be a less helpful person), whereas learning that she helped because she was instructed to by her boss (the departmental secretary) does not lead to such discounting. Although the situational cause was rated equally highly in both cases, the kind of motive they respectively implied had different effects on dispositional attributions. The finding that dispositional attributions were still quite strong in both motive conditions indicates that people can consider situational factors to be strongly causal *and* to have major implications about the actor's personality. Such patterns of causal explanation and dispositional attribution sit uneasily with the hydraulic model of internal versus external causality and seem more consistent with the idea that people form complex representations of how events come about.

Summary: Discounting and the Fundamental Attribution Error

Having accepted that people often do discount (see McClure, 1998, for a review), work in the area seems to be moving in the direction of (1) distinguishing causal discounting from causal backgrounding; (2) establishing normative models that allow us to decide when people do discount appropriately; and (3) identifying factors such as resource limitations and processing style that may truncate this discounting process. At the descriptive level, the universality of the correspondence bias has been called into question; rather, people discount either dispositional or situational causes as a function of the information-processing task, whether they have adequate information-processing resources, and the perceived relevance of information. Even if, as Lieberman and colleagues (2002) argue, people make more dispositional inferences in Western society because these kinds of questions get asked more often, this is a *sociological* fact about cultural orientations, not a psychological one about cognitive biases.

Nevertheless, in at least one domain, that of Milgram's experiment, people fail to discount situational pressures adequately, such as when they attribute obedience to the brutal personality of the confederate. This may well be because people hold inaccurate "situational theories" about behavior in this situation rather than due to faulty inference processes from these informational premises. More research on the accuracy of situational theories thus seems to be warranted, especially as there is still very little research that investigates either the accuracy of situational theories across domains or how these situational theories would be deployed in a Bayesian judgment process of the kind described by Morris and Larrick (1995). Other lines of inquiry that seem to warrant investigation are whether people reason from qualitatively different expectations or "norms" in certain situations, such as moral behavior, or whether correspondence bias will be particularly pronounced when the attributor will wish to exert social control over the actor by holding him or her responsible for his actions (Tetlock, 2000). In sum, much theoretical and empirical work remains to be done to

fully understand whether, when, and why discounting happens.

CONCLUSIONS

Causal explanation remains a fundamental topic of research in philosophy, psychology, and cognitive science. Research in these fields has added to our understanding of attribution and explanation processes since the classic "lay scientist" formulations. It has shown that although the human mind is equipped with associative learning heuristics that are economical (though not infallible) general-purpose methods for detecting covariations that are shared with other animals, humans seem to have a specific social cognition module, are capable of counterfactual simulations to evaluate explanations, and are able communicate explanations that take into account others' knowledge and viewpoints.

Improvements in clarifying the nature of explanation and in formulating normative models of the attribution process have shown that even when people are required to perform "scientific" causal analyses, they do so rather better than earlier work would suggest (e.g., Jones & Harris, 1967; McArthur, 1972). So it is the layperson who uses world knowledge to integrate missing base-rate information into the ANOVA-style attribution process when the experimenter omits to provide it (Cheng & Novick, 1992), and who uses this world knowledge to discount situational causes in a way that is compatible with Bayes's theorem (Morris & Larrick, 1995). When required to, the layperson seems better able to perform full-blown "scientific" analyses of his environment in order to arrive at reliable and accurate representations of reality than he would be credited with by early formulations of the "lay scientist" model.

I conclude by noting that the fact that people can be quite accurate in causal inference invites reconsideration of issues in attribution*al* theory, the study of how attributions impact emotions and behaviors (see Anderson, Krull, & Weiner, 1996; Försterling, 2001; Kelley & Michela, 1980). For example, Försterling (2002) reviews evidence that differences in patterns of attribution reflect "logical" inferences from the assumed patterns of covariation that people have, not differences in their causal inference processes.

In addition, accurate causal inference processes can help people adapt better to their world, in contrast to Taylor and Brown's (1988) suggestion that attributional biases in the form of "positive illusions" can be adaptive for health outcomes.

Thus Försterling and Morgenstern (2002) have shown that realism in causal attributions can lead to more favorable performance outcomes. They show that giving participants accurate rather than self-enhancing feedback for task performance in a learning phase enables individuals to specialize in tasks that they are good at in the subsequent test phase, and thus improve their chances of gaining a financial reward for the best performance. One of the beneficial effects of realism is that accurate feedback may help individuals to identify and invest in areas

of competitive advantage, whereas positive illusions may lead individuals to overestimate their chances of success and thus lose their investments. A question for future research will be to identify environments in which accuracy in attributions is more likely to be adaptive than positive illusions, and vice versa.

REFERENCES

Abelson, R. P. (1975). Concepts for representing mundane reality in plans. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science.* New York: Academic Press.

Abelson, R. P., & Lalljee, M. G. (1988). Knowledge structures and causal explanation. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality.* (pp. 175–203). Brighton, UK: Harvester Press.

Adolphs, R., Bechara, A., Tranel, D., & Damasio, A. (1995). Neuropsychological approaches to decision-making. In A. R. Damasio, H. Damasio, & Y. Christen (Eds.), *Neurobiology of decision-making* (pp. 157–169). New York: Springer-Verlag.

Alicke, M., & Insko, C. (1984). Sampling of similar and dissimilar comparison objects as a function of the generality of the attribution goal. *Journal of Personality and Social Psychology*, *46*, 763–777.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Science*, *4*, 267–278.

Anderson, C., Krull, D., & Weiner, B. (1996). Explanations: Processes and consequences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 271–297). New York: Guilford Press.

Anderson, N. H. (1978). Progress in cognitive algebra. In L. Berkowitz (Ed.), *Cognitive theories in social psychology* (pp. 103–126). New York: Academic Press.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*, 379–398.

Bassili, J., & Regan, D. T. (1977). Attributional focus as a determinant of information selection. *Journal of Social Psychology*, *101*, 113–121.

Bierbrauer, G. (1979). Why did he do it? Attribution of obedience and the phenomenon of dispositional bias. *European Journal of Social Psychology*, *11*, 427–38.

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. In C. Frith & D. Wolpert (Eds.), *The neuroscience of social interaction: Decoding, imitating, and influencing the actions of others* (pp. 241–264). Oxford, UK: Oxford University Press.

Blakemore, S.-J., Fonlupt, P., Pachot-Clouard, M., Darmon, C., Boyer, P., Meltzoff, A. N., et al. (2001). How the brain perceives causality: An event-related fMRI study. *NeuroReport, 12*, 3741–3746.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*, 237–273.

Brown, R., & van Kleeck, R. (1989). Enough said: Three principles of explanation. *Journal of Personality and Social Psychology*, *57*, 590–604.

Carbonell, J. G. (1981). Politics. In R. C. Schank & C. K. Riesbeck (Eds.), *Inside computer understanding: Five programs plus miniatures*. Hillsdale, NJ: Erlbaum.

Carpendale, J., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding of mind within social interaction. *Behavioral and Brain Sciences*, 27, 79–150.

Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perceptions and interpretation of complex intentional movement patterns. *NeuroImage*, 12, 314–325.

Catellani, P., Alberici, I. A., & Milesi P. (2004). Counterfactual thinking and stereotypes: The nonconformity effect. *European Journal of Social Psychology*, 34, 421–436.

Cha, J.-H., & Nam, K.-D. (1985). A test of Kelley's cube theory of attribution: A cross-cultural replication of McArthur's study. *Korean Social Science Journal*, 12, 151–180.

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, 58, 545–567.

Cheng, P. W., & Novick, L. R.(1991). Causes versus enabling conditions. *Cognition*, 40, 83–120.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, 99, 65–82.

Choi, I., & Nisbett, R. E. (1998). Situational salience and cultural differences in the correspondence bias. *Personality and Social Psychology Bulletin*, 24, 949–960.

Chun, W-Y., Spiegel, S., & Kruglanski, A. (2002). Assimilative behavior identification can also be resource dependent: A unimodel perspective on personal attribution phases. *Journal of Personality and Social Psychology*, 83, 542–555.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99, 1–19.

Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology*, 70, 1164–1184.

Fincham, F. D. (1985). Attribution processes in distressed and nondistressed couples: 2. Responsibility for marital problems. *Journal of Abnormal Psychology*, 94, 183–190.

Fiske, S., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Försterling, F. (1989). Models of covariation and attribution: How do they relate to the analysis of variance? *Journal of Personality and Social Psychology*, 57, 615–26.

Försterling, F. (1992). The Kelley model as an analysis of variance analogy: How far can it be taken? *Journal of Experimental Social Psychology*, 28, 475–490.

Försterling, F. (2001). *Attribution: An introduction to theories, research, applications*. Hove, UK: Psychology Press.

Försterling, F. (2002). Does scientific thinking lead to success and sanity? An integration of attribution and attributional models. *European Journal of Social Psychology*, 13, 217–258.

Försterling, F., & Morgenstern, M. (2002). Accuracy of self-assessment and task performance: Does it pay to know the truth? *Journal of Educational Psychology*, 94, 576–585.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. In C. Frith & D. Wolpert (Eds.), *The neuroscience of social interaction: Decoding, imitating, and influencing the actions of others* (pp. 45–76). Oxford, UK: Oxford University Press.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and non-verbal tasks. *Neuropsychologia*, 38, 11–21.

Garland, H., Hardy, A., & Stephenson, L. (1975). Information search as affected by attribution type and response category. *Personality and Social Psychology Bulletin*, 1, 612–615.

Gawronski, B. (2003). On difficult questions and evident answers: Dispositional inference from role-constrained behavior. *Personality and Social Psychology*, 29, 1459–1475.

Gawronski, B. (2004). Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology*, 15, 183–217.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995).Taking the intentional stance at 12 months of age. *Cognition*, 56, 165–193.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 89–150). New York: McGraw-Hill.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias: The what, when, how and why of unwarranted dispositional inference. *Psychological Bulletin*, 117, 21–38.

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, 54, 733–740.

Graesser, A. C., Robertson, S. P., & Anderson, P. A. (1981). Incorporating inferences in narrative representations: A study of how and why. *Cognitive Psychology*, 13, 1–26.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics. 3: Speech acts* (pp. 41–58). New York: Wiley.

Grossman, E., & Blake, R. (2002). Brain areas active during visual perception of biological stimuli. *Neuron*, 35, 1167–1175.

Halpern, J., & Pearl, J. (2005a). Causes and explanations. A structural model approach. Part I: Causes. *British Journal for the Philosophy of Science*, 56(4), 843–887.

Halpern, J., & Pearl, J. (2005b). Causes and explanations. A structural model approach. Part II: Explanations. *British Journal for the Philosophy of Science*, 56(4), 889–911.

Hansen, R. D. (1980). Commonsense attribution. *Journal of Personality and Social Psychology*, 39, 996–1009.

Harris, P. (1996). Desires, beliefs and language. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 200–220). Cambridge, UK: Cambridge University Press.

Hart, H. L. A., & Honoré, A. M. (1985). *Causation in the law* (2nd ed.). Oxford, UK: Oxford University Press.

Hastie, R. (1984). Causes and effects of causal attributions. *Journal of Personality and Social Psychology*, 46, 44–56.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–259.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. London: Collier Macmillan.

Hesslow, G. (1983). Explaining differences and weighting causes, *Theoria*, 49, 87–111.

Hesslow, G. (1988). The problem of causal selection. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 11–32). Brighton, UK: Harvester Press.

Hewstone, M. R. C., & Jaspars, J. M. F. (1987). Covariation and causal attribution: A logical model of the intuitive analysis of variance. *Journal of Personality and Social Psychology*, 51, 663–672.

Hilton, D. J. (1988). Logic and causal attribution. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 33–65). New York: New York University Press.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, 107, 65–81.

Hilton, D. J. (2002). Commonsense and scientific thinking about causality: Cognitive and social bases of reasoning. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 211–231). Cambridge, UK: Cambridge University Press.

Hilton, D. J., & Erb, H.-P. (1996). Mental models and causal explanation: Judgments of probable cause and explanatory relevance. *Thinking and Reasoning*, 2, 33–65.

Hilton, D. J., & Jaspars, J. M. F. (1987). The explanation of occurrences and non-occurrences: A test of the inductive logic model of causal attribution. *British Journal of Social Psychology*, 26, 189–201.

Hilton, D. J., McClure, J. L., & Slugoski, B. R. (2005). The course of events: Counterfactuals, causal sequences and explanation. In D. Mandel, D. J. Hilton, & P. Catellani (Eds.), *The psychology of counterfactual thinking* (pp. 44–60). London: Psychology Press.

Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, *93*, 75–88.

Hilton, D. J., & Slugoski, B. R. (2001). The conversational perspective in reasoning and explanation. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Vol. 1. Intrapersonal processes* (pp. 181–206). Oxford, UK: Blackwell.

Hilton, D. J., Smith, R. H., & Alicke, M. D. (1988). Knowledge-based information acquisition: Norms and the functions of consensus information. *Journal of Personality and Social Psychology*, *55*, 79–92.

Hilton, D. J., Smith, R. H., & Kim, S-H. (1995). The processes of causal explanation and dispositional attribution. *Journal of Personality and Social Psychology*, *68*, 377–387.

Hortacsu, N. (1987). Attributional focus and information selection in relation to chronological age. *Child Development*, *58*, 225–233.

Hume, D. (1888). *A treatise of human nature* (L. A. Selby-Bigge, Ed.). Oxford, UK: Oxford University Press. (Original work published 1739)

Iacobucci, D., & McGill, A. L. (1990). Analysis of attribution data: Theory testing and effects estimation. *Journal of Personality and Social Psychology*, *59*, 426–441.

Jaspars, J. M. F. (1983). The process of attribution in common-sense. In M. R. C. Hewstone (Ed.), *Attribution theory: Social and functional extensions* (pp. 28–44). Oxford, UK: Blackwell.

Jaspars, J. M. F., Hewstone, M. R. C., & Fincham, F. D. (1983). Attribution theory and research: The state of the art. In J. M. F. Jaspars, F. D. Fincham, & M. R. C. Hewstone (Eds.), *Attribution theory: Conceptual, developmental and social dimensions* (pp. 3–36). London: Academic Press.

Jones, E. E. (1990). *Interpersonal perception*. New York: Freeman.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology*, *63*, 220–266.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1–24.

Jones, E. E., & McGillis, D. (1976). Correspondent inference and the attribution cube: A comparative reappraisal. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 389–420). Hillsdale, NJ: Erlbaum.

Kahneman, D. E., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Kanazawa, S. (1992). Outcome of expectancy? Antecedent of spontaneous causal attribution. *Personality and Social Psychology Bulletin*, *18*, 659–668.

Kashima, Y., McKintyre, A., & Clifford, P. (1998). The category of the mind: Folk psychology of belief, desire, and intention. *Asian Journal of Social Psychology*, *1*, 289–313.

Kashima, Y., Siegal, M., Tanaka, K., & Kashima, E. S. (1992). Do people believe behaviours are consistent with attitudes?: Towards a cultural psychology of attribution processes. *British Journal of Social Psychology*, *31*, 111–124.

Kelley, H. H. (1967). Attribution in social psychology. *Nebraska Symposium on Motivation*, *15*, 192–238.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist*, *28*, 107–128.

Kelley H. H. (1983). Perceived causal structures. In J. M. F. Jaspars, M. R. C. Hewstone, & F. D. Fincham (Eds.), *Attribution theory and research*: *Conceptual, developmental and social dimensions* (pp. 343–359). New York: Academic Press.

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457–501.

Keltner, D., Ekman, P., Gonzaga, G. C., & Beer, J. (2003). Facial expression of emotion. In R. J. Davidson, K. R. Scherer, & H. Hill Goldsmith (Eds.), *Handbook of affective sciences* (pp. 415–432). Oxford, UK: Oxford University Press.

Kruglanski, A. W. (1988). Knowledge as a social psychological construct. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 109–141). Cambridge, UK: Cambridge University Press.

Krull, D. S. (1993). Does the grist change the Mill? The effect of the perceiver's inferential goal on the process of social inference. *Personality and Social Psychology Bulletin*, *19*, 340–348.

Krull, D. S., & Erickson, D. J. (1995). Judging situations: On the effortful process of taking dispositional information into account. *Social Cognition*, *13*, 417–438.

Krull, D. S., Hui-Min Loy, M., Lin, J., Wang, C-F., Chen, S., & Zhao, X. (1999). The fundamental attribution error: Correspondence bias in individualist and collectivist cultures. *Personality and Social Psychology Bulletin*, *25*, 1208–1219.

Lalljee, M. G., & Abelson, R. P. (1983). The organization of explanations. In M. R. C. Hewstone (Ed.), *Attribution theory: Social and functional extensions* (pp. 65–80). Oxford, UK: Basil Blackwell.

Lalljee, M. G., Watson, M., & White, P. (1982). Explanations, attributions and the social context of unexpected behaviour. *European Journal of Social Psychology*, *12*, 17–29.

Lehnert, W. G. (1978). *The process of question-answering*. Hillsdale, NJ: Erlbaum.

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). New York: Elsevier.

Mackie, J. L. (1980). *The cement of the universe* (2nd ed.). London: Oxford University Press.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, *3*, 23–48.

Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.

Malle, B. F. (in press-a). The actor–observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin*.

Malle, B. F. (in press-b). Attributions as behavior explanations: Toward a new theory. In D. Chadee & J. Hunter (Eds.), *Toward a new theory: Current themes and perspective in social psychology*.

Mandel, D. R. (2003). Judgment dissociation theory: An analysis of differences in causal, counterfactual, and covariational reasoning. *Journal of Experimental Psychology: General*, *132*, 419–434.

McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attributions. *Journal of Personality and Social Psychology*, *22*, 171–193.

McArthur, L. Z. (1976). The lesser influence of consensus than distinctiveness information: The person–thing hypothesis. *Journal of Personality and Social Psychology*, *33*, 733–742.

McClure, J. L. (1998). Discounting causes of behavior: Are two reasons better than one? *Journal of Personality and Social Psychology*, *47*, 7–20.

McClure, J. L., & Hilton, D. J. (1997). For you can't always get what you want: When preconditions are better explanations than goals. *British Journal of Social Psychology*, *36*, 223–240.

McClure, J. L., & Hilton, D. J. (1998). Are goals or preconditions better explanations? It depends on the question. *European Journal of Social Psychology*, *28*, 897–911.

McClure, J. L., Hilton, D. J., Cowan, J., Ishida, L., & Wilson, M. (2001). When rich or poor people buy expensive objects is the causal question how or why? *Journal of Language and Social Psychology*, *20*, 339–357.

McClure, J. L., Lalljee, M., & Jaspars, J. (1993). Explanations of moderate and extreme events. *Journal of Research in Personality*, *25*, 146–166.

McGill, A. L. (1989). Context effects on causal judgment. *Journal of Personality and Social Psychology*, *57*, 189–200.

McGill, A. L. (1990). Conjunction effects: The effect of comparison of the target episode to a contrasting background instance. *Social Cognition*, *8*, 362–82.

McGill, A. L. (1991). Conjunction effects: Accounting for events that differ from several norms. *Journal of Experimental Social Psychology*, *27*, 527–49.

Medcof, J. W. (1990). PEAT: An integrative model of attribution processes. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 111–210). New York: Elsevier.

Michotte, A. E. (1963). *The perception of causality*. London: Methuen.

Milgram, S. (1974). *Obedience to authority: An experimental view*. London: Tavistock.

Mill, J. S. (1973). System of logic. In J. M. Robson (Ed.), *Collected works of John Stuart Mill* (8th ed., Vols. 7 and 8). Toronto: University of Toronto Press. (Original work published 1872).

Miller, F. D., Smith, E. R., & Uleman, J. S. (1981). Measurement and interpretation of dispositional and situational attributions. *Journal of Experimental Social Psychology*, *32*, 932–943.

Morris, M. W., Ames, D. R., & Knowles, E. D. (2001). What we theorize when we theorize that we theorize: The "lay theory" construct in developmental, social and cultural psychology. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 143–161). Mahwah, NJ: Erlbaum.

Morris, M. W., & Larrick, R. (1995).When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review, 102*, 331–355.

Morris, M. W., Nisbett, R. E., & Peng, K. (1995). Causal attribution across domains and cultures. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 577–612). Oxford, UK: Clarendon Press.

Nisbett, R. E., & Borgida, E. (1975). Attributions and the psychology of prediction. *Journal of Personality and Social Psychology, 32*, 932–943.

Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. (1976). Popular induction: Information is not always informative. In J. Carroll & J. Payne (Eds.), *Cognition and social behavior* (pp. 113–134). Potomac, MD: Erlbaum.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Novick, L., Fratianne, A., & Cheng, P. W. (1992). Knowledge-based assumptions in causal attribution. *Social Cognition, 10*, 299–344.

Oestermeier, U., & Hesse, F. (2001). Singular and general causal arguments. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the Cognitive Science Society* (pp. 720–725). Mahwah, NJ: Erlbaum.

Orvis, B. R., Cunningham, J. D., & Kelley, H. H. (1975). A closer examination of causal inference: The role of consensus, distinctiveness and consistency information. *Journal of Personality and Social Psychology, 27*, 154–164.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.

Pelham, B., & Krull, D. S. (2000). *Practice makes perfect? Judgmental expertise and the correspondence bias*. Unpublished paper.

Peng, K., & Knowles, E. D. (2004). Culture, education and the attribution of physical causality. *Personality and Social Psychology Bulletin, 29*, 1272–1284.

Premack, D., & Premack, A. (1995). Intention as a psychological cause. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 185–199). Oxford, UK: Clarendon Press.

Read, S. J. (1987) Constructing causal scenarios: A knowledge-structure approach to causal reasoning. *Journal of Personality and Social Psychology, 39*, 39–58.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1*, 26–53.

Reeder, G. D., Kumar, S., Hesson-McInnis, M., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology, 83*, 789–903.

Reeder, G. D., Vonk, R., Ronk, M. J., Ham, J., & Lawrence, M. (2004). Dispositional attribution: Multiple inferences about motive-related traits. *Journal of Personality and Social Psychology, 86*, 530–544.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton-Century-Croft.

Rips, L. J., & Turnbull, W. M. (1980). How big is big?: Relative and absolute properties in memory. *Cognition, 8*(2), 145–174.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 173–200). New York: Academic Press.

Ross, L., Amabile, T. M., & Steinmetz, J. L. (1977). Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880–892.

Ross, L., & Anderson, C. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 129–152). Cambridge, UK: Cambridge University Press.

Rudolph, U., & Försterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin, 121*, 192–218.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An enquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Shanks, D. R., & Dickinson, A. (1988). The role of selective attribution in causality judgment. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 94–126). Brighton, UK: Harvester Press.

Shweder, R. A. (1977). Likeness and likelihood in everyday thought: Magical thinking in everyday judgments about personality. *Current Anthropology, 18*, 637–658.

Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379–396). Cambridge, UK: Cambridge University Press.

Sloman, S. A., & Lagnado, D. (2005). Do we "do"? *Cognitive Science, 29*, 5–39.

Slugoski, B. R., Lalljee, M. G., Lamb, R., & Ginsburg, J. (1993). Attribution in conversational context: Effect of mutual knowledge on explanation-giving. *European Journal of Social Psychology, 23*, 219–238.

Smith, R. H., Hilton, D. J., Kim, S-H., & Garonzik, R. (1992). Knowledge-based causal inference: Norms and the use of consensus information. *British Journal of Social Psychology, 31*, 239–248.

Snyder, M. L., & Frankel, A. (1976). Observer bias: A stringent test of behavior engulfing the field. *Journal of Personality and Social Psychology, 34*, 857–864.

Smith, A. (2002). *The theory of moral sentiments* (K. Haakonssen, Ed.). Cambridge, UK: Cambridge University Press. (Original work published 1789)

Spellman B. (1997). Crediting causality. *Journal of Experimental Psychology: General, 126*, 323–348.

Taylor, S., & Brown, J. (1988). Illusion and well being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193–210.

Tetlock, P. E. (2000). Cognitive biases and organizational correctives: Do both disease and cure depend on the politics of the beholder? *Administrative Science Quarterly, 45*, 293–326.

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians and prosecutors. *Psychological Review, 109*, 451–71.

Tetlock, P. E., & Lebow, R. N. (2001). Poking counterfactual holes in covering laws: Cognitive styles and historical reasoning. *American Political Science Review, 95*, 829–43.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees under-

stand psychological states: The question is which ones and to what extent. *Trends in Cognitive Science, 7*, 153–156.

Tomasello, M., Carpenter, M., Call, J., Behne, J., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–691.

Trabasso, T., & Sperry, L. L. (1985). The causal basis for deciding importance of story events. *Journal of Memory and Language, 24*, 595–611.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and story comprehension. *Journal of Memory and Language, 24*, 612–630.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93*, 239–257.

Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility in dispositional judgment processes. *Journal of Personality and Social Psychology, 73*, 662–674.

Trope, Y., & Gaunt, R. (1999). A dual-process model of overconfident attributional inferences. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 161–178). New York: Guilford Press.

Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology, 79*, 344–354.

Trope, Y., & Gaunt, R. (2003). Attribution and person perception. In M. A. Hogg & J. Cooper (Eds.), *The Sage handbook of social psychology* (pp. 190–208). London: Sage.

Turnbull, W. M. (1986). Everyday explanation: The pragmatics of puzzle resolution. *Journal for the Theory of Social Behaviour, 16*, 141–160.

Turnbull, W. M., & Slugoski, B. R. (1988). Conversational and linguistic processes in causal attribution. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 66–93). New York: New York University Press.

van Overwalle, F. (1997). A test of the joint model of causal attribution. *European Journal of Social Psychology, 27*, 221–236.

van Overwalle, F. (2003). Acquisition of dispositional attributions. Effects of sample size and covariation. *European Journal of Social Psychology, 33*, 525–533.

van Overwalle, F., & Timmermans, B. (2005). Discounting and the role of the relation between causes. *European Journal of Social Psychology, 35*, 199–224.

van Overwalle, F., & van Rooy, D. (2001a). How one cause discounts or augments another: A connectionist account of causal competition. *Personality and Social Psychology Bulletin, 27*, 1613–1625.

van Overwalle, F., & van Rooy, D. (2001b) . When more observations are better than less: A connectionist account of the acquisition of causal strength. *European Journal of Social Psychology, 31*, 155–175.

Webster, D. M. (1993). Motivated augmentation and reduction of the overattribution bias. *Journal of Personality and Social Psychology, 65*, 261–271.

Weiner, B. (1985). "Spontaneous" causal thinking. *Psychological Bulletin, 109*, 74–84.

Wilensky, R. W. (1981). Meta-planning: Representing and using knowledge about planning in problem-solving and natural language understanding. *Cognitive Science, 5*, 192–234.

Wilensky, R. W. (1983). *Planning and understanding: A computational approach to human reasoning*. Reading, MA: Addison-Wesley.

Wimer, S., & Kelley, H. H. (1982). An investigation of the dimensions of causal attribution. *Journal of Personality and Social Psychology, 43*, 1142–1162.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs. Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Woolfe, T., Want, S. C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development, 73*, 768–778.

Zauberbühler, K. (2000a). Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition, 76*, 195–207.

Zauberbühler, K. (2000b). Causal knowledge of wild predators' behavior in Diana monkeys. *Animal Behaviour, 59*, 209–220.

The Role of Metacognition in Social Judgment

RICHARD E. PETTY
PABLO BRIÑOL
ZAKARY L. TORMALA
DUANE T. WEGENER

The capacity for complex thought is one of the traits that distinguish us as humans. Although social psychologists have been concerned with thinking and thought processes from our earliest history (e.g., Allport, 1924; James, 1890), the contemporary social cognition movement catapulted the importance of thought processes to center stage within the field (e.g., Fiske & Taylor, 1991; Moscowitz, 2004). This thinking is sometimes very effortful and deliberate, and at other times it is more simple and automatic (Chaiken, Liberman, & Eagly, 1989; Kruglanski & Thompson, 1999; Petty & Cacioppo, 1986). Furthermore, our thinking is on occasion quite objective but at other times contaminated by various biases (see Eagly & Chaiken, 1993; Petty & Wegener, 1998, for reviews).

Although many distinctions such as those just noted can be made about thoughts, in this chapter we are concerned with a distinction between primary and secondary cognition. Primary thoughts involve our initial associations of some object with some attribute, or a projection of some object on some dimension of judgment such as "the flower is red" or "I like the flower" (McGuire & McGuire, 1991). Our thoughts can be directed at any object including other people, the environment, and ourselves. Following a primary thought, people can also generate other thoughts that occur at a second level, which involve reflections on the first-level thoughts (e.g., "Is that flower really red or pink?" and "I am not sure how much I like that flower"). *Metacognition* refers to these second-order thoughts, or our thoughts about our thoughts or thought processes. Some authors have con-

ceived metacognition more broadly as people's knowledge about their own and *others*' mental states, processes, and beliefs (Jost, Kruglanski, & Nelson, 1998; Wright, 2001). In this chapter, however, we focus mostly on research examining the role of thinking about one's own thoughts and thought processes because that is the domain in which most of the social metacognitive research has been conducted.

The topic of metacognition has received considerable theoretical and empirical attention in the past decade, being considered one of the top 100 topics of psychological research (Nelson 1992). Within the field of cognitive psychology, the study of metacognition has traditionally focused on how people monitor and control their own mental functioning, especially in the domain of memory (e.g., Koriat & Goldsmith, 1996; Nelson & Narens, 1994). For example, research has found that the stronger one's feeling of knowing about a piece of information, the more time one is willing to spend searching for it (e.g., Costermans, Lories, & Ansay, 1992). The motivation to complete the search is particularly intense when one has the subjective experience that the information is on the tip of the tongue and thus about to emerge into consciousness (see Yzerbyt, Lories, & Dardenne, 1998a). Thus, metacognitive processes are consequential in guiding further thinking and action.

In addition to cognitive psychology, where the formal study of metacognition emerged, the idea that people think about their thoughts and thought processes is prevalent in a number of other psychological domains. For instance evaluating one's thoughts is critical to some forms

of clinical practice. Indeed, the main goal of cognitive-behavioral therapy is to get individuals to further think about their maladaptive thoughts with the goal of inducing doubt in them (e.g., Beck & Greenberg, 1994; Ellis, 1962). Metacognition also plays a prominent role in the context of consumer psychology (Alba & Hutchinson, 2000; Wright, 2001). For instance, Friestad and Wright (1994) have noted that people's naive theories of attitude change play an important role in determining how individuals deal with persuasion attempts (see also Campbell & Kirmani, 2000; Friestad & Wright, 1995).

Before beginning our review of social psychology's research on metacognition, it is useful to discuss the dimensions along which metacognitive thought might be organized. In particular, we note that the dimensions of metacognitive thought can be organized along many of the same dimensions that have proven useful for understanding primary thoughts, as well as some unique dimensions. Within social psychology, thought coding is particularly prominent in research on attitude change. In this literature, thoughts generated in response to a persuasive message are typically classified into the following categories by judges or the participants themselves: target (is the thought about the message content or the source, etc.), origin (e.g., does the thought stem from the message content or is it uniquely generated by the message recipient), valence (e.g., is the thought favorable or unfavorable toward the proposal), and number (e.g., are there many or few thoughts—see, e.g., Cacioppo, Harkins, & Petty, 1981; Greenwald, 1968; Wright, 1973). Coding thoughts for target, origin, valence, and number has provided a very fruitful approach for understanding some of the psychological processes that underlie attitude formation and change (see Eagly & Chaiken, 1993; Petty, Ostrom, & Brock, 1981).[1] Interestingly, when participants are asked to judge their own thoughts in persuasion research, they are in essence being asked to engage in metacognition because they are asked for their thoughts about their thoughts (e.g., how favorable toward the issue is your thought?). There is no presumption in the literature on primary cognition, however, that people necessarily think about their thoughts in this way on their own—only that these post hoc categorizations are useful for predicting what attitudes people will adopt. Because these categories have already proven effective for classifying primary thoughts, we postulate that they can also serve to understand thoughts at the metacognitive level of thinking.

Thus, in this chapter we use the same categories to describe metacognitive thought. In primary cognition, the dimensions often refer to some objective reality that judges or the participants themselves are asked to determine. For example, with respect to the target of the thought, a judge or the participant might be asked if the thought is about the message itself or about the source of the communication (e.g., Chaiken, 1980). With respect to origin, a judge could determine whether the thought reflected original thinking or merely restated the message content (e.g., Greenwald, 1968). With respect to valence, a judge or the participant might be asked if the thought reflects some positive or negative reaction to the

message (e.g., Petty, Wells, & Brock, 1976). Perhaps attesting to the rather objective nature of these assessments, judges' and participants' ratings often correlate quite well (see Cacioppo et al., 1981). As we describe further below, with metacognitive thought, the focus is on the individual's *perception* of his or her own thought, regardless of its relation to any objective reality. In metacognition, the issue is whether the person spontaneously thinks about the target, origin, valence, quantity, or some other dimension of his or her thoughts and whether these second-order thoughts are consequential.

In this chapter, we refer back to these various dimensions of thought as needed in reviewing the empirical literature. The main function of our organizing structure is to facilitate and highlight the similarities between studies coming from diverse domains in social psychology (e.g., research on the self vs. attitude change). Our review of the literature is organized into four substantive content areas. Within each area we focus on the types of thoughts people have about their primary thoughts and the consequences of these thoughts. We also address, when relevant, thoughts people have about their thought processes and issues for future research. Before reviewing the core areas in which metacognition has been applied, we elaborate the dimensions of metacognition that are most studied in the literature.

TYPES OF THOUGHTS ABOUT THOUGHTS

Briefly described, the *target* dimension of metacognition refers not to what the thought is actually about but what the person perceives it to be about (cf. Higgins, 1998a). For example, a person might wonder: Is the thought about my friend? Is the thought about me? Among other things, this type of metacognition can help individuals to classify their thoughts into categories as a first step in marking them for further control and change (e.g., Ellis, 1962). The *origin* of a thought refers to its source. Where did the thought come from? That is, a person can wonder if the thought is his or her own, or if it merely reflects the statements or sentiments of others (e.g., Greenwald, 1968). People are more likely to act on thoughts that are perceived as connected to or originating from the self (see Wheeler, DeMarree, & Petty, 2005). Perhaps the most commonly studied aspect of thought is its *valence*. That is, regardless of the target of the thought or its origin, does the person perceive it to reflect something positive or negative with respect to its target? Finally, *amount* refers not to how many thoughts are actually generated but to the perceived quantity of thoughts. For example, people might think that they possess very few or many thoughts about a given topic, and as we review shortly, such thought attributions have important implications for social judgment and behavior (e.g., Schwarz, Bless, et al., 1991).

In addition to these aspects of thought, two additional dimensions are uniquely metacognitive and have achieved the most conceptual and empirical attention. First, one can assess one's *evaluation* of a thought. That is, regardless of the perceived target, origin, valence, or

number of thoughts generated, people can assess their thoughts as good or bad, desirable or undesirable, appropriate and wanted, or not. When thoughts are unwanted, people might try to suppress them (Wegner, 1994). Or, if the thoughts are seen as inappropriate or bad, people might try to correct for their anticipated impact on judgments or action (Wegener & Petty, 1997).

Second, people can have varying degrees of *confidence* in their thoughts, ranging from extreme certainty to a high amount of doubt.[2] Thus, two people might have the *same* thought with respect to a given proposal or target, but one person might have considerably more confidence in that thought than the other. Thoughts held with more confidence have a larger impact on judgments (e.g., Petty, Briñol, & Tormala, 2002). Given the attention that research in social psychology has dedicated to the study of confidence, in the next section we describe this dimension in more detail.

Before turning to confidence, however, it is worth noting that the different categories of metacognitive thoughts likely relate to each other. For example, a thought whose origin is perceived as the self might be evaluated more favorably than one with a perceived external origin (e.g., Greenwald & Albert, 1968). Or, thoughts about targets for which we are perceived to have more knowledge than other people might be held with more confidence (Kruglanski et al., 2005). Furthermore, it is possible to consider third-order cognition in which people are asked to make one metacognitive judgment about another. Thus, after asking about the evaluation of a particular thought, one can ask about the confidence in that evaluation. Or, after asking about the confidence in one's thought, one can ask whether this confidence is appropriate, and so forth. To our knowledge, no research to date has systematically examined third-order cognition.

METACOGNITIVE CONFIDENCE

A subjective experience that constitutes one of life's greatest metacognitive challenges is the sense of epistemic certainty or uncertainty (e.g., Kruglanski 1980, 1989; Nelson, Kruglanski, & Jost, 1998). Although certainty could presumably apply to various aspects of one's judgment (e.g., am I certain that this is *my* belief?), social psychological theory has paid special attention to the perceived *validity* of one's judgments. For example, in lay epistemic theory (Kruglanski, 1980), following the generation of a thought or hypothesis, people are said to want to test and validate the thought. Within dual-process models of persuasion such as the elaboration likelihood model (Petty & Cacioppo, 1986; Petty & Wegener, 1999) and the heuristic–systematic model (Chaiken, Liberman, & Eagly, 1989), people are postulated to seek a certain level of validity, confidence, or certainty in their judgments. When the gap between a person's current level of judgmental confidence and the desired level is high, they should engage in greater information-processing activity (see Chaiken et al., 1989). Underlying these notions is the assumption that people generally seek to hold accurate judgments (Festinger, 1954), that uncertainty is aversive, and that deliberative information processing is often a good way of obtaining accuracy and reducing uncertainty (cf. Sorrentino, Bobocel, Gitta, Olson, & Hewitt, 1988). In fact, inducing a general feeling of doubt, whether explicit (Tiedens & Linton, 2001) or implicit (Petty, Tormala, Brinol, & Jarvis, 2006), has provoked greater information processing. Indeed, incongruencies of all sorts have increased information-processing activity (e.g., Baker & Petty, 1994; Maheswaran & Chaiken, 1991; Smith & Petty, 1995; Ziegler, Diehl, & Ruther, 2002), presumably by inducing some doubt or confusion that might be resolved with thinking.

Thus, metacognitive confidence generally refers to a sense of validity regarding one's thoughts or judgments, though it is possible for certainty to be applied to other aspects of attitude-relevant beliefs (e.g., certainty that the attitude is my own). Studying metacognitive confidence is important mainly because confidence affects whether people translate their individual thoughts into more general judgments or evaluations, and whether these judgments in turn are influential in guiding behavior. There is great deal of empirical evidence suggesting that beliefs held with great conviction are a more potent foundation for judgment and behavior than more tentatively held beliefs (e.g., Berger & Mitchell, 1989; Briñol & Petty, 2004; Fazio & Zanna, 1978; Pieters & Verplanken, 1995; Swann & Ely, 1984; Tormala & Petty, 2002).

Logically, one might expect that highly confident thoughts, beliefs, and attitudes determine actions because confident beliefs stem from a truly accurate perception of reality. That is, common sense would suggest that confidence should emerge from objective sources such as the amount and quality of information people have. For example, the more time a person has to see something, the more confident the person should be in his or her identification of it. However, an objective basis to confidence seems to be the exception rather than the rule. For example, when people were asked to predict the behavior or the personality of patients, prosecutors, dating partners, roommates, or strangers, there was no relation or only a modest correlation between confidence and predictive accuracy (e.g., Deffenbacher, 1984; Dunning, Griffin, Milojkovic, & Ross, 1990; Jacoby, Bjork, & Kelley, 1994; Oskamp, 1965; Swann & Gill, 1997; Wells & Murray, 1983). Numerous studies have shown that the confidence with which people hold their beliefs can be affected by different factors that do not necessarily increase the validity of beliefs. For example, the more frequently people think about their thoughts, the more confidence they have in them (e.g., Koriat, Lichtensetein, & Fischhoff, 1980), the more a judgment is repeated, the more confidence is increased (e.g., Shaw, 1996), the more details that are included in a given thought, the more confidence is increased (e.g., Gill, Swann, & Silvera, 1998). In addition to these factors, confidence has been found to be affected by several other situational (e.g., Yzerbyt, Schadron, Leyens, & Rocher, 1994) and personality (Schaefer, Williams, Goodie, & Campbell, 2004) variables.

Research on decision making has also provided a number of important examples of how peoples' judgments of confidence can be independent of their accuracy. According to Griffin and Tversky (1992), for instance, confidence judgments require the integration of different kinds of evidence such as the extremity of the available information and the weight or predictive validity of that information. For example, when evaluating a letter of recommendation for a graduate student written by a former teacher, a person may consider how positive the letter is (i.e., extremity) and how credible or knowledgeable the writer is (i.e., weight). According to Tversky and Kahneman (1974), people tend to focus mostly on the extremity of the evidence leading them to underutilize other variables that control predictive validity. As a result, Griffin and Tversky argue that overconfidence occurs when extremity is high and weight is low, whereas underconfidence takes place when extremity is low and weight is high.

In this research, as in most of the literature on decision making, confidence judgments refer exclusively to the estimation of how likely it is for an answer (e.g., a judgment or a decision) to be correct, and criteria of accuracy are typically available. In social psychological research it is much less common to use objective criteria of accuracy as people's thoughts often relate to judgments or actions involving people, groups, political views, and so forth. Although one can determine whether one's confidence in the likelihood of a red ball coming out of an urn is well placed, it is not possible to determine whether one's confidence in an attitude toward a presidential candidate is accurate in any objective sense. Yet, a subjective sense of accuracy or confidence has important implications such as determining if one will vote for the candidate based on one's attitude. Within the social psychological literature, judgmental confidence has been tied to the extremity of the attributes a target possesses and the certainty that the target does or does not possess those attributes (Griffin & Tversky, 1992; Petty et al., 2002), one's confidence in the likelihood that a target possesses a certain attribute, and one's confidence in the desirability of that attribute (Briñol, Petty, & Tormala, 2004), as well as other sources such as the ease with which information comes to mind (Haddock, Rothman, Reber, & Schwarz, 1999).

Although metacognitive ideas have been applied to various topics in social judgment, in the next sections we focus on some of the most heavily researched areas of social psychology in which the role of metacognition has been examined. These are memory and cognitive fluency, attitudes and persuasion, the self and individual differences, and bias correction. In the final section, we outline several key conclusions and general principles of the reviewed work.

MEMORY AND COGNITIVE FLUENCY

Metacognition is deeply rooted in the study of human memory—in particular, people's theories about and perceptions of their memory (e.g., Costermans et al., 1992; Strack & Förster, 1998), their knowledge (e.g., Koriat, 1993), and their learning (e.g., Dunlosky & Nelson, 1994). Jacoby and colleagues, for instance, proposed that memory often involves not a literal search for stored information, but rather a series of inferences based on "cognitive feelings" (e.g., Jacoby, Kelley, & Dywan, 1989; see also Clore & Parrot, 1994). One such feeling is familiarity, which people interpret as indicating that something is known, or remembered (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989). Another such feeling involves the ease or fluency with which information can be retrieved from memory (e.g., Benjamin & Bjork, 1996; Koriat & Levy-Sadot, 1999; Nelson & Narens, 1990). It is this feeling of fluency that has received the most attention in social metacognition research. Next, we provide an overview of classic and contemporary findings on fluency. In so doing, we also review the current state of knowledge with respect to two very influential effects in social psychological research: ease of retrieval effects and mere exposure effects.

Ease of Retrieval

We begin with the notion that people sometimes base judgments on the subjective experience of fluency with which information comes to mind. In their seminal research on the availability heuristic, Tversky and Kahneman (1973) found that people sometimes perceive events as more likely or common if examples of them come to mind easily. In their now classic work on ease of retrieval, Schwarz, Bless, and colleagues (1991) expanded on the earlier research by directly pitting the subjective ease of information retrieval against the actual content of the information retrieved. Schwarz, Bless, and colleagues asked participants to rate their assertiveness after recalling 6 versus 12 examples of their own assertive behavior. They found that people viewed themselves as more assertive after retrieving 6 rather than 12 examples. The logic behind this effect was that people based their judgments of assertiveness on the subjective experience of the ease with which assertive behaviors could be retrieved from memory. When it was easy, because only a few behaviors were requested, people concluded that they must be pretty assertive. When it was difficult, because many behaviors were requested, people inferred that they must not be very assertive. What made this finding so intriguing was that it demonstrated that in at least some contexts people forsake the content of accessible information in memory (e.g., having many assertive behaviors activated) and instead base judgments on the subjective experience of memory (e.g., ease or difficulty). Of course, these effects depend on the perceived diagnosticity of the feeling of ease or difficulty (see Schwarz, 1998, for a review). In the Schwarz, Bless, and colleagues studies, for example, ease of retrieval had no impact on judgments of assertiveness when ease was attributed to an external source.

Since the Schwarz, Bless, and colleagues (1991) research, the ease of retrieval notion has been applied to a number of domains. In the attitudes domain, for instance, it had been found that the easier it feels for people to generate positive thoughts about an object or issue, the more people like that object or issue (Haddock,

Rothman, & Schwarz, 1996; Tormala, Petty, & Briñol, 2002; Wänke, Bless, & Biller, 1996; Wänke, Bohner, & Jurkowitsch, 1997). In addition to attitudes, ease of retrieval can influence a variety of other judgments, such as likelihood estimates (e.g., Hirt, Kardes, & Markman, 2004; Wänke, Schwarz, & Bless, 1995), risk assessments (Grayson & Schwarz, 1999; Rothman & Schwarz, 1998), stereotypes (Dijksterhuis, Macrae, & Haddock, 1999; see also Rothman & Hardin, 1997), attitude certainty (Haddock et al., 1999), judgments of interpersonal closeness (Broemer, 2001), and feelings of self-doubt (Hermann, Leonardelli, & Arkin, 2002). Recent research has shown that ease of retrieval can affect not only explicit but also implicit measures of attitudes (Gawronski & Bodenhausen, in press). Behavior can also be affected. For example, Keller and Bless (2005) found that when people were asked to recall negative stereotypes about their own group, their performance on an ability test was worse when they recalled few (easy) rather than many (difficult) pieces of information.

Although the basic ease effect appears to be quite robust, there have been divergent findings with respect to moderating and mediating factors. For instance, are ease effects most likely to occur under high or low thinking conditions? Researchers originally assumed that ease effects are heuristic in nature and, thus, most likely to operate when thinking is low (see Schwarz, 1998). Some evidence has been produced that is consistent with this notion (Grayson & Schwarz, 1999; Rothman & Schwarz, 1998; Ruder & Bless, 2003). Other research, however, has pointed to the exact opposite conclusion—that ease effects are more likely to operate under high thinking conditions, when people have the motivation and ability to attend to and interpret their own cognitive experience (Hirt et al., 2004; Tormala et al., 2002; Wänke & Bless, 2000).

This controversy may stem in part from different perspectives on the mechanism responsible for ease of retrieval effects. The argument for low thought conditions is based on the notion that ease effects stem from availability inferences (Schwarz, 1998). Difficulty in generating favorable arguments for a tax cut, for example, indicates that few favorable arguments exist, so the tax cut is not worth supporting. Effects might be viewed as emerging mainly when thinking is low because that is when heuristics are most likely to impact judgments (Chaiken, 1987). Other research (Tormala et al., 2002), however, suggests that ease effects can be mediated by feelings of confidence or validity associated with the particular arguments or thoughts retrieved. The easier it is to generate a list of arguments supporting a tax cut, the more confident people are that those arguments are valid (see also Wänke & Bless, 2000). People have been found to be particularly attuned to thought confidence and validity under high thinking conditions (Petty et al., 2002). Although research has identified multiple mechanisms by which ease effects can occur, each begins with the assumption that people perceive their ease or difficulty in thinking—a metacognition. What differs in the accounts is what inferences people make based on this perceived ease. It seems likely that the different mechanisms each contribute to ease effects in different situations.

Mere Exposure

A well-known phenomenon that predates ease of retrieval research is that of mere exposure. The mere exposure effect occurs when attitudes toward stimuli become more favorable as a consequence of repeated exposure to those stimuli (Zajonc, 1968). In one early demonstration of this phenomenon, Kunst-Wilson and Zajonc (1980) presented people with a series of polygon images and found that even when these images could not be consciously recognized, the more often they had been presented, the more they were liked. This effect has now been demonstrated with a wide variety of stimuli such as foreign words, photographs, music, ideographs, and nonsense syllables (see Bornstein, 1989, for a review). Moreover, it has been shown that mere exposure can affect mood, and that this mood can spread to other, related stimuli that were not even presented (Monahan, Murphy, & Zajonc, 2000). To account for mere exposure effects, two general explanations have been proposed: perceptual fluency and perceived familiarity.

A great deal of research suggests that previous or repeated exposure to stimuli can make those stimuli easier to process, and that this *perceptual fluency* enhances subsequent liking. Specifically, the feeling of perceptual fluency, or ease of processing, is thought to be misattributed to a positive evaluation of the stimulus (Bornstein, 1989; Bornstein & D'Agostino, 1994; Jacoby, Kelley, Brown, & Jasechko, 1989). Of importance, though, perceptual fluency can also be attributed to other stimulus dimensions (Mandler, Nakamura, & Van Zandt, 1987). For instance, previously presented names seem more famous (Jacoby, Woloshyn, & Kelley, 1989) and previously presented statements seem more true (Begg, Armour, & Kerr, 1985), even when those statements are explicitly identified as false (Skurnik, Yoon, Park, & Schwarz, 2005). When stimuli already have some meaning, or tend to elicit a dominant response in one direction or another, repeated exposure can accentuate that dominant response (Brickman, Redfield, Harrison, & Crandall, 1972). Repeatedly presenting negative information, for instance, can make that information seem more negative (Cacioppo & Petty, 1989; Grush, 1976). One possible reason for these polarization effects is that one's positive assessments of positive information seem more valid or plausible as exposure increases, as do one's negative assessments of negative information (Kruglanski, Freund, & Bar-Tal, 1996).[3]

As an alternative to the perceptual fluency account, some research suggests that mere exposure effects might at least partially derive from feelings of *familiarity* (see Bornstein & D'Agostino, 1994; Lee, 2001). In fact, Lee (1994, 2001) found that people generally prefer old (familiar) to new (unfamiliar) stimuli, even when the old stimuli have not been *repeatedly* presented. Furthermore, some research suggests that the feeling of familiarity enhances liking even when the familiarity does not stem from any prior exposure at all. Moreland and Zajonc (1982) found that people responded more favorably to faces when those faces felt familiar because they were similar to other ones that had been viewed. The associa-

tion between familiarity and liking is so strong that stimuli that are already positive are also likely to be perceived as familiar (Corneille, Monin, & Pleyers, 2005; Garcia-Marques, Mackie, Claypool, & Garcia-Marques, 2004; Monin, 2003). As with perceptual fluency, however, some studies have qualified this view, suggesting that the feeling of familiarity stemming from repeated exposure can foster liking or disliking, depending on other contextual factors (e.g., Klinger & Greenwald, 1994; for a similar finding, see Smith et al., 2006).

Although the perceptual fluency and familiarity explanation of mere exposure effects have been pitted against one another in the literature, these mechanisms might ultimately be somewhat intertwined. For example, Whittlesea, Jacoby, and Girard (1990) found that stimuli presented with greater visual clarity were perceived as being more familiar, and more likely to have appeared previously, than stimuli presented with less visual clarity. Furthermore, perceptual fluency stemming from repeated exposure can make a stimulus feel more familiar and enhance liking in this manner (e.g., Jacoby & Kelley, 1987; Jacoby, Kelley, Brown, & Jasechko, 1989).

Other Sources of Fluency

In addition to ease of retrieval and mere exposure, there are a number of other sources of processing fluency. The classic perceptual fluency effect—that easy-to-process stimuli are evaluated more favorably than difficult-to-process stimuli—has been produced using a diverse set of experimental manipulations. Line drawings, for example, tend to be liked more and produce more positive affect when they have greater rather than less figure–ground contrast, when they have been presented for a longer rather than shorter amount of time, and when they have been preceded by a similar rather than dissimilar prime (Reber, Winkielman, & Schwarz, 1998; see also Winkielman & Cacioppo, 2001). In addition, words tend to be easier to process and judged as more pleasant when they are embedded in a predictive rather than nonpredictive semantic context (e.g., Whittlesea, 1993).

Other Effects of Fluency

Confidence

In addition to affecting people's evaluations of stimuli, as noted earlier, processing fluency can also influence feelings of confidence. For example, recall that the ease of generating thoughts can affect the confidence with which those thoughts are held (Tormala et al., 2002). Other forms of processing fluency have also been shown to affect confidence (see Koriat & Levy-Sadot, 1999, for a review). For example, Gill and colleagues (1998) used a priming procedure to make certain kinds of impressions of a person easier to form following a paragraph about that person. Gill and colleagues found that when judgments were made more easily, because of priming, those judgments were also held with greater confidence. Norwick and Epley (2003) found that participants were more confident that a given statement was true when the statement was easy rather than difficult to read. Busey, Tunnicliff, Loftus, and Loftus (1995) presented participants with a series of faces to study for a later recognition test. When participants were tested under bright rather than dim viewing conditions, which made the faces easier to see, participants were more confident in their recognition judgments. Borrowing from the mere exposure paradigm, researchers have also found various forms of repetition to increase confidence. Repeating questions, for instance, increases confidence in the answer retrieved (e.g., Hastie, Landsman, & Loftus, 1978). Similarly, repeated expression of one's attitude increases attitude certainty, and this effect can stem from the objective ease, or accessibility, with which the attitude comes to mind (Holland, Verplanken, & van Knippenberg, 2003).

Perceived Knowledge

Another metacognition associated with cognitive fluency involves people's perceptions of their own knowledge (see Koriat & Levy-Sadot, 1999, for a review). As but one example, Werth and Strack (2003) presented people with trivia-type questions and answers in an easy-to-read or difficult-to-read color scheme. When the colors made the question and answer easy rather than difficult to read, participants were more likely to assume they knew the answer all along. This *feeling of knowing*, it turns out, can have important implications for thought and behavior. Under some conditions, the feeling of knowing has a positive impact on information processing. The greater one's feeling of knowing an elusive item in memory, the more time one will spend searching for that item before giving up (e.g., Costermans et al., 1992; Koriat, 1993; Nelson & Narens, 1990; Yzerbyt, Lories, & Dardenne, 1998). Under other conditions, though, perceived knowledge has a negative impact on information processing. In particular, the more one thinks one knows about a topic, the less likely one is to seek new information on that topic (Radecki & Jaccard, 1995). These effects are particularly intriguing given that the correlation between perceived and actual knowledge tends to be quite low (e.g., Glenberg, Wilkinson, & Epstein, 1982; Krosnick, Boninger, Chuang, Berent, & Carnot, 1993; Radecki & Jaccard, 1995).

Caveat

As the preceding review indicates, much of the research on processing fluency suggests that the experience of ease is positive. In fact, some recent work explicitly concludes that fluency is by its very nature a pleasurable experience (for a review, see Reber, Schwarz, & Winkielman, 2004; Schwarz, 2004). Based on most of the research we have reviewed, this conclusion makes sense. Indeed, easy-to-process stimuli tend to be rated more favorably than difficult-to-process stimuli (e.g., Bornstein, 1989; Jacoby, 1983; Lee & Aaker, 2004). Easy-to-process pictures have been shown to elicit greater positive affect than do hard-to-process pictures (Winkielman & Cacioppo, 2001). Easy-to-generate thoughts or argu-

ments, whether positive or negative, tend to be viewed as more valid (Tormala et al., 2002), frequent (Schwarz, 1998), or generally diagnostic of what one thinks. In short, processing fluency has typically been viewed as inherently positive in some way.

However, recent research has determined that the meaning of ease is malleable (e.g., Winkielman & Schwarz, 2001), and changing the meaning of ease can modify the normal ease of retrieval effect. In one study, for instance, Briñol, Petty, and Tormala (2006) reported that under some conditions people can be induced to view easy-to-generate thoughts as *less* rather than more diagnostic of what they think. Briñol and colleagues manipulated the meaning of ease by leading some participants to believe that cognitive ease was a positive sign of mental functioning whereas difficulty was a negative sign. Other participants received the exact opposite information. When ease was described as positive, Briñol and colleagues replicated the usual ease effect—participants were more favorable toward an issue after they generated an easy rather than difficult number of arguments in favor of it. When ease was described as negative, however, this effect was reversed—participants were less reliant on their thoughts when those thoughts were easy rather than difficult to generate. In other words, although there may be a default tendency to think easy-to-generate thoughts are more plentiful, valid or desirable, this tendency can be changed (see also Freitas, Azizan, Travers, & Berry, 2005; Schwarz, 2004). The implications of this effect for other fluency effects have yet to be fully explored, but this initial finding suggests that the impact of the cognitive experience of ease on social judgment may ultimately prove more complex than is typically presumed.

ATTITUDES AND PERSUASION

One of the first areas in social psychology to be concerned with metacognition was that of attitudes and persuasion. In particular, researchers have been interested in various metacognitive properties of attitudes as indicators of the attitude's strength, though the early work on this topic did not link explicitly to the "metacognitive" moniker. *Attitude strength* refers to the extent to which an attitude persists over time, guides behaviors and other judgments, and is resistant to influence attempts (Krosnick & Petty, 1995). Wegener, Downing, Krosnick, and Petty (1995) noted that for many aspects of attitude strength, there were both objective and subjective indicators. That is, for virtually every postulated objective indicator of an attitude's strength such as the actual speed with which an attitude comes to mind (accessibility; see Fazio, 1995), or the amount of information people can generate regarding their attitudes (knowledge; see Wood, Rhodes, & Bick, 1995), or the amount of thinking they have done about their attitude (Petty, Haugtvedt, & Smith, 1995), there is a parallel measure of the *perceived* ease of attitude access or amount of knowledge or thought. However, there are some subjective perceptions, such as attitude certainty and importance, for which there are no objective counterparts.

In an influential paper, Bassili (1996) referred to the subjective perceptions regarding one's attitudes as *meta-attitudinal* indicators and contrasted them with the more *operative* or objective indicators that tapped more directly into attitude structure or process. Notably, Bassili was rather critical of meta-attitudinal features of attitudes, arguing that they are typically not represented with the attitude object or closely related to the factors that determine attitudes. He argued that because people are often not aware of the processes leading to their judgments (Nisbett & Wilson, 1977), reports of how much knowledge was behind one's attitude or how quickly one's attitude would come to mind would not relate very well to operative measures of these constructs. Furthermore, he argued that "second order judgment (i.e., judgment about judgment) requires a level of intrapsychic awareness that people seldom have" (p. 640; see also Roese & Olson, 1994).

After conducting two studies comparing meta to operative indicators of strength in their abilities to predict attitude stability and resistance to change, he concluded that "operative measures have more predictive validity than meta-attitudinal measures" (p. 651). Although this conclusion is consistent with the data he reports for indices collapsed across various meta and operative measures, it is important to note that his data also show that one metameasure in particular—attitude certainty—either did as well as or outperformed the operative measures in predicting attitude consequences.

Thus, our point of view on the utility of metacognitive measures of attitude strength is a bit different. Consistent with Bassili (1996), we think it is unlikely that there are a large number of subjective assessments that are directly linked to attitudes. That is, it may not be the case that perceptions of accessibility, knowledge, and so forth are linked to most attitudes. In contrast to Bassili, however, we argue that many attitudes at least have a validity, certainty, or confidence tag (Petty, 2006; Petty & Briñol, 2006; Petty et al., 2006). Indeed, considerable research is consistent with the idea that when people come to disbelieve or have some doubt in an idea (i.e., hold the idea with less than complete certainty), this idea can be marked with a negation or doubt tag (Clark & Chase, 1972; Gilbert, Tafarodi, & Malone, 1993; Mayo, Schul, & Burnstein, 2004).

We review research in four topic areas. First, we examine research that assesses metacognitive strength features of attitudes and how these features affect various consequences. Second, we discuss research that examines how metaperceptions of an attitude might be affected by accepting or rejecting a persuasive message. Next, we discuss research in which metaperceptions of one's attitude-relevant thoughts determine persuasion. Finally, we briefly mention work regarding metacognition about the processes of persuasion.

Metacognitive Properties of Attitudes: Certainty and Others

As noted earlier, there are many potential thoughts people could have about their attitudes such as how quickly

they come to mind, how many others share their view, and so forth (Wegener et al., 1995). Nevertheless, several perceptions of attitudes have received the most attention and we discuss these next.

Attitude Certainty

The most studied metacognitive aspect of attitudes and the one of most long-standing interest (e.g., Allport, 1924) is the certainty or confidence with which an attitude is held. Certainty generally refers to a sense of validity concerning one's attitudes (Gross, Holtz, & Miller, 1995). Although certainty naturally covaries with extremity (i.e., people tend to feel more certain as their attitudes deviate from neutrality; e.g., Raden, 1989), certainty and extremity are conceptually distinct such that a person can have high certainty in the validity of a neutral attitude, or express an extreme attitude with low confidence. A number of determinants of attitude certainty have been examined. People tend to be more certain of their attitudes when they are based on direct experience (e.g., Fazio & Zanna, 1981), when they come to mind easily (Haddock et al., 1999), when others agree with the attitude (e.g., McGarty, Turner, Oakes, & Haslam, 1993; Visser & Mirabile, 2004), and when people have done much prior thinking about the attitude object (Abelson, 1988) unless that thinking arouses conflicting thoughts (Liberman & Chaiken, 1991). Some individual differences have also been related to attitude certainty (e.g., uncertainty orientation–Sorrentino & Short, 1986; dogmatism–Palmer & Kalin, 1991), though it is not clear whether the certainty effects of these individual differences are independent of attitude extremity. Attitude certainty, at least with respect to political issues, tends to increase from young adulthood to middle age but then declines sharply as one becomes elderly (Visser & Krosnick, 1998). Given the declines in memory accessibility as one ages (Varhaeghen & Salthouse, 1997), it would be interesting if declines in attitude accessibility are partially responsible for the declines in certainty in one's later years.

Attitude certainty has been associated with a number of important attitude-relevant outcomes. In particular, attitudes held with greater certainty are more resistant to change (e.g., Tormala & Petty, 2002), persistent in the absence of a persuasive attack (Bassili, 1996), and more predictive of behavior (Fazio & Zanna, 1978) than attitudes about which there is doubt. In fact, attitudes may have to reach a certain level of certainty before action is initiated (Gerard & Orive, 1987). Certain attitudes may be more resistant to change because certainty induces a confirmatory information-seeking style (e.g., Swann & Ely, 1984), and certain people are more likely to assume that others agree with them (Marks & Miller, 1985). Gross and colleagues (1995) suggest that it is useful to distinguish "true confidence" in one's attitude from "compensatory confidence." The former is based on knowledge or social support whereas the latter actually reflects an absence of confidence. We suggest that the latter might be revealed in low confidence on an implicit measure. Gross and colleagues argue that those with compensatory confidence

may be most likely to project their attitudes onto others and respond to counterattitudinal messages with a feeling of threat rather than challenge (see Blascovich, 1992).

Importance, Ambivalence, and Knowledge

Other metacognitive features of attitudes that have been examined with some frequency include attitude importance, ambivalence, and knowledge. Although these constructs are often determined by the same factors as attitude certainty and produce similar effects, the accumulated evidence suggests that there are sufficiently different antecedents and consequences to treat them as distinct (see Visser, Bizer, & Krosnick, 2006, for a comprehensive review). For example, in one study (Visser, Krosnick, & Simmons, 2003), despite an overall positive correlation between importance and certainty on the issue of global warming, it was found that as people were exposed to more media reports on global warming, the importance of global warming increased but certainty in one's attitudes decreased.

Attitude importance has been defined as the extent to which people attach significance to their attitude and care about it (Krosnick, 1988). Because importance is viewed as something attached to the attitude (Boninger, Krosnick, & Berent, 1995), it is appropriately viewed as a metacognitive feature. However, in much research on attitude importance, what is typically measured is the perceived importance of the attitude *object* rather than the attitude itself (e.g., how important is the *issue* of global warming to you personally"; Bizer & Krosnick, 2001, emphasis added; see also Holbrook, Berent, Krosnick, Visser, & Boninger, 2005; Krosnick et al., 1993; Visser et al., 2003).[4] Because of this, the construct as operationalized is closely related to the construct of issue involvement, which concerns the personal importance people attach to particular attitude objects and issues (e.g., Thomsen, Borgida, & Lavine, 1995; Petty & Cacioppo, 1990; Petty, Cacioppo, & Haugtvedt, 1992). Furthermore, when manipulations of attitude importance have been attempted, they too have been closely related (if not identical) to manipulations of issue involvement (i.e., manipulating the personal relevance of the *issue*; e.g., Bizer & Krosnick, 2001; Boninger et al., 1995). To add further confusion, some researchers have measured attitude importance (e.g., "I consider my attitude toward smoking to be important") but have referred to it as *issue* importance (van Harreveld & van der Pligt, 2004).

At the conceptual level, it seems likely that attitude importance and issue importance would often go together (see Boninger et al., 1995). Nevertheless it also seems quite easy to distinguish between attitude and issue importance in some circumstances. For example, a person might think that the issue of capital punishment is very important but care little about what particular attitudinal position is adopted with respect to it. On the other hand, another person might be committed to adopting particular opinions that agree with his or her ingroup but the issues themselves are of little consequence. In an early paper, Zimbardo (1960) distinguished the involvement

people can have with an issue and the involvement with one's particular attitudinal response. The more recent conflation of attitude and issue importance is potentially misleading if the constructs are conceptually distinct and can induce different psychological consequences, at least in some situations. For example, Petty and Cacioppo (1990) argued that the increased importance of any particular attitudinal *issue* would lead to enhanced information seeking and processing with respect to the issue, whereas the increased importance of any particular attitudinal position (i.e., caring about whether you favored or opposed a topic) was more likely to lead to a biased seeking and processing style.

Recognizing a distinction between the importance of the attitude issue (or object) and the particular attitudinal position may help to clarify what appear to be surprising findings in the literature. For example, Bassili (1996) found that attitude importance was positively rather than negatively related to attitude change as would be expected if importance tapped into attitude strength. However, because the measure of attitude importance he used actually assessed the importance of the issue rather than the importance of the person's attitudinal position, this positive relation makes sense if increased importance of the issue led people to pay more attention to messages about the issue and they changed in response to processing these messages (Petty & Cacioppo, 1979). Indeed, rated issue importance has been associated with enhanced information-seeking activities (e.g., Visser et al., 2003).

Similar points might be made about other traditional attitude strength assessments. For instance, although researchers have referred to *attitude ambivalence*, the actual assessment of ambivalence, both objective and subjective, concerns whether or not people actually have (e.g., Kaplan, 1972) or perceive themselves to have thoughts about the attitude *object* or *issue* that are mixed in valence or one sided (e.g., "I do not find myself feeling torn between the two sides of the *issue* of capital punishment"; Newby-Clark, McGregor, & Zanna, 2002, emphasis added; see also Connor, Povey, Sparks, James, & Shepherd, 2003; Priester & Petty, 1996; Tourangeau, Rasinski, Bradburn, & D'Andrade, 1989). Thus, what has been studied in the literature might more properly be called objective and subjective *issue ambivalence* in that the person's thoughts are mixed with respect to the attitude object rather than the attitudinal position per se. Although it has not been studied, it is presumably possible for people to be ambivalent about their attitudinal positions per se, but this would require assessment of whether people had both positive and negative thoughts *about their attitude* rather than about the issue.[5] It seems quite possible for a person to be rather ambivalent about an issue (i.e., there are some good and bad things about capital punishment) but quite *un*ambivalent with respect to the particular attitude position held (i.e., my thoughts about being neutral are all positive; I do not feel ambivalent about having an ambivalent attitude). Similarly, people might have little ambivalence with respect to the attitude object (i.e., all thoughts about oneself are positive) but recognize that there are both good and bad things about being

so favorable (e.g., others might think you are competent but conceited). This might well explain why ambivalence and certainty have been only weakly correlated. That is, ambivalence has been assessed with respect to the attitude object or issue whereas certainty has been about the attitude itself.

Finally, it is worth noting that the same distinction can be made regarding *attitude knowledge*. Again, what has typically been assessed has been perceived knowledge about the issue under consideration (see Davidson, 1995; Wood et al., 1995) rather than the attitude per se. The distinction in this case has to do with whether people are being asked to make a subjective assessment of what they know about an issue (e.g., capital punishment) versus what they know about a particular attitudinal position (e.g., strongly opposing capital punishment). Although these two constructs are likely to share much variance, it is also easy to imagine divergence. For example, people could feel that they have a lot of information about the attitude issue in general but relatively little in support of the particular position they favor. Or, they might think they have much information in support of their position but relatively little information overall because they are completely ignorant of the other side. The more overall information they believe they have (on both sides), the more certain they might be in the validity of the attitude, whereas the more information they believe they have in support of a particular side, the more extreme their attitude might be.

The effects of perceived knowledge have paralleled the effects of amounts of actual knowledge (e.g., see Johnson, 1994; Wood et al., 1995). First, high perceptions of knowledge lead attitudes to be more predictive of behavior (see Davidson, 1995, for a review). In addition, people who believe they are already well informed on an issue are less likely to seek additional information on that topic, especially if they are also high in the need for cognitive closure. But, if those high in perceived knowledge are provided with information, they are more likely to think about it, differentiating strong from weak arguments (see Kruglanski et al., 2005), and are less reliant on issue-irrelevant cues (Ellis, 1996). Kruglanski and colleagues (2005) note that one's perception of one's own knowledge is an important factor to consider in understanding source credibility effects. That is, they argue that it is the gap between one's own perceived knowledge and that of the source that determines the source's impact (see also Ellis & Kruglanski, 1992).

Attitude Certainty and Persuasion

Having reviewed some definitional issues with respect to metacognitive features of attitudes, we now turn to how the most studied metacognitive strength feature, attitude certainty, can be affected by receipt of a persuasive message. The voluminous research on persuasion has naturally focused on whether or not an attitude changes in valence or extremity as a result of message exposure. More recent research, however, has studied how the certainty with which people hold their attitudes can be affected following persuasion or resistance.

Confidence in Old and New Attitudes as a Result of Exposure to Persuasive Messages

First, consider a situation in which people receive a persuasive message and do not change their attitudes in response to it. Traditional analyses assume that when attitudes fail to change, the persuasive message has been ineffective. Yet, Tormala and Petty (2002) argued that when people resist a persuasive message they sometimes contemplate their own resistance and as a result, the confidence with which they hold their initial attitude can change. In particular, in several studies, Tormala and Petty showed that when people believe they have effectively resisted a strong persuasive message, certainty in their original attitudes is increased. When people believe they have resisted a weak message, certainty does not increase because people are uncertain as to whether they would have been able to resist a stronger message.[6]

Subsequent research (Tormala & Petty, 2004a) showed that this metacognitive effect was only present among individuals who were prone to engaging in considerable amounts of thought (i.e., those high in need for cognition; Cacioppo & Petty, 1982) or in situations that fostered high amounts of thinking (e.g., a topic high in personal relevance; Petty & Cacioppo, 1979). Furthermore, just as certainty in one's original attitude can be increased when people believe they have resisted a message with strong arguments, so too can their certainty be increased when they believe they have resisted a message from an expert (Tormala & Petty, 2004b). Resisting a message from a nonexpert is not very diagnostic with respect to one's attitude because one does not know whether the attitude could have survived an attack from someone more knowledgeable.

What if people are trying to counterargue a message but are unsuccessful in doing so? In a series of studies, Rucker and Petty (2004) found that people who process a message with an intent to find fault, but are unsuccessful in counterarguing, become more convinced in the validity of their *new* attitude than people who processed the same message in a more objective manner. Rucker and Petty argued that when the arguments in a message are very strong, both people who are processing objectively and those trying to find fault will realize that there are many positives to the proposal. Only people who are trying to find fault and fail, however, will realize that there are few negatives to the proposal. Because these individuals will realize that they considered both sides of the issue (i.e., positives and negatives), they will have more confidence in their new attitudes. This logic implies that if the message arguments are very weak, it is people who are trying to be favorable and fail who will have more confidence in their old attitudes than people who are being objective. This is because everyone will realize that the proposal has many flaws (due to the very weak arguments) but only people who were trying to be favorable will realize that there are no positives. Because these individuals will have considered both sides of the issue, confidence in their attitudes will be enhanced (Rucker & Petty, 2004).

The studies we just reviewed suggest quite clearly that following an attempt to resist persuasion, people sometimes reflect on their resistance and its meaning for their attitudes. In the examples just provided, certainty in either an old or a new attitude was increased as a result of reflection on one's attempted resistance. These outcomes are depicted in Figure 11.1, along with situations in which confidence in either old or new attitudes can be

Trying to Counterargue

Able to Counterargue → Resist (no change)

- Impressed with ability (resisted strong arguments; seemed; easy to resist; counterarguments seemed good) → **More Confidence** in Old Attitude
- Unimpressed with ability (seemed very difficult to counterargue; counterarguments seemed bad) → **Less Confidence** in Old Attitude

Unable to Counterargue → Succumb (change)

- Impressed with inability (tried very hard; was difficult to counterargue; counterarguments seemed bad) → **More Confidence** in New Attitude
- Unimpressed with inability (did not try very hard; many distractions; counterarguments seemed good) → **Less Confidence** in New Attitude

FIGURE 11.1. How confidence in one's old and new attitude can be affected when people are able to counterargue and resist or are unable to counterargue and change. From Petty, Tormala, and Rucker (2004). Copyright 2004 by the American Psychological Association. Adapted by permission.

decreased as a result of reflection on one's resistance.[7] For example, Figure 11.1 suggests, and recent research confirms, that if people resist changing to a message but believe that their counterarguing attempt was flawed in some way (Tormala, Clarkson, & Petty, 2006), or that they resisted by illegitimate means (e.g., rejecting a minority source; see Tormala, Petty, & DeSensi, in press), certainty in one's attitude can be decreased (see Petty, Tormala, & Rucker, 2004, for additional discussion).

A Metacognitive Approach to Attitude Change

The notion that people can hold both old and new attitudes with varying degrees of confidence has important implications for an analysis of what happens when attitudes change. Traditional models of persuasion hold that when attitudes change, the old attitude just disappears or is incorporated into the new attitude (e.g., see Anderson, 1971). More recent constructivist perspectives on attitudes (e.g., Schwarz & Bohner, 2001) hold that attitudes are newly constructed on each occasion when they are needed and thus there is no residue from prior attitudes. In contrast, our metacognitive model of attitudes (MCM) incorporates the notion that both old and new attitudes can coexist with differing levels of confidence (Petty & Brinol, 2006; Petty, Wheeler, & Tormala, 2003; Petty et al., 2006).

In particular, consistent with various "negation tag" models of judgment (e.g., Gilbert et al., 1993; Mayo et al., 2004), the MCM holds that when people shift from one attitude to another, the old attitude can be tagged as invalid or low in confidence (see Figure 11.2 for a person who starts out liking chocolate cake but then comes to dislike it after reading a message about its fat content). An important implication of this model is that if people do not retrieve the associated invalidity tag that is linked to a rejected attitude, then the two opposing evaluations (old and new) could be jointly activated producing an ambivalent-like state. Notably, this ambivalence would not be *explicit* because when thinking carefully, people would recognize that they only truly endorse one side. However, when not thinking carefully, activation of both evaluations (without invalidity tag) could produce a state of *implicit ambivalence*.[8]

To examine this possibility, in a series of studies, Petty and colleagues (2006) created initial attitudes in participants and then changed them for one group, and reinforced them for another group. Thus, at time 2, there were two groups of individuals who felt positively or negatively about some attitude object. The only difference was that one group had always felt positively or negatively whereas the other group used to feel the opposite way (i.e., their attitudes were changed). Petty and colleagues reasoned that if people who used to feel differently experienced some implicit ambivalence, they should engage in greater information processing with respect to a message relevant to the attitude, just as do individuals who experience explicit ambivalence (e.g., Maio, Bell, & Esses, 1996). Consistent with this idea, when people had a new attitude that conflicted with their old one,
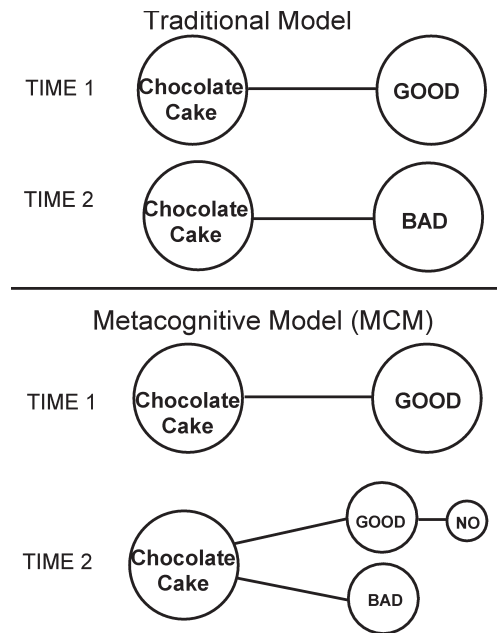
FIGURE 11.2. Comparison of the metacognitive model (bottom panel) with the traditional model (top panel) of attitude change. From Petty, Tormala, Briñol, and Jarvis (2006). Copyright 2006 by the American Psychological Association. Adapted by permission.

their attitudes were more influenced by message quality than when there was no conflict with an old attitude. In research on the MCM to date, confidence in old and new attitudes was not assessed. Yet, the results from this research are consistent with the notion that attitudes have validity tags that, depending on their activation, can influence peoples' reactions to attitude-relevant objects.

Thought Certainty and Persuasion: Self-Validation Effects

In addition to considering the consequences for attitude confidence of attempts at attitude change, persuasion researchers have also begun to study how metacognitive processes might contribute to attitude change itself. Perhaps the best example of this comes from the *self-validation hypothesis* (Petty et al., 2002). The key idea of this hypothesis is that just as attitude confidence is an important determinant of which attitudes predict behavior, thought confidence is an important determinant of which thoughts predict attitudes.

Considerable research has demonstrated that when people care about an issue or are motivated and able to think for other reasons, the number and valence of thoughts they have in response to a message determines the extent of attitude change (see Eagly & Chaiken, 1993; Petty et al., 1981; Petty & Wegener, 1998, for reviews). The self-validation hypothesis suggests that in addition to number and valence, it is also important to consider thought confidence. Research has shown that measuring the confidence people have in their thoughts to a persua-

sive message enhances prediction of the attitudes that are formed over and above a consideration of the valence and number of thoughts (e.g., Petty et al., 2002).

Furthermore, direct manipulations of thought confidence have a similar impact. In one study, for instance, following exposure to a message containing strong or weak arguments and a typical thought listing task (see Cacioppo, Harkins, & Petty, 1981), people were asked to think about situations in which they had felt confident or doubtful in their thinking (see Petty et al., 2002). Those who generated instances of confidence became more certain of the validity of their thoughts than those who generated instances of doubt. Furthermore, this confidence led to greater persuasion when the message arguments were strong and to less persuasion when the arguments were weak. This is because confidence led people to rely on the favorable thoughts generated to the strong arguments and the unfavorable thoughts they generated to the weak arguments. Individuals who were induced to doubt the validity of their thoughts were less reliant on them in forming attitudes even though the number and valence of thoughts was the same as those induced to feel confidence.

The self-validation framework provides a new explanation for how variables can impact attitudes. For example, prior research on head nodding had assumed that nodding one's head in a vertical (vs. horizontal manner) produced more positive attitudes either because vertical head nodding biased thinking in a favorable direction (Wells & Petty, 1980) or because head nodding served as a relatively simple affective cue (Tom, Pettersen, Lau, Burton, & Cook, 1991). The self-validation hypothesis suggested another possibility—that just as vertical head movements from others give us confidence in what we are saying, our own vertical head movements could give us confidence in what we are thinking. In a series of studies, Briñol and Petty (2003) found that head movements affected the confidence people had in their thoughts, and thereby had an impact on attitudes. Thus, when thoughts to a message were mostly favorable, vertical movements led to more confidence in the favorable thoughts generated and to more favorable attitudes than when horizontal movements were made. When thoughts were mostly unfavorable, however, vertical movements led to more confidence in the unfavorable thoughts generated and to less favorable attitudes than when horizontal movements were made.

The self-validation hypothesis proposes a new role that variables can play in persuasion situations in the context of the elaboration likelihood model of persuasion. That is, in addition to serving as cues, arguments, or affecting the number and valence of thoughts that come to mind (see Petty & Cacioppo, 1986), variables can also influence the confidence people have in their thoughts and thus whether they are relied on in forming attitudes. The self-validation hypothesis therefore provides a new mechanism by which classic persuasion variables can have an impact on attitudes.

For example, prior research has shown that the expertise of the source and a person's mood can serve in multi-ple roles (e.g., argument, cue, and affecting processing; see Petty & Wegener, 1998, for a review). Research on the self-validation hypothesis has shown that expertise and mood can also affect thought confidence. Source expertise presumably affects thought confidence because expert sources are more likely to present accurate information. If people can be more confident in the information presented by experts (Kaufman, Stasson, & Hart, 1999), then they can also be more confident in their thoughts regarding this information (see also Kruglanski et al., 2005). The mood prediction follows directly from the finding that people feel more confident when in some moods (e.g., happy) than in others (e.g., sad; Tiedens & Linton, 2001). In relevant research, Briñol, Petty, and Tormala (2004) showed that when the likelihood of thinking is high and people learn of the source's expertise after they have processed the message, source expertise influences thought confidence (see also Tormala, Briñol, & Petty, 2006). Briñol, Petty, and Barden (2006) have shown that in similar situations, people are more confident in their thoughts when placed in a happy rather than a sad mood following a message. If people are more confident in their thoughts about a message when they learn the message is from an expert and when they are in a happy mood, this means that source expertise and a happy mood can increase persuasion relative to a nonexpert and a sad mood when the arguments are strong but decrease persuasion when the message arguments are weak.

Finally, we remind readers that the self-validation hypothesis provides an alternative explanation for some cognitive fluency effects. For example, Tormala and colleagues (2002) found that people were more reliant on their thoughts when they were asked to generate an easy rather than a difficult number of them. In addition to the classic ease-of-retrieval effect, perceptions of fluency and thought confidence might be involved in other persuasion phenomena. Consider the voluminous work on matching and tailoring in persuasion. Matching refers to instances in which a message is matched to some aspect of an individual. For example, an image appeal might be presented to a person high in self-monitoring (Snyder & DeBono, 1985), or an emotional message might be presented to a person whose attitudes are based primarily on affect (e.g., Fabrigar & Petty, 1999; see Briñol & Petty, 2005; Petty, Wheeler, & Bizer, 2000, for reviews of matching work). According to a fluency account, if a message matches the person in some way, it may be easier to process. In one demonstration of this logic, regulatory fit (see Higgins, 2000) was shown to enhance processing fluency. Lee and Aaker (2004) presented promotion- and prevention-focused individuals with persuasive messages framed in terms of gains or losses. Participants reported that it was easier to process the messages when they matched (gain frame-promotion focus or loss frame-prevention focus) rather than mismatched (gain frame-prevention focus or loss frame- promotion focus) participants' regulatory focus. Furthermore, participants were more persuaded by the messages under matched rather than mismatched conditions. Although there are several

possible explanations for this effect (see Petty et al., 2000), it may be that the processing fluency stemming from regulatory fit leads message thoughts to be held with more confidence (see also Cesario, Grant, & Higgins, 2004). To the extent that the thoughts to the message are primarily favorable, enhanced confidence would increase persuasion. This fluency explanation for various matching effects should be examined in future research.

It is noteworthy that the self-validation findings results have been most pronounced under high thinking conditions. For example, research on head nodding (Briñol & Petty, 2003), source expertise (Briñol, Petty, & Tormala, 2004), and ease of retrieval (Tormala et al., 2002) showed that these variables affected confidence in thoughts for high- but not for low-need-for-cognition individuals and affected confidence when issue involvement was high but not when it was low. Relatively high elaboration presumably enhances self-validation effects for at least two reasons. First, if people have few thoughts, then thought confidence will have little effect. Second, the same variables that would increase elaboration (e.g., issue importance) would also likely increase thinking about one's thoughts. If people do not care enough to generate thoughts in the first place, they are hardly likely to care enough to think about the validity of their thoughts.

Metacognition about Persuasion Processes

Earlier in this section we noted that people sometimes think about their resistance to persuasion and generate inferences about their attitudes based on this resistance. More generally, research suggests that people have developed naive theories about the persuasion process, including what kinds of strategies they would use to persuade others (e.g., Bisanz & Rule, 1989), what strategies might be effective in persuading them (Friestad & Wright, 1994, 1995), and what strategies they use to resist influence (Jacks & Cameron, 2003). However, there is little (if any) empirical work on how these naive theories influence actual persuasion. Such research could be quite informative.

Although there is little accumulated research about how people's chronic persuasion theories affect attitudes, some work has begun to explore how people's online theories of attitude change can modify the persuasion process. For example, research by Mazursky and Schul (2000) suggests that people sometimes reflect on their information-processing strategy following receipt of a persuasive message so that they can modify it for future messages if the initial strategy proves ineffective. In two studies investigating this issue, they gave people information about a consumer product along with information about the source of the information. For half of these individuals, one piece of information provided by the source was said to be invalid, whereas the other half did not learn this. Then, all participants received another message about a new consumer product either from this source or from a more credible one. Some of the decision makers were highly involved with the consumer product whereas others were not. When consumers had not received any prior invalid information, judgments of

the involved participants appeared to be based on their effortful consideration of the evidence presented whereas those less involved relied on simple cues, consistent with much work on dual-process models of persuasion (e.g., Chaiken et al., 1989; Petty & Cacioppo, 1986). That is, high-involvement participants took longer to make their decisions than low-involvement individuals (an indication of their more deliberative processing), and low-involvement individuals relied more on the credibility of the information source (see Petty, Cacioppo, & Goldman, 1981). However, when recipients had received some invalid prior information, participants appeared to change their information-processing strategy. That is, now it was the low-involvement individuals who took more time to evaluate the messages, and it was the high-involvement individuals who relied more on the source credibility information.

THE SELF AND INDIVIDUAL DIFFERENCES

We have seen that metacognition can play a role in various attitudinal phenomena. One of the most studied attitude objects is the self (Baumeister, 1998). The self includes one's cognitive representation of oneself, composed of self-schemas and self-knowledge, as well as one's evaluation of oneself, or self-esteem. Similar to any other attitude object, the relevant beliefs (self-concept) and evaluations (self-esteem) toward the self can be accompanied by metacognitions about their origin, content, evaluation, amount, and so on. This section reviews the importance of considering various metacognitive aspects of the self.

Like research on attitudes, most of the metacognitive research in the self domain has focused on the *confidence* dimension. That is, researchers have studied peoples' subjective certainty about the validity of their cognitive representation of themselves and their self- evaluations. Similar to the literature in other domains, confidence typically has been measured by asking people to rate the degree to which they were certain or uncertain about their self-beliefs or their self-esteem.[9] Across different measurement and induction techniques, research reviewed in this section reveals that it is critical to consider confidence in order to understand the functioning of the self-concept and self-esteem. In the first part of this review, we cover research examining the consequences of *self-concept confidence* for various outcomes. Next, we describe studies dealing with *self-esteem confidence* and its implications for a variety of relevant areas. In the final section, we move from the study of metacognition regarding general beliefs and evaluations about the self to the domain of metacognition about more specific self-dimensions. Thus, the last section explores metacognition and individual differences.

Consequences of Self-Concept Confidence

Similar to the literature in attitude strength (e.g., Petty & Cacioppo, 1986; Petty & Krosnick, 1995), self-beliefs that are held with greater confidence are also more stable,

more resistant to change, and more predictive of behavior. For example, people who report greater certainty about their self-beliefs have been found to be more stable in their self-views (Pelham, 1991). Furthermore, Pelham and Swann (1994) showed that people are more likely to actively solicit self-consistent social feedback (i.e., to actively seek out feedback that supports their existing self-views) in domains in which they are most certain. Pelham and Swann also showed that interaction partners are more likely to confirm peoples' confident rather than doubtful self-views. In short, when people report high confidence in their self-views, they are likely to behave and to be perceived in ways that are consistent with those personal views (see also Pelham, 1991).

The confidence with which self-beliefs are held can influence not only individuals' stability and consistency but also other important dimensions, such as resistance to change. Similar to the literature in attitude strength (e.g., Bassili, 1996; Gross et al., 1995), self-beliefs that are held with greater confidence are more difficult to change. Swann and Ely (1984) found, for instance, that people who reported being relatively uncertain about their personality (e.g., extraversion) showed more change in response to a series of leading questions (i.e., a technique that leads people to provide evidence that confirms the premises in the leading questions) than those who were certain about their self-concept. In follow-up research, Swann, Pelham, and Chidestu (1988) replicated this finding and showed not only that individuals high in self-belief certainty resist leading questions better than those with relatively lower certainty but that boomerang effects can also occur. That is, leading questions can cause confident people to change in a direction opposite to the leading questions when encouraged to make statements that are consistent with but more extreme than their own beliefs.

Subsequent research has shown that self-belief confidence can also influence how people resist other forms of influence. For example, Sedikides (1995) found that relative to participants who reported uncertainty about their traits, those who expressed more self-belief certainty were more resistant to the biasing influences of a mood induction (happy, sad, or neutral). This research suggests that being certain about oneself can lead to more resistance to information about both the specific beliefs about which one is certain and the more general factors (e.g., mood) capable of influencing those beliefs.

In a different line of research, Baumgardner (1990) found that certainty in self-attributes not only causes people to be more resistant and to behave in a more consistent and stable way but can also promote a sense of control over future outcomes, thus generating positive affect. Specifically, in one study Baumgardner manipulated confidence by providing participants with bogus feedback about their personality in which they were led to believe that based on their previous responses, the experimenter was uncertain or certain about the assessment of them. Although not empirically tested, the implied assumption in this experiment was that the confidence expressed by the experimenter influenced

participants' self-certainty. As anticipated, self-concept certainty led to positive self-affect as assessed by participants' self-rating. This suggests that the confidence with which self-beliefs are held can influence not only what people think and do but also how they feel.

Also paralleling the literature of attitude strength, the confidence with which people hold their self-related beliefs (e.g., "I'm intelligent") has a number of implications for behavior. In general, to the extent that individuals are certain of their self-beliefs, they are more likely to act in ways that are consistent with them. For example, a person who is certain that he is humorous and lazy is likely to choose situations that allow him to be funny and avoid those that demand his being productive. In research conducted to test this idea, Setterlund and Niedenthal (1993) manipulated self-concept certainty by asking participants to describe three times in which they acted in a way consistent (confidence) or inconsistent (doubt) with traits previously rated as highly self-descriptive. The result was that individuals who were manipulated to feel certain about their self-concept were more likely to use the self to guide decisions in a subsequent task in which they had to choose situations that allowed them to express aspects of their identity.[10]

In another study, Briñol and Petty (2003) found that self-belief confidence can influence self-esteem. As part of a supposed graphology study, participants were required to think about and then write down their best or worst qualities using their dominant or nondominant hand. Writing with the dominant hand was presumed to induce more confidence in the self-beliefs generated compared to the nondominant hand. Then, participants rated the confidence in their self-beliefs and reported their self-esteem. As expected, using the nondominant hand decreased the confidence with which people held the self-beliefs they listed. This occurred despite the fact that the actual quality of the self-beliefs did not vary across the hand conditions. As a consequence of the differential self-belief confidence, the effect of the direction of self-beliefs (best vs. worst qualities) on self-esteem was significantly greater when participants wrote their beliefs with their dominant rather than their nondominant hand. This study demonstrated that inducing doubts about possessing positive qualities tended to undermine self-esteem whereas inducing doubt about possessing negative qualities tended to enhance self-esteem. Furthermore, this study showed that the changes in self-esteem were mediated by changes in the certainty of the self-beliefs listed.

Finally, it is worth noting that although most of the research described in this section deals with confidence as the main metacognitive dimension, other metacognitive aspects occasionally have been explored in relation to self-beliefs. For example, Pelham (1991) found that positive self-beliefs rated as important were associated with more stability than positive self-beliefs considered relatively less important (see also Sedikides, 1995). Although this result shown for *self-belief importance* is similar to those found for self-belief confidence, importance and confidence are two relatively independent forms of metacognition (see discussion in the attitude section).

Consequences of Self-Esteem Confidence

As noted earlier, people can have confidence in their self-related beliefs as well as in their overall evaluation of the self. Similar to the outcomes described in the foregoing section, the confidence with which people hold their self-esteem has been found to have a number of important implications for different domains.

First of all, self-esteem confidence seems to moderate a variety of well-established findings in the self-esteem literature. For example, there is ample evidence that relative to individuals with high self-esteem, those with low self-esteem feel uncomfortable with success (presumably because of its inconsistency with their negative self-evaluation) and that as a consequence further success tends to be avoided. In a pioneering study, Marecek and Mettee (1972) showed that only participants with low self-esteem who also reported relatively high self-esteem certainty avoided success. Another illustration of the moderating role of self-esteem certainty can be found in the literature on self-handicapping. Self-handicapping refers to actions oriented to inhibit (or handicap) one's own performance in order to protect oneself from others' potential attributions that a failure, if it occurred, was due to the lack of ability (e.g., Arkin & Oleson, 1998). In their landmark study, Berglas and Jones (1978) found that individuals who were induced to doubt their abilities (by failing to solve insoluble problems) handicapped their own performance (by choosing a supposedly inhibiting drug in anticipation of a second set of similar problems) more than individuals induced to trust their abilities (by receiving success feedback on previous soluble problems). Extending this notion to the domain of self-esteem confidence, Harris and Snyder (1986) found that participants who were uncertain of their self-esteem were more likely to self-handicap (by low preparation for an upcoming test) than those who reported higher self-esteem certainty.

As another example of the moderating role of self-esteem confidence, consider work on discrepancy reduction. Social psychological literature has clearly documented that people can simultaneously hold incompatible beliefs, attitudes, feelings, and behavioral tendencies regarding oneself and others, and that these internal discrepancies are unpleasant and often result in negative affect and psychologically undesirable outcomes (e.g., Abelson & Rosenberg, 1958; Heider, 1958; Higgins, 1987; Kaplan, 1972; Newcomb, 1968; Norton, 1975; Osgood & Tannenbaum, 1955; Priester & Petty, 1996). A common approach to dealing with discrepancy is enhanced thinking or information processing (e.g., Abelson et al., 1968; Aronson, 1969; Festinger, 1957). By considering additional information, individuals may hope to gain enough information for one or the other side of the discrepancy in order to resolve or minimize the inconsistency, or at least the subjective discomfort that results from the discrepancy (e.g., Briñol, Petty, & Wheeler, 2006; Katz, Wackenhut, & Hass, 1986; Maio et al., 1996). For example, Woike and Baumgardner (1993) found that participants whose global and specific self-esteem were incongruent expressed greater interest in learning more about themselves than those whose self-worth was congruent. Importantly, however, this effect was only evident for participants who reported high confidence in their global and specific self-evaluations (see also Marsh, 1993).

Perhaps the domain of self-consistency is the area in which self-esteem certainty has been studied most extensively. Specifically, research suggests that self-esteem certainty can help to shed light on the classic debate concerning the potentially opposing human motivations for self-verification and self-enhancement. For example, people prefer feedback that is consistent with their own views when they report being certain of those views, but they prefer positive feedback when those views are less confidently held (Pelham, 1991). In line with this finding, as noted earlier, Swann and Ely (1984) found that participants who reported being certain of their traits tended to convince others to see them as they saw themselves. When participants reported being relatively uncertain about their traits, however, others saw them according to their own expectations. Extending this work from perceptions by strangers to perceptions by known others, Pelham and Swann (1994) found that self-views (both self-beliefs and self-esteem) matched other's views more strongly when those self-views were reported to be relatively certain. Taken together, these findings suggest that the need to be consistent operates mostly for aspects of the self-concept that are held with more confidence (see, Wright, 2001, for a review).[11]

In closing this section, it is important to consider not only the consequences of self-concept confidence but also some of its potential antecedents (for an extensive review of antecedents and consequences of self-certainty, see, DeMarree, Petty, & Briñol, in press). Similar to the literature on attitude certainty, theory and common sense suggest that the certainty of people's self-views is likely to be grounded in the *amount* of information they have about themselves as well as the *consistency* of this information. However, as described earlier in this section, self-concept confidence can depend on and be affected by other more transitory variables in the situation, such as the hand with which self views are written (Briñol & Petty, 2003) and the behavior of the experimenter (Setterlund & Niedenthal, 1993). In addition to these factors, and given that certainty in self-views is considered to be socially desirable, it seems possible that self-concept confidence might result from other operating motives related to impression management.[12]

Individual Differences

As just reviewed, people can differ in their self-concepts and self-esteem as well as the confidence with which these constructs are held. There are other differences among individuals in which certainty plays a part. The term "individual differences" refers to how people vary with respect to factors such as personality, motives, and abilities. Importantly, some conceptualizations of individual differences can be understood as metacognitions to the extent that they refer to thoughts people have about their thoughts or thought processes. Consider, for

example, the *need for cognition* (NC; Cacioppo & Petty, 1982), which refers to stable individual differences in the tendency to engage in and enjoy effortful thought. Consistent with the schema outlined in this chapter and an analysis of the NC scale items, people's judgments about their own NC represent an *evaluation* of thinking (e.g., "I find satisfaction in deliberating hard for long hours"). Although the NC scale has focused mostly on enjoyment of thinking (i.e., the evaluative component), NC also is related to the assessment of the *amount* of thinking (e.g., "I usually end up deliberating about issues even when they do not affect me personally"). That is, judging our own NC might imply thinking about how much we tend to think and how much we enjoy thinking. Importantly, in the NC scale, these assessments are not about any particular thoughts (as in most metacognition research) but are about thought processes in general.

People appear to have good insight into their own enjoyment of thinking and tendency to engage in it as the NC scale has proven to be a robust predictor of a wide variety of cognitive activities (see Cacioppo, Petty, Feinstein, & Jarvis, 1996, for a review). Individuals high in NC not only tend to think more about any given attitude object (e.g., Cacioppo, Petty, & Morris, 1983) but also devote more attention to their own thinking. As a result, high NC has been related to metacognitive processes described in this chapter such that individuals high in NC are more likely to evaluate their own thoughts for validity (Briñol & Petty, 2003; Briñol, Petty, & Tormala, 2004; Petty et al., 2002; Tormala et al., 2002), to engage in controlled (Martin, Seta, & Crelia, 1990; see Wegener & Petty, 1997) and automatic (Petty, DeMarree, Briñol, & Horcajo, 2005) bias correction processes, and to draw different metacognitive inferences based on their responses to persuasive messages (e.g., Tormala & Petty, 2004a).

Need for closure (NFC; Webster & Kruglanski, 1994) is another individual difference variable that could be understood as metacognitive in nature. NFC refers to people's desire for a definitive answer on some topic as opposed to confusion and ambiguity. Thus, NFC taps into a person's thoughts about their own thinking with some items explicitly linked to confidence in thinking ("I usually make important decisions quickly and confidently"). If a person high in NFC generates a thought that is assessed as a definitive answer to address the situation, that person should have confidence in that thought and should consider it an appropriate thought. Importantly, NFC represents a stable individual difference as well as a situationally evocable state. As a chronic dimension, the desire for definitive knowledge has been measured with the NFC scale (for properties of the scale, see Webster & Kruglanski, 1994; see also Neuberg, West, Judice, & Thompson, 1997). In general, being high in NFC has been shown to reduce the extent of information processing, to magnify primacy effects, to increase reliance on theory-driven versus data-driven processing, and also to enhance reliance on initial anchors and primes (see Kruglanski & Webster, 1996, for a review).

There are a variety of other individual differences variables that may be relevant for metacognition because they refer to different aspects of mental activity, such as causal uncertainty (Weary & Edwards, 1994), field dependence (Witkin et al., 1954), self-monitoring (Snyder, 1974), uncertainty orientation (Sorrentino & Short, 1986), need to evaluate (Jarvis & Petty, 1996), self-awareness (Carver, & Scheier, 1981), preference for consistency (Cialdini, Trost, & Newsom, 1995), resistance to persuasion (Briñol, Rucker, Tormala, & Petty, 2004), and defensive confidence (Albarracín & Mitchell, 2004). Although these variables deal in one way or another with how people think about their thinking, they do so in a relatively direct way. That is, most of the items contained in the scales consist of direct statements about thinking, rather than second-order thoughts such as the ones described earlier for NC. For example, the self-doubt scale (Oleson, Pohlmann, Yost, Lynch, & Arkin, 2000) measures individual differences in uncertain feelings about one's competence and ability. This scale contains items, such as "More often than not I feel unsure of my abilities." These self-doubt judgments take place at the direct, first level of cognition and are different from, for example, the second-order judgments described in the section on self-concept certainty.

A clear illustration of this distinction between people's self-perceptions and metacognitions can be found in the literature of individual differences relevant to minority groups. For example, social psychologists have developed numerous measures to assess individual differences in attitudes toward many groups considered to be stigmatized in some way (see Briñol & Petty, 2005, for a review). However, there are not only individual differences in evaluations of minority groups (which constitutes a direct judgment) but also individual differences in chronic motivations to control for prejudice toward these groups (which refers to the *evaluation* dimension or appropriateness of the former judgment). Among this second group of measures are the Motivation to Control Prejudiced Reactions scale (Dunton & Fazio, 1997), the Internal and External Motivation to Respond without Prejudice scale (Plant & Devine, 1998), and the Humanitarianism–Egalitarianism and Protestant Ethic Scales (Katz & Hass, 1988). As described in more detail in the next section, these instruments are effective in predicting differences in public and private endorsement of stereotypes as well as motivation to correct one's social judgments for inappropriate content.

BIAS CORRECTION

Our last domain of metacognition relates a number of types of thoughts about thoughts, including assessments of thought content (valence and source of the thought), evaluation, and validity. Although these various metacognitive assessments come into play in bias correction, the essence of correction rests in evaluation of the thought or judgment as relatively good or bad, wanted or unwanted, appropriate or inappropriate. As noted earlier, this evaluation of the thought can be distinguished from identification of the thought as being relatively positive versus negative toward the target. Thus, one can perceive

the thought as good or bad, appropriate or not, regardless of whether the thought itself is a positive or a negative one (e.g., "it is inappropriate to be favorable toward criminals"). Individuals' evaluations of their thoughts and perceptions can have sweeping effects on judgment and behavior. When a thought or perception is viewed as bad, unwanted, or inappropriate, people might try to avoid or to modify the thought or perception. People might also try to limit the effects of that thought or perception on subsequent judgments and behavior.

Thoughts or perceptions can be viewed as unwanted or inappropriate for a variety of reasons. In general, these reasons can be summarized by saying that perceptions are viewed as inappropriate or unwanted when they do not serve the perceiver's current judgment goals (Wegener & Petty, 1997). As noted earlier, in many cases, the judgment goal is likely to be arriving at a "correct" or "accurate" view of the target (Chaiken et al., 1989; Festinger, 1954; Petty & Cacioppo, 1986). Therefore, if a thought or perception is viewed as inaccurate, people are likely to be motivated to improve those thoughts or perceptions by using one of a number of possible strategies for bias correction. Of course, many other goals and motivations are possible. People might want to view themselves as possessing generally positive qualities (see Kunda, 1990) or a specific positive quality such as good health (e.g., Ditto & Lopez, 1992). In contrast, as reviewed earlier, people with negative self-views might be motivated to continue to view themselves negatively (e.g., Swann & Ely, 1984). In some cases, goals are more socially oriented, as when people are motivated to uphold procedural justice for courtroom defendants, even if accuracy is better served by judging the defendant as guilty (e.g., Fleming, Wegener, & Petty, 1999). In research on need for closure, people have been shown to view certain thoughts and ideas as unwanted precisely because they oppose the goal of making a quick decision (e.g., Kruglanski & Webster, 1991).

Corrections can be distinguished from the previously discussed metacognitions by considering the extent to which the issue of bias is salient to individuals (see Wegener & Petty, 2001; Wegener, Petty, Smoak, & Fabrigar, 2004). When bias is not salient, people may seek goal-appropriate perceptions of targets (often valid or correct perceptions). Seeking correctness would bear similarities to the *promotion* orientation described by Higgins (1998b). When seeking correctness, many of the previous types of metacognitions are likely (e.g., how confident am I in my judgment). However, to the extent that potential for bias is salient, people become more oriented toward taking steps to identify and avoid any biases at work (Wegener & Petty, 1997; similar to Higgins's *prevention* orientation).

We discuss three different types of corrections guided by a sense that one's cognition are somehow unwanted or inappropriate. These corrections are attempts at *subtracting* reactions to biasing factors when forming perceptions of targets, attempts to use *theories of bias* in seeking to formulate appropriate (goal-consistent) perceptions of targets, and attempts at *suppressing* or *inhibiting* a particular thought from coming to mind in the first place.

Subtraction of Reactions to Biasing Factors

Much of the research on bias correction from the 1980s and early 1990s focused on *partialing* or *subtraction* of reactions that are viewed as inappropriate because they are inferred to be responses to some irrelevant contextual variable rather than to the target. For example, research on Martin's (1986) set–reset model generally begins with blatant priming of responses consistent with one interpretation of an ambiguous target. Overlap in reactions to the prime and to the target result in *setting*, which is a default misattribution to the target of reactions that were really reactions to the prime (i.e., the context in which the target is judged). That is, the default perception is that reactions are "about the target" (Higgins, 1998a) rather than about the context.[13] Setting is said to produce assimilation to the context, which is consistent with traditional priming effects (e.g., Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979). For example, Martin and colleagues (1990, study 3) blatantly primed participants with the concept *persistent* or *stubborn* and then gave participants target materials that were ambiguous with regard to whether the target was persistent or stubborn. Consistent with the idea that "setting" to blatant primes requires relatively little thought, perceivers who were low in NC (Cacioppo & Petty, 1982) assimilated judgments of the target to the primed concept.

However, when motivation and ability to think are sufficiently high, the set–reset approach suggests that perceivers *reset* (i.e., partial out reactions to the context). Overlap in valence between reactions to the target and reactions to the context can lead to confusion over which reactions belong to the context and which to the target. Thus, attempts to partial out reactions to the context could result in some reactions to the target being misattributed to the context. When this occurs, resetting not only reduces assimilation to the context but can result in contrast (i.e., judgments of the target that are even less like the context than if the context were not present). Consistent with this idea, perceivers high in NC in the Martin and colleagues (1990, study 3) research contrasted their judgments of the target away from the blatantly primed concepts.

Similarly, the inclusion–exclusion model (Schwarz & Bless, 1992a) treats *inclusion* of information in one's representation of the target as the default mental operation and treats *exclusion* of the information as requiring greater cognitive effort. Inclusion–exclusion studies typically begin with a target that can either be a superordinate category (within which a specific member of the category can be included or excluded) or a subordinate member (or subset) of a category (to which characteristics of the category as a whole can be ascribed or excluded). When the context information is *subordinate* to the target category, including that information in the representation of the category leads to assimilation of the category to the exemplar. For example, thinking of an extreme exemplar, such as a popular athlete, could increase perceivers' positivity toward the athlete's gender or ethnic group (e.g., Bodenhausen, Schwarz, Bless, & Wänke, 1995; Coats & Smith, 1999; Schwarz & Bless,

1992b). Exclusion of the exemplar can reduce the assimilation (if the exemplar is simply subtracted) or can lead to contrast (if oversubtraction occurs, similar to resetting, or if the excluded exemplar serves as a standard of comparison for the category; Schwarz & Bless, 1992a). The standard of comparison could directly influence perceptions of the target (as in judgment theories like adaptation-level theory, Helson, 1964, or social judgment theory, Sherif & Hovland, 1961) or could redefine the meaning of the response scale anchors (as in variable perspective theory, Ostrom & Upshaw, 1968).

If information *superordinate* to the target comes to mind, that superordinate information can lead to assimilation if the target is included in the superordinate category. This inclusion also sets the stage for generation of additional features of the target using the category. Similar to encountering subordinate information, exclusion of superordinate information also leads to a decrease in assimilation or to contrast. This could be because of subtraction of features associated with the category or because of establishment of an extreme standard of comparison.

Like the set–reset approach, inclusion versus exclusion is determined, in part, by whether a particular type of information is regarded as appropriate for inclusion in the representation of the target. Most studied determinants of inclusion versus exclusion directly reflect attention to the role of categorization in context effects. For example, information that is representative of the target category should generally be included, whereas information not representative of the target category should be excluded. Lack of representativeness could be caused by factors such as temporal distance between the context and target (Strack, Schwarz, & Gschneidinger, 1985), lack of feature overlap between context and target (Herr, Sherman, & Fazio, 1983), or receipt of information from separate, potentially inconsistent sources rather than from a single source (Hilton & von Hippel, 1990). Similarly, information is more likely to be included when category width is high (e.g., asking about politicians in general after asking people about three political scandals) but excluded when category width is low (e.g., asking about a particular politician not involved in the previously rated scandals; Schwarz & Bless, 1992b). Finally, when stimuli are presented together, they are more likely to be perceived as a unit, resulting in assimilation, but when stimuli are presented sequentially, they are perceived separately, resulting in contrast (e.g., Seta, Martin, & Capehart, 1979; Wedell, Parducci, & Gieselman, 1987).

Similar to the set–reset approach, Schwarz and Bless (1992a) also noted that exclusion of reactions can occur when participants realize that previously encountered stimuli (e.g., primes) other than the target may have created the reactions (e.g., Lombardi, Higgins, & Bargh, 1987; Strack, Schwarz, Bless, Kübler, & Wänke, 1993). Schwarz and colleagues have also shown that conversational norms can motivate people to deliberately exclude information from target categories. For example, Schwarz, Strack, and Mai (1991) asked participants about their marital satisfaction and about their general life satisfaction. When people reported their life satisfaction af-

ter their marital satisfaction, the correlation between these measures was high, presumably because people included the recently activated information about marital satisfaction when reporting life satisfaction. However, when the questions were described as assessing two areas of life that may be important for overall well-being, the correlation between the items decreased substantially because marital satisfaction information was presumably excluded from ratings of life satisfaction (see also Strack, Martin, & Schwarz, 1988).

Attempts to partial or subtract perceptions include a variety of features that are highly metacognitive. People have to identify the likely sources of thoughts and reactions and have to determine whether those reactions are informative about the target as currently construed. If the thought or reaction is to be subtracted, there likely has to be some cognitive mechanism for setting the thought aside and focusing attention on the included thoughts instead. In the subtraction research conducted thus far, few measurements of such metacognitive mechanisms have been included. For example, one might imagine that in order to set aside a particular reaction, some type of monitoring of that reaction might have to take place. If so, then as in the thought suppression work described shortly (e.g., Wegner, 1994), the thought or reaction that is set aside might actually become rather accessible in memory as the monitoring occurs. Similarly, few studies have measured perceptions of appropriateness or of attributions of reactions to targets versus contexts. Therefore, although many of the manipulations used in this research would seem to suggest that these types of metacognitions are involved, future research could benefit from more direct assessment of them.

Theory-Based Correction

An alternative view on the metacognitive activity of bias correction relies on peoples' perceptions of the bias(es) at work in a given setting.[14] For some time, researchers have noted that people might realize that a bias is at work and might make efforts to overcome that bias (e.g., Higgins et al., 1977; Strack, 1992; Thompson, Fong, & Rosenhan, 1981; Wyer & Budesheim, 1987). Early attention to lay beliefs or theories about biases focused on the fallibility of such perceptions. For example, Wilson and his colleagues have shown that people often believe that factors affect their perceptions even when the factors do not, and people often believe that they can resist influences that they cannot (e.g., Nisbett & Wilson, 1977; Wilson & Brekke, 1994; Wilson, Houston, & Meyers, 1998).

Despite the potential inaccuracy of beliefs about bias, social perceivers may use such perceptions in attempts to avoid bias. The flexible correction model is a theory of bias correction based on social perceivers' use of naive theories of bias (Petty & Wegener, 1993; Wegener & Petty, 1995, 1997). Research guided by this approach has shown that people correct their judgments in different directions when they hold theories of opposite biases (e.g., Wegener & Petty, 1995; Wegener, Petty, & Dunn, 1998), even when those opposite theories of bias are for different people perceiving the same context and target

(Wegener & Petty, 1995). People correct for biases they believe exist, even if there is no real bias (e.g., Wegener & Petty, 1995). This can then create the opposite bias, as when corrections for perceived negativity toward the dislikable source of a persuasive message leads that source to be more persuasive than a likable source (Petty, Wegener, & White, 1998; see also Schul & Goren, 1997).

Corrections for perceived bias also mean that people sometimes correct primarily for one bias (the one most salient or for which clear beliefs exist), even if other biases truly influence perceptions of that same target. For example, Sczesny and Kühnen (2004) showed that people believe that gender can bias judgments of leadership qualifications but do not hold similar beliefs about physical features associated with masculine versus feminine appearance. When encountering mock application materials (including photos), research participants showed different corrections for gender versus physical features. When cognitive load was high (and metacognitive activity was likely curtailed), participants were more likely to employ male than female applicants and people who possessed masculine rather than feminine features. When cognitive load was low, however, research participants overcorrected effects of gender (the bias for which relevant naive theories existed), such that female applicants were viewed more favorably than male applicants. In contrast, the biasing impact of masculine versus feminine features (for which no naive theories existed) remained unchanged.

Research on theory-based correction also illustrates the breadth of domains in which corrections for bias occur. Studies of theory-based correction have been conducted in domains as diverse as affect and judgment (e.g., Berkowitz, Jaffee, Jo, & Troccoli, 2000; DeSteno, Petty, Wegener, & Rucker, 2000), impression formation (Isbell & Wyer, 1999; Stapel, Martin, & Schwarz, 1998), stereotyping (Lepore & Brown, 2002; Strack & Mussweiler, 2001), persuasion (Petty et al., 1998; Schul & Goren, 1997), recognition memory (Förster & Strack, 1998), and courtroom judgment (Thompson et al., 1981; Wegener, Kerr, Fleming, & Petty, 2000).

Consistent with the idea that online metacognition is most likely when motivation and ability to think are high, most studies of theory-based correction that have manipulated or measured motivation or ability to think have shown greater theory-based correction with high levels of thinking (e.g., DeSteno et al., 2000; Sczesny & Kühnen, 2004). There may be important exceptions to this general pattern, however. For example, if certain corrections are performed repeatedly, they may become less effortful (see Wegener & Petty, 1997; cf. Glaser & Banaji, 1999), even automatic. In one study, for instance, Maddux, Barden, Brewer, and Petty (2005) assessed the automatic evaluations that Whites had of Black versus White targets in particular contexts. The White participants varied in their motivation to control racial prejudice as assessed with a scale developed by Dunton and Fazio (1997). People who are high in motivation to control prejudice are presumably highly practiced in controlling any prejudiced reactions they might be feeling.

When people high and low in motivation to control prejudice evaluated Blacks versus Whites in innocuous settings (e.g., church and garden), neither group showed much prejudice replicating prior research (see Barden, Maddux, Petty, & Brewer, 2004; Wittenbrink, Judd, & Park, 2001). However, when the same Black and White targets were evaluated in settings that might suggest anti-Black bias (e.g., jail and dingy factory), those low in motivation to control prejudice showed an anti-Black bias, but those high in motivation to control prejudice showed a significant pro-Black bias, consistent with the idea that they were (over)correcting their judgments. Because the attitude measure tapped automatic evaluative responding (see Fazio, Sanbonmatsu, Powell, & Kardes, 1986), this research is consistent with the idea that highly practiced corrections can be executed automatically in certain contexts.

Some biases might be so salient or obvious that people would adjust their ratings without much additional consideration of the target. This might occur in studies of the sleeper effect in persuasion when a *discounting cue* says that the previous message was false and, in fact, the opposite position has stronger support (e.g., Gruder et al., 1978). Such strong discounting cues are necessary for production of the sleeper effect (see Kumkale & Albarracin, 2004). From a theory-based correction point of view, strong discounting cues might produce a sleeper effect because they alert message recipients to an obvious bias that brings about an initial adjustment of ratings without theory-guided "reprocessing" of the initial information (after all, why go back to pay attention to the false information). If initial processing of the information was high and the message is encountered prior to the discounting cue (two other requirements for the sleeper effect, see Kumkale & Albarracin, 2004; Petty & Cacioppo, 1986), the relatively nonthoughtful correction only temporarily shifts ratings, whereas attitudes based on initial thinking about the information are more likely to persist over time. In other circumstances, however, more thoughtful corrections could lead to persisting "corrected" views (Wegener & Petty, 1997), which would undermine the sleeper effect (see Priester, Wegener, Petty, & Fabrigar, 1999, for additional discussion). Whether corrections are accompanied by considerable or little thought is a fertile area for future research.

There might also be circumstances in which high levels of initial thinking make biases harder to identify and to correct (Petty & Wegener, 1993). For example, high levels of elaboration should lead to a great deal of integration of perceptions with existing knowledge structures (Petty & Cacioppo, 1986). Depending on the type of knowledge available about (or related to) the target, the highly integrated view of the target may seem justified by the existing information. In other words, if the reactions to the target seem to be an accurate reflection of available information, it may be that the people would view their opinions of the target as relatively appropriate and unbiased. Of course, this would undermine any need to correct the perception of the target (see Wegener, Clark, & Petty, 2006; see also Schul & Burnstein, 1985). High

levels of integration could also spread the bias across many disparate perceptions, which might also make identification of bias and correction more difficult.

Although significant research supports the possibility of corrections based on peoples' perceptions of bias, much work remains to be done. Future research will likely address more directly the metacognitive mechanisms at work when people spontaneously identify potential biases. It could be that people use accessible or salient theories to guide searches for potential bias, but there could be a variety of additional cues to bias such as a mismatch between past and present perceptions of the target, matches between the valence of a salient situational factor and the current perceptions, or others (see Wegener et al., 2001, for additional discussion).

Thought Suppression

On some level, if it were possible, keeping biased thoughts from coming to mind at all would be the ideal way to avoid bias. Yet, research on thought suppression attempts have suggested limited utility in this strategy. For example, Wegner and his colleagues have found that many types of thoughts become hyperaccessible after initial attempts to suppress them (e.g., Wegner, 1994; Wegner & Erber, 1992). In one study, Macrae, Bodenhausen, Milne, and Jetten (1994) asked participants to write a paragraph about a day in the life of a skinhead either with or without instructions to avoid stereotypical thoughts about the target. Later, a lexical decision task showed that the stereotype of skinheads was more accessible (i.e., people were faster to recognize stereotype consistent words as words) when the initial essay had been written while attempting to suppress stereotypical thoughts. According to Wegner and colleagues, this ironic hyperaccessibility of the suppressed content comes from a monitoring process that scans conscious thought for the presence of the unwanted material. As this process checks for the to-be-suppressed material, it repeatedly activates the concept in memory.

Consistent with online metacognitive activity being more likely when motivation and ability to think are high, attempting to suppress thoughts is an attention-consuming activity (Macrae, Bodenhausen, Milne, & Wheeler, 1996), and ability to suppress is undermined with time pressure (e.g., Wegner, Erber, & Bowman, 1993, described in Wegner, 1994). Thought suppression may become easier and more effective in avoiding rebound with practice (Kelly & Kahn, 1994; Monteith, Sherman, & Devine, 1998; cf. Smith, 1994; Wegner, 1994). Some reasons for this may include that people who practice suppressing particular thoughts might also have more replacement thoughts (i.e., thoughts that distract one from the to-be-suppressed concept) at the ready (e.g., Monteith et al., 1998; Wegner, Schneider, Carter, & White, 1987) and more practiced individuals may become less likely to have the unwanted thoughts come to mind in the first place (e.g., Fazio, Jackson, Dunton, & Williams, 1995; Lepore & Brown, 1997).

Although there is sufficient evidence that attempts to suppress thoughts can make the thoughts more accessible at a later point in time, it is important to note that many studies of thought suppression allow for other types of correction in addition to or instead of suppression per se. For example, Wyer, Sherman, and Stroessner (2000) showed that suppression of stereotypes toward African Americans resulted in application of primed stereotype-consistent concepts to a race-unspecified target but not to an African American target. Moreover, this restriction in use of primed stereotype-consistent material for a race-identified target only occurred when people had sufficient cognitive capacity to engage in the metacognitive work of correction. Although it is possible that people resuppressed the primed concepts when encountering the African American target, it seems more likely that research participants were using one of the other corrective mechanisms to limit the impact of the primed concepts on judgments of the target. For example, Dunton and Fazio (1997) noted that motivation to control prejudice might create overcorrections for automatically activated negative reactions through the types of theory-based correction proposed in the flexible correction model (Petty & Wegener, 1993; Wegener & Petty, 1997; see also Maddux et al., 2005). When the race of the target was unspecified in the Wyer and colleagues research, perceivers might not have realized that their perceptions of the target could be biased by the activated reactions, thereby opening the door to the bias (see Monteith et al., 1998; Sczesny & Kühnen, 2004; Strack & Hannover, 1996; Wegener & Petty, 1997).

In addition to suppression of thoughts per se, one could also talk about suppression of judgments or behaviors. That is, rather than focusing on keeping a thought from coming to mind, people might focus on making sure a thought does not become an action (cf. Monteith et al., 1998). In the *social judgability* approach, Yzerbyt and his colleagues (1994) have pointed out that people might withhold judgments when they do not feel justified in making a judgment. This might often occur when people perceive the amount of information about the target to be insufficient for forming an accurate view of the target (e.g., Darley & Gross, 1983; Yzerbyt, Leyens, & Corneille, 1998). Interestingly, this caution in judgment can be overcome by people believing they have received individuating information about the target even if they really did not (Yzerbyt et al., 1994). Therefore, the perception of having received sufficient information can lead to greater willingness to use stereotypes in formulating judgments. The willingness to make stereotype-consistent judgments is likely due, in part, to perceiving the judgment as more appropriate or less biased when based on an information base that seems sufficient. The social judgeability research also provides a potential bridge between thought evaluation and perceptions of confidence/validity. As noted earlier, perceived lack of accuracy can lead to perceptions of thoughts or reactions as unwanted or inappropriate. In the Yzerbyt and colleagues research, people express greater confidence in their judgments when they believe they have received in-

dividuating information. That is, when people lack confidence in the validity of their reactions, they are unwilling to use those reactions for judgments, but when people possess greater confidence, they are willing to use those reactions.

Even when people prefer not to be prejudiced toward a target, inability to inhibit the use of stereotype-related information can influence the likelihood of making judgments biased by the target's group membership. For example, elderly perceivers, who have trouble inhibiting a variety of types of thoughts about available information (Hasher & Zacks, 1988), also are more likely to provide stereotypical judgments, even if they are instructed to avoid race-related information and even when older perceivers profess greater motivation to avoid prejudiced responding (von Hippel, Silver, & Lynch, 2000).

Comparison of Strategies

To date, there is precious little research directly comparing subtraction, suppression, and theory-based corrections. In an initial attempt, Strack and Mussweiler (2001) compared *recomputation* of judgments (i.e., setting aside of biased information and formulation of perceptions based on remaining information; Strack & Hannover, 1996; Strack & Mussweiler, 2001; cf. Schwarz & Clore, 1983) with *adjustment* of responses (generally consistent with theory-based corrections or subtraction). Strack and Mussweiler designed studies so that recomputation would lead to more stereotypical judgments (because available information was highly stereotype consistent), but adjustment would lead judgments to be less stereotypical. When participants received a brief resume about a hypothetical job applicant, an instruction asking people not to be influenced by gender led to less stereotypical ratings of the applicant. This could have been due to theory-based correction or to subtraction effects. However, when a large number of stereotype-consistent behavioral episodes were presented about the target, the same correction instruction led to an increase in stereotypical ratings (consistent with recomputation). In a follow-up study, Strack and Mussweiler provided participants with the large number of behavioral episodes but varied whether perceivers were given a goal of forming an accurate judgment or of being fair and not using gender to make judgments. Similar to the earlier study, the accuracy goal led to judgments that were consistent with the presented information. However, the fairness goal led to judgments that were less consistent with gender stereotypes.

In two studies, Dove, Wegener, and Petty (2001, 2003) used a similar design but made the same amount of information either available or unavailable just prior to judgment. When the information was available, judgments were consistent with recomputation. However, when information was unavailable, judgments were consistent with theory-based corrections. It is possible that recomputation is more likely when perceivers form their opinions in an online rather than memory-based fashion (Hastie & Park, 1986). Yet, there might also be a variety of settings in which memory-based recomputation is feasible. When participants received a conditional correction instruction (i.e., they were asked to correct if they perceived anything as biasing them; see Stapel et al., 1998), both the recomputation and theory-based correction patterns were more pronounced for people high in NC (Cacioppo & Petty, 1982).

How to interpret such effects remains a matter for discussion. Strack and Mussweiler (2001) interpreted their results as signifying a preference for recomputation when both options were open to social perceivers. However, some type of adjustment might happen more often, with recomputation limited to situations in which social perceivers see recomputation as feasible and reliable (Dove et al., 2001, 2003). At any rate, comparisons among the correction processes are just beginning. There are likely to be a host of moderators of such outcomes. For example, it seems unlikely that accuracy-based instructions would always result in recomputation when information is available about the target. This should depend on factors such as the accessibility and salience of the theory of bias as well as the salience of the biasing agent itself (Wegener et al., 2001). Undoubtedly, future work will include additional comparisons of the correction processes, including studies of moderators of when each process dominates.

BASIC PRINCIPLES AND NEW DIRECTIONS

In this chapter, we described social psychology's major research findings on metacognition. Consistent with most prior literature (e.g., Jost et al., 1998), we defined metacognition as second-order thoughts, or thoughts about our primary thoughts or thought processes. That indeed constitutes a basic principle in human cognition:

- *Principle 1: There is primary and secondary cognition. Primary thoughts are those that occur at the direct level of cognition, involving initial associations. Following a primary thought, people can also generate other thoughts that occur at a second level which involve reflections on the first-level thoughts or the process that generated these thoughts.* Perhaps the most important conclusion of the work we reviewed is that in general, second-order cognition has an impact on first-order cognition. This is summarized in principle 2.

- *Principle 2: Second-order cognition can magnify, attenuate, or even reverse first-order cognition.* When increasing confidence in an attitude causes it to exert a larger impact on behavior than it did previously, metacognitive factors are exerting a magnifying effect on one's attitudes (see Fazio & Zanna, 1978). When people shaking their heads while listening to a message rely on their thoughts less than they would have if they had not been shaking, metacognitive factors are exerting an attenuating factor on one's thoughts (see Briñol & Petty, 2003). When people want to control for their perceived biases and then show an opposite bias, metacogitive processes are reversing the effects of first-order cognition (see Petty et al., 1998; Wegner, 1994).

In reviewing the relevant literature, we argued that the same categories that have proven effective for classifying primary thoughts could also be used to categorize metacognitive thoughts. This is summarized in principle 3.

• *Principle 3: Second-order thoughts can be coded into the same categories that have already proven effective for classifying primary thoughts, such as target, evaluation, number, and confidence.* By grouping the many specific metacognitions into meaningful categories, we aimed to provide a useful guide to organize and facilitate access to key findings across diverse literatures in social psychology.

The main areas of social psychological research in which metacognition has been examined most extensively are (1) memory and cognitive fluency, (2) attitudes and persuasion, (3) the self and individual differences, and (4) bias and correction processes. Although different categories of metacognition were examined in these different areas, it is clear that across areas, the most studied dimension has been the confidence one has in one's thoughts and judgments. As reviewed, confidence has proven to be relatively independent of accuracy and has been shown to be affected by a wide variety of situational and individual variables.

One area that has not been explored much in prior research concerns the various bases of confidence. For example, in one study Haugtvedt and Petty (1992) found that although both high and low NC individuals changed their attitudes in response to a message from a high credible source and developed equivalent levels of confidence in these attitudes, the confidence stemmed from different sources. In particular, for those high in NC, confidence increased as the number of message arguments they could recall increased. For those low in NC, however, confidence was tied to the perceived credibility of the source. Though speculative at the moment, it seems quite plausible to argue that different bases of confidence might lead to different outcomes (see also Gross et al., 1995). For example, it could be that some bases of confidence are more likely to be consequential than others.

In addition to examining the content underlying confidence, it would also be important to examine the processes leading to confidence judgments. For example, just as primary judgments can be based on much or little thought (Kruglanski & Thompson, 1999; Petty & Cacioppo, 1986), so too can second-order cognition. That is, people's judgments of confidence can be based on a careful consideration of the reasons why they should be confident, or they can result from a simple inference based on shaking one's head. To the extent that second-order cognition follows the same principles as primary cognition, one would expect thoughtful confidence to be more long lasting and consequential than confidence that stems from simple cues and inferences.

Finally, it would be worth studying whether metacognitive judgments are formed online or are retrieved from memory (Hastie & Park, 1986). Just as primary judg-

ments can operate in either way, so too do we propose that this holds for metacognitive judgments. To date, there is no research conclusively demonstrating that metacognitive tags (e.g., certainty and importance) have any structural basis in memory, though some research is certainly consistent with this possibility. These considerations lead us to our first speculative postulate that is worthy of research attention.

• *Principle 4: The content and process bases of metacognitive judgments are likely to be as consequential as are the bases of primary cognition.* In addition to examining confidence (and other metacognitive constructs) based on different content and process considerations, a potential avenue for future research would be to assess various metacognitive properties with implicit measures. This raises the possibility that just as some researchers have argued that primary cognition can be held at explicit and implicit levels, so too might this be the case with secondary cognition. That is, a person might have high confidence in some judgment at the explicit level but low confidence at the implicit level. Furthermore, this implicit uncertainty could potentially guide thinking and behavior (e.g., Briñol, Petty, & Wheeler, 2006). As described earlier in this chapter, for example, when attitudes are changed, people may sometimes have high confidence in their new attitudes at the explicit level but low confidence at the implicit level due to the conflict with the old attitude (Petty et al., 2006). This leads to our second speculative principle.

• *Principle 5: Although explicit metacognitive activity is generally more likely to take place when people have the motivation and ability to attend to and interpret their own cognitive experiences, metacognition might also operate outside awareness with important consequences for social judgment and behavior.* Indeed, the possibility of assessing mental constructs at the implicit level might open the door for a new generation of research relevant to metacognition. For instance, a common characteristic of most of the research covered in our review is that metacognitive activity is more likely to take place when people have the motivation and ability to attend to and interpret their own cognitive experience. Research on bias correction has shown, for example, that in order to correct for mental biases, people need to identify the likely source of their thoughts, determine whether those reactions are informative, and spend time and attention trying to modify, suppress, substitute, and/or correct those thoughts (Petty & Wegener, 1993; Wilson & Brekke, 1994). In contrast to all these mental activities that require extensive cognitive effort, recent research has shown that correction processes can also occur automatically if the correction is a highly practiced one (Maddux et al., 2005). Future research should explore what variables moderate the relationship between explicit and implicit metacognition. In sum, although most of the research covered in this chapter has focused on the power of explicit metacognition to modify the impact of explicit primary cognition, both primary and secondary thoughts may also operate outside awareness with important consequences for social judgment and behavior.

NOTES

1. People's primary thoughts can also be classified into other dimensions such as fluency (e.g., how accessible the thought is) or time (e.g., how recent the thought is; see also Cornoldi & Vianello 1992), but these dimensions are used infrequently in coding thoughts.

2. With respect to these dimensions, it is possible for objective judges to try to infer the person's perception, but this has not been attempted in research to date.

3. The positive effect of mere exposure is most likely when conditions favor minimal processing of the repeated stimuli (Bornstein, 1989). As stimulus processing increases, it becomes more likely that only positive items will enhance in favorability whereas negative items will decrease. When tedium or boredom sets in with excessive exposure, even positive items can be rated more negatively (Cacioppo & Petty, 1979).

4. Boninger et al. (1995) argue that this is to avoid confusion on the part of research participants (see also Abelson, 1988), and because measures of attitude importance and issue importance are highly correlated. Yet, in some individual studies, researchers have asked about attitude importance per se (e.g., "how important is your attitude toward the issue of doctor assisted suicide to you personally"; Haddock et al., 1999).

5. For example, one might imagine that with respect to objective ambivalence, one could get very different ambivalence scores if one asked: What are the good and bad things about same-sex marriage (the attitude object) versus What are the good things (bad things) about being very favorable toward (or being very opposed to) same-sex marriage?

6. In a somewhat similar vein, when people refuse to engage in an attitude consistent behavior for a high incentive, they become even more extreme in the direction of their initial opinions than when they refuse for a lower incentive. Presumably, people reason that if they did not go along and were given a high incentive to do so, their original attitude must really be valid. Increasing attitude extremity is one way to assert the validity of the attitude (Darley & Cooper, 1972). If attitudes did not become more extreme as a result of the manipulation, it would be reasonable to expect that confidence in the original attitude would have increased via similar attributional reasoning.

7. Of course, if people counterargue their own position rather than an external message, they can lose confidence in their attitude; Koriat et al., 1980).

8. The MCM is compatible with a "dual-attitudes" approach (i.e., the proposition that a person can hold conflicting implicit and explicit attitudes) in many respects. It diverges in part because the dual-attitudes approach emphasizes that each attitude operates in different situations and has ignored the possibility of joint activation (e.g., see Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Wilson, Lindsey, & Schooler, 2000).

9. Although directly asking people to rate their certainty has been the most used procedure to assess confidence, other, more indirect techniques sometimes have been used such as the intraindividual standard deviation of self-esteem scores (see, e.g., Wright, 2001, for a review). Alternatively, these indirect measures can be considered *consequences* of self-certainty.

10. Although the manipulation in this study was intended to influence certainty, it might have just primed consistency or inconsistency, thus accounting for the results. Because a manipulation check for confidence was not provided, the issue remains unclear.

11. Because most of the studies conducted in this domain have been correlational, the presumed directionality of some of the effects remains unclear. For example, consider the studies in which self-certainty was described as leading to different forms of self-consistency, such as asking for confirmatory feedback (e.g., Pelham, 1991; Pelham & Swann, 1994). It would be plausible to argue the opposite directionality. That is, self-belief confidence might be the product, rather than the precursor, of interpersonal congruence. This argument can also be applied to the relationship between self-esteem certainty and other concepts, such as self-esteem stability and self-esteem clarity. Future research would benefit from designs in which confidence is manipulated rather than measured (e.g., Briñol & Petty, 2003).

12. To the extent that certainty in self-views is considered to be socially desirable, those with higher impression management concerns should be more likely to report greater certainty in their self-conceptions. Also, because individuals with higher self-esteem are more likely to present themselves in a positive way than are those with lower self-esteem (e.g., Baumeister, Tice, & Hutton, 1989), it may be the case that esteem differences in confidence reflect differing impression management concerns (e.g., Wright, 2001, see also Baumgardner, 1990; Story, 2004). Indeed, some other consequences of self-concept certainty described in this chapter (e.g., stability, resistance, and prediction of behavior) might also be partially due to impression management. Furthermore, people may report feeling confident in a given self-view in order to compensate for other unrelated internal doubts (e.g., McGregor, 2003; Rhodewalt & Morf, 1995).

13. Attributing reactions to the target or the context is quite similar to source monitoring in research on memory, wherein perceivers attempt to determine the cause of recollective experiences such as familiarity of a target (see Mitchell & Johnson, 2000).

14. For comparisons of theory-based correction with the "partialing" approaches, see Strack (1992), Wegener and Petty (1997), and Wegener, Dunn, and Tokusato (2001).

REFERENCES

Abelson, R. P. (1988). Conviction. *American Psychologist*, *43*, 267–275.

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, Th. M., Rosenberg, M. J., & Tannenbaum, P. H. (1968). *Theories of cognitive consistency: A sourcebook.* Chicago: Rand McNally.

Abelson, R. P., & Rosenberg, M. J. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, *3*, 1–13.

Alba, J. W., & Hutchinson, W. (2000). Knowledge calibration: What consumers know and what they think they know. *Journal of Consumer Research*, *22*, 62–74.

Albarracín, D., & Mitchell, A. L. (2004). The role of defensive confidence in preference for proattitudinal information: How believing that one is strong can be a weakness. *Personality and Social Psychology Bulletin*, *30*, 1565–1584.

Allport, F. H. (1924). *Social psychology*. Boston: Houghton Mifflin.

Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*, 171–206.

Arkin, R. M., & Oleson, K. C. (1998). Self-handicapping. In J. M. Darley & J. Cooper (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 313–371). Washington, DC: American Psychological Association.

Aronson, E. (1969). Cognitive dissonance: A current perspective. In L. Bercowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 1–34). New York: Academic Press.

Baker, S. M., & Petty, R. E. (1994). Majority and minority influence: Source–position imbalance as a determinant of message scrutiny. *Journal of Personality and Social Psychology*, 67, 5–19.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, 87, 5–22.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, 31, 637–653.

Baumeister, R. F. (1998). The self. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 680–740). New York: McGraw-Hill.

Baumeister, R. F., Tice, D. M., & Hutton, D. G. (1989). Self- presentational motivations and personality differences in self-esteem. *Journal of Personality*, 57, 547–579.

Baumgardner, A. H. (1990). To know oneself is to like oneself: Self-certainty and self-affect. *Journal of Personality and Social Psychology*, 58, 1062–1072.

Beck, A. T., & Greenberg, R. L. (1994). Brief cognitive therapies. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 230–249). New York: Wiley.

Begg, I., Armour, V., & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioral Science*, 17, 199–214.

Benjamin, A. S., & Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 309–338). Hillsdale, NJ: Erlbaum.

Berger, I. E., & Mitchell, A. A. (1989). The effect of advertising on attitude accessibility, attitude confidence, and the attitude–behavior relationship. *Journal of Consumer Research*, 16, 269–279.

Berglas, S., & Jones, E. E. (1978). Drug choice as a self-handicapping strategy in response to noncontingent success. *Journal of Personality and Social Psychology*, 36, 405–417.

Berkowitz, L., Jaffee, S., Jo, E., & Troccoli, B. (2000). On the correction of possible feeling-induced judgmental biases. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in cognition and judgments* (pp. 131–152). New York: Cambridge University Press.

Bisanz, G. L., & Rule, B. G. (1989). Bender and the persuasion schema: A search for cognitive invariants. *Personality and Social Psychology Bulletin*, 15, 4–18.

Bizer, G. Y., & Krosnick, J. A. (2001). Exploring the structure of strength-related attitude features: The relation between attitude importance and attitude accessibility. *Journal of Personality and Social Psychology*, 81, 566–586.

Blascovich, J. (1992). A biopsychosocial approach to arousal regulation. *Journal of Social and Clinical Psychology*, 11, 213–237.

Bodenhausen, G. V., Schwarz, N., Bless, H., & Wänke, M. (1995). Effects of atypical exemplars on racial beliefs: Enlightened racism or generalized appraisals? *Journal of Experimental Psychology*, 31, 48–63.

Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality and Social Psychology*, 68, 61–80.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106, 265–289.

Bornstein, R. F., & D'Agostino, P. R. (1994). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, 63, 545–552.

Brickman, P., Redfield, J., Harrison, A. A., & Crandall, R. (1972). Drive and predisposition as factors in the attitudinal effects of mere exposure. *Journal of Experimental Social Psychology*, 8, 31–44.

Briñol, P., & Petty, R. E (2003). Overt head movements and persuasion: A self-validation analysis. *Journal of Personality and Social Psychology*, 84, 1123–1139.

Briñol, P., & Petty, R. E. (2004). Self-validation processes: The role of thought confidence in persuasion. In G. Haddock & G. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 205–226). Philadelphia: Psychology Press.

Briñol, P., & Petty, R. E. (2005). Individual differences in persuasion. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes and attitude change* (pp. 575–616). Hillsdale, NJ: Erlbaum.

Briñol, P., Petty, R. E., & Barden, J. (2006). *The effects of mood on persuasion: A self-validation analysis*. Manuscript under review.

Briñol, P., Petty, R. E., & Tormala, Z. L. (2004). The self-validation of cognitive responses to advertisements. *Journal of Consumer Research*, 31, 559–573.

Briñol, P., Petty, R. E., & Tormala, Z. L. (2006). The malleable meaning of subjective ease. *Psychological Science*, 17, 200–206.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, 91, 154–170.

Briñol, P., Rucker, D., Tormala, Z. L., & Petty, R. E. (2004). Individual differences in resistance to persuasion: The role of beliefs and meta-beliefs. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 83–104). Mahwah, NJ: Erlbaum.

Broemer, P. (2001). Ease of recall moderates the impact of relationship-related goals on judgments of interpersonal closeness. *Journal of Experimental Social Psychology*, 37, 261–266.

Busey, T. A., Tunnicliff, J. L., Loftus, G. R., & Loftus, E. F. (1995, November). *Predicting picture memory performance: Not all confidence judgments are equal*. Paper presented at the 36th annual meeting of the Psychonomic Society, Los Angeles.

Cacioppo, J. T., Harkins, S. G., & Petty, R. E. (1981). The nature of attitudes and cognitive responses and their relationships to behavior. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 31–54). Hillsdale, NJ: Erlbaum.

Cacioppo, J. T., & Petty, R. E. (1979). The effects of message repetition and position on cognitive responses, recall, and persuasion. *Journal of Personality and Social Psychology*, 37, 97–109.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116–131.

Cacioppo, J. T., & Petty, R. E. (1989). Effects of message repetition on argument processing, recall, and persuasion. *Basic and Applied Social Psychology*, 10, 3–12.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119, 197–253.

Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, 45, 805–818.

Campbell, M. C., & Kirmani, A. (2000). Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent. *Journal of Consumer Research*, 27, 69–83.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right." *Journal of Personality and Social Psychology*, 86, 388–404.

Chaiken, S. (1980). Heuristic versus systematic information processing in the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752–766.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discov-

ery of surprising behavioral implications. *Journal of Personality and Social Psychology*, *69*, 318–328.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*, 472–517.

Clore, G. L., & Parrott, W. G. (1994). Cognitive feelings and meta-cognitive judgments. *European Journal of Social Psychology*, *24*, 101–115.

Coats, S., & Smith, E. R. (1999). Perceptions of gender subtypes: Sensitivity to recent exemplar activation and in-group/out-group differences. *Personality and Social Psychology Bulletin*, *25*, 515–526.

Conner, M., Povey, R., Sparks, P., James, R., & Shepherd, R. (2003). Moderating role of attitudinal ambivalence within the theory of planned behavior. *British Journal of Social Psychology*, *42*, 75–94.

Corneille, O., Monin, B., & Pleyers, G. (2005). Is positivity a cue or a response option? Warm glow vs. evaluative matching in the familiarity for attractive and not-so-attractive faces. *Journal of Experimental Social Psychology*, *41*, 431–437.

Cornoldi, C. F., & Vianello, R. (1992). Meta-cognitive knowledge, learning disorders and mental retardation. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (Vol. 7, pp. 87–134). Greenwich, CT: JAI Press.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 142–150.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

Darley, S. A., & Cooper, J. (1972). Cognitive consequences of forced noncompliance. *Journal of Personality and Social Psychology*, *24*, 321–326.

Davidson, A. R. (1995). From attitudes to actions to attitude change: The effects of amount and accuracy of information. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 315–336). Mahwah, NJ: Erlbaum.

Deffenbacher, K. A. (1984). Experimental psychology actually can assist triers of fact. *American Psychologist*, *39*, 1066–1068.

DeMarree, K. G., Petty, R. E., & Briñol, P. (in press). Self-certainty: Parallels to attitude certainty. *International Journal of Psychology and Psychological Therapy*.

DeSteno, D., Petty, R. E., Wegener, D. T., & Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion specificity. *Journal of Personality and Social Psychology*, *78*, 397–416.

Dijksterhuis, A., Macrae, C. N., & Haddock, G. (1999). When recollective experiences matter: Subjective ease of retrieval and stereotyping. *Personality and Social Psychology Bulletin*, *25*, 760–768.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584.

Dove, N., Wegener, D. T., & Petty, R. E. (2001, February). *Information availability and bias correction strategies*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, San Antonio, TX.

Dove, N., Wegener, D. T., & Petty, R. E. (2003, February). *Information availability, need for cognition, and bias correction strategies*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, Universal City, CA.

Dovido, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, *33*, 545–565.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, *58*, 568–581.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference mea-

sure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326.

Eagly A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Ellis, S. (1996). Selecting information on job content or job context: The moderating role of one's own epistemic authority. *Journal of Applied Social Psychology*, *26*, 1643–1657.

Ellis, S., & Kruglanski, A. W. (1992). Self as an epistemic authority: Effects on experiential and instructional learning. *Social Cognition*, *10*, 357–375.

Fabrigar, L. R., & Petty, R. E. (1999). The role of the affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin*, *25*, 363–381.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences.* (pp. 247–282). Hillsdale, NJ: Erlbaum.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Fazio, R. H., & Zanna, M. P. (1978). Attitudinal qualities relating to the strength of the attitude–behavior relationship. *Journal of Experimental Social Psychology*, *14*, 398–408.

Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. *Advances in Experimental Social Psychology*, *14*, 161–202.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Fleming, M., Wegener, D. T., & Petty, R. E. (1999). Procedural and legal motivations to correct for perceived judicial biases. *Journal of Experimental Social Psychology*, *35*, 186–203.

Förster, J., & Strack, F. (1998). Subjective theories about encoding may influence recognition. *Social Cognition*, *16*, 78–92.

Freitas, A. L., Azizan, A., Travers, S., & Berry, S. A. (2005). The evaluative connotation of processing fluency: Inherently positive or moderated by motivational context. *Journal of Experimental Social Psychology*, *41*, 636–644.

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, *21*, 1–31.

Friestad, M., & Wright, P. (1995). Persuasion knowledge: Lay people's and researches' beliefs about the psychology of persuasion. *Journal of Consumer Research*, *27*, 123–156.

Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality and Social Psychology Bulletin*, *30*, 585–593.

Gawronski, B., & Bodenhausen, G. V. (in press). Accessibility effects on implicit social cognition: The role of knowledge activation and retrieval experiences. *Journal of Personality and Social Psychology*.

Gerard, H. B., & Orive, R. (1987). The dynamics of opinion formation. *Advances in Experimental Social Psychology*, *20*, 171–202.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You cannot believe everything you read. *Journal of Personality and Social Psychology*, *65*, 221–233.

Gill, M. J., Swann, W. B., & Silvera, D. H. (1998). On the genesis of confidence. *Journal of Personality and Social Psychology*, *75*, 1101–1114.

Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair:

Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology, 77*, 669–687.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition, 10*, 597–602.

Grayson, C. E., & Schwarz, N. (1999). Beliefs influence information processing strategies: Declarative and experiential information in risk assessment. *Social Cognition, 17*, 1–18.

Greenwald, A. G. (1968). Cognitive learning, cognitive response persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147–170). New York: Academic Press.

Greenwald, A. G., & Albert, R. D. (1968). Acceptance and recall of improvised arguments. *Journal of Personality and Social Psychology, 8*, 31–34.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24*, 411–435.

Gross, S. R., Holtz, R., & Miller, N. (1995). Attitude certainty. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 215–246). Mahwah, NJ: Erlbaum.

Gruder, C. L., Cook, T. D., Hennigan, K. M., Flay, B. R., Alessis, C., & Halamaj, J. (1978). Empirical tests of the absolute sleeper effect predicted from the discounting cue hypothesis. *Journal of Personality and Social Psychology, 36*, 1061–1074.

Grush, J. E. (1976). Attitude formation and mere exposure phenomena: A nonartificial explanation of empirical findings. *Journal of Personality and Social Psychology, 33*, 281–290.

Haddock, G., Rothman, A. J., Reber, R., & Schwarz, N. (1999). Forming judgments of attitude certainty, intensity, and importance: The role of subjective experiences. *Personality and Social Psychology Bulletin, 25*, 771–782.

Haddock, G., Rothman, A. J., & Schwarz, N. (1996). Are (some) reports of attitude strength context dependent? *Canadian Journal of Behavioral Science, 24*, 313–317.

Harris, R. N., & Snyder, C. R. (1986). The role of uncertain self-esteem in self-handicapping. *Journal of Personality and Social Psychology, 51*, 451–458.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Hastie, R., Landsman, R., & Loftus, E. F. (1978). Eyewitness testimony: The dangers of guessing. *Jurimetrics Journal, 19*, 1–8.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*, 258–268.

Haugtvedt, C. P., & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology, 63*, 308–319.

Heider, F. (1958). *The psychology of interpersonal relations*. New York Wiley.

Hermann, A. D., Leonardelli, G. J., & Arkin, R. M. (2002). Self-doubt and self-esteem: A threat from within. *Personality and Social Psychology Bulletin, 28*, 395–408.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology, 19*, 323–340.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319–340.

Higgins, E. T. (1998a). The aboutness principle: A pervasive influence on inference. *Social Cognition, 16*, 173–198.

Higgins, E. T. (1998b). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology, 46*, 1–46.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist, 55*, 1217–1233.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*, 141–154.

Hilton, J. L., & von Hippel, W. (1990). The role of consistency in the judgment of stereotype-relevant behaviors. *Personality and Social Psychology Bulletin, 16*, 430–448.

Hirt, E. R., Kardes, F. R., & Markman, K. D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *Journal of Experimental Social Psychology, 40*, 374–383.

Holbrook, A. L., Berent, M. K., Krosnick, J. A., Visser, P.S., & Boninger, D. S. (2005). Attitude importance and the accumulation of attitude-relevant knowledge in memory. *Journal of Personality and Social Psychology, 88*, 749–769.

Holland, R., W., Verplanken, B., & van Knippenberg, A. (2003). From repetition to conviction: Attitude accessibility as a determinant of attitude certainty. *Journal of Experimental Social Psychology, 39*, 594–601.

Isbell, L. M., & Wyer, R. S., Jr. (1999). Correcting for mood-induced bias in the evaluation of political candidates: The roles of intrinsic and extrinsic motivation. *Personality and Social Psychology Bulletin, 25*, 237–249.

Jacks, J. Z., & Cameron, K. A. (2003). Strategies of resisting persuasion. *Basic and Applied Social Psychology, 25*, 145–161.

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 21–38.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 57–80). Washington, DC: National Academy Press.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13*, 314–336.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56*, 326–338.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*, 115–125.

James, W. (1890). *The principles of psychology* (Vol. 1). Cambridge, MA: Harvard University Press.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology, 70*, 172–194.

Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin, 116*, 274–292.

Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review, 2*, 137–154.

Kaplan, K. J. (1972). On the ambivalence–indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin, 77*, 361–372.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55*, 893–905.

Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality, and behavior. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 35–59). New York: Academic Press.

Kaufman, D. Q., Stasson, M. F., & Hart, J. W. (1999). Are the tabloids always wrong or it that just what we think? Need for cognition and perceptions of articles in print media. *Journal of Applied Social Psychology, 29*, 1984–1997.

Keller, J., & Bless, H. (2005). When negative expectancies turn into negative performance: The role of ease of retrieval. *Journal of Experimental Social Psychology, 41*, 535–541.

Kelly, A. E., & Kahn, J. H. (1994). Effects of suppression of personal intrusive thoughts. *Journal of Personality and Social Psychology*, *66*, 998–1006.

Klinger, M. R., & Greenwald, A. G. (1994). Preferences need no inferences?: The cognitive basis of unconscious mere exposure effects. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 67–85). San Diego, CA: Academic Press.

Koriat, A. (1993). How do we know that we know?: The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483–502). New York: Guilford Press.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118.

Krosnick, J. A. (1988). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. *Journal of Personality and Social Psychology*, *55*, 196–210.

Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K., & Carnot, C. G. (1993). Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology*, *65*, 1132–1151.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). Mahwah, NJ: Erlbaum.

Kruglanski, A. W. (1980). Lay epistemologic process and contents: Another look at attribution theory. *Psychological Review*, *87*, 70–87.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.

Kruglanski, A. W., Freund, T., & Bar-Tal, D. (1996). Motivational effects in the mere-exposure paradigm. *European Journal of Social Psychology*, *26*, 479–499.

Kruglanski, A. W., Raviv, A., Bar-Tal, D., Raviv, A., Sharvit, K., Bar, R., et al. (2005). Says who? Epistemic authority effects in social judgment. *Advances in Experimental Social Psychology*, *37*, 345–392.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Kruglanski, A. W., & Webster, D. M. (1991). Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of environmental noise. *Journal of Personality and Social Psychology*, *61*, 212–225.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: Seizing and Freezing. *Psychological Review*, *103*, 263–283.

Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin*, *130*, 143–172.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, *207*, 557–558.

Lee, A. Y. (1994). The mere exposure effect: Is it a mere case of misattribution? In C. T. Allen & D. R. John (Eds.), *Advances in consumer research* (Vol. 21, pp. 270–275). Provo, UT: Association for Consumer Research.

Lee, A. Y. (2001). The mere exposure effect: An uncertainty reduction explanation revisited. *Personality and Social Psychology Bulletin*, *27*, 1255–1266.

Lee, A. Y., & Aaker, J. L. (2004). Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion. *Journal of Personality and Social Psychology*, *86*, 205–218.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, *72*, 275–287.

Lepore, L., & Brown, R. (2002). The role of awareness: Divergent automatic stereotype activation and implicit judgment correction. *Social Cognition*, *20*, 321–351.

Liberman, S., & Chaiken, S. (1991). Value conflict and thought-induced attitude change. *Journal of Experimental Social Psychology*, *27*, 203–216.

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task. *Personality and Social Psychology Bulletin*, *13*, 411–429.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*, 808–817.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Wheeler, V. (1996). On resisting the temptation for simplification: Counterintentional effects of stereotype suppression on social memory. *Social Cognition*, *14*, 1–20.

Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology*, *41*, 19–35.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effects of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, *61*, 13–33.

Maio, G. R., Bell, D. E., & Esses, V. M. (1996). Ambivalence and persuasion: The processing of messages about immigrant groups. *Journal of Experimental Social Psychology*, *32*, 513–536.

Mandler, G., Nkamura, Y., & Van Zandt, B. J. (1987). Nonspecific effects of exposure to stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 646–648.

Marecek, J., & Mettee, D. R. (1972). Avoidance of continued success as a function of self-esteem, level of esteem certainty, and responsibility for success. *Journal of Personality and Social Psychology*, *22*, 98–107.

Marks, G., & Miller, N. (1985). The effect of certainty on consensus judgments. *Personality and Social Psychology Bulletin*, *11*, 165–177.

Marsh, H. W. (1993). Relations between global and specific domains of self: The importance of individual importance, certainty, and ideals. *Journal of Personality and Social Psychology*, *65*, 975–992.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, *59*, 27–37.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs. "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*, 433–449.

Mazursky, D., & Schul, Y. (2000). In the aftermath of invalidation: Shaping judgment rules on learning that previous information was invalid. *Journal of Consumer Psychology*, *9*, 213–222.

McGarty, C., Turner, J. C., Oakes, P. J., & Haslam, S. A. (1993). The creation of uncertainty in the influence process: The roles of stimulus information and disagreement with similar others. *European Journal of Social Psychology*, *23*, 17–38.

McGregor, I. (2003). Defensive zeal: Compensatory conviction about attitudes, values, goals, groups, and self-definitions in the face of personal uncertainty. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 73–92). Mahwah, NJ: Erlbaum.

McGuire, W. J., & McGuire, C. V. (1991). The content, structure, and operation of thought systems. In R. S. Wyer, Jr., & T. Srull

(Eds.), *Advances in social cognition* (Vol. 4, pp. 1–78). Hillsdale, NJ: Erlbaum.

Mitchell, K. J., & Johnson, M. K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179–195). New York: Oxford University Press.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, *11*, 462–466.

Monin, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology*, *85*, 1035–1048.

Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review*, *2*, 63–82.

Moreland, R. L., & Zajonc, R. B. (1982). Exposure effects in person perception: Familiarity, similarity, and attraction. *Journal of Experimental Social Psychology*, *18*, 395–415.

Moskowitz, G. B. (2005). *Social cognition: Understanding self and others*. New York: Guilford Press.

Nelson, T. O. (Ed.). (1992). *Meta-cognition: Core readings*. Boston: Allyn & Bacon.

Nelson, T. O., Kruglanski, A. W., & Jost, J. T. (1998). Knowing thyself and others: Progress in metacognitive social psychology. In V. Y. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Meta-cognition: Cognitive and social dimensions* (pp. 69–89). Thousand Oaks, CA: Sage.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 125–173). New York: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. Shimamura (Eds.), *Meta-cognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Neuberg, S. L., West, S. G., Judice, T. N., & Thompson, M. M. (1997). On dimensionality, discriminant validity, and the role of psychometric analyses in personality theory and measurement: Reply to Kruglanski et al.'s (1997) defense of the Need for Closure Scale. *Journal of Personality and Social Psychology*, *73*, 1017–1029.

Newby-Clark, I. R., McGregor, I., & Zanna, M. P. (2002). Thinking and caring about cognitive inconsistency: When and for whom does attitudinal ambivalence feel uncomfortable? *Journal of Personality and Social Psychology*, *82*, 157–166.

Newcomb, T. M. (1968). Interpersonal balance. In R. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 28–51). Chicago: Rand McNally.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.

Norton, G. R. (1975). Measurement of ambiguity tolerance. *Journal of Personality Assessment*, *39*, 607–619.

Norwick, R., & Epley, N. (2003, February). *Experiential determinants of confidence*. Poster presented at the Society for Personality and Social Psychology, Los Angeles.

Oleson, K. C., Poehlmann, K. M., Yost, J. H., Lynch, M. E., & Arkin, R. M. (2000). Subjective overachievement: Individual differences in self-doubt and concern with performance. *Journal of Personality*, *68*, 491–524.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*, 42–55.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, *29*, 261–265.

Ostrom, T. M., & Upshaw, H. S. (1968). Psychological perspective and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 217–242). New York: Academic Press.

Palmer, D. L., & Kalin, R. (1991). Predictive validity of the Dog-

matic Rejection Scale. *Personality and Social Psychology Bulletin*, *17*, 212–218.

Pelham, B. W. (1991). On confidence and consequence: The certainty and importance of self-knowledge. *Journal of Personality and Social Psychology*, *60*, 518–530.

Pelham, B. W., & Swann, W. B., Jr. (1994). The juncture of intrapersonal and interpersonal knowledge: Self-certainty and interpersonal congruence. *Personality and Social Psychology Bulletin*, *20*, 349–357.

Petty, R. E. (2006). A metacognitive model of attitudes. *Journal of Consumer Research*, *33*, 22–24.

Petty, R. E., & Briñol, P. (2002). Attitude change: The Elaboration Likelihood Model of persuasion. In G. Bartels & W. Nielissen (Eds.), *Marketing for sustainability: Towards transactional policy making* (pp. 176–190). Amsterdam: IOS Press.

Petty, R. E., & Briñol, P. (2006). A meta-cognitive approach to "implicit" and "explicit" evaluations: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin*, *132*, 740–744.

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*, *82*, 722–741.

Petty, R. E., & Cacioppo, J. T. (1979). Issue-involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, *37*, 1915–1926.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Petty, R. E., & Cacioppo, J. T. (1990). Involvement and persuasion: Tradition versus integration. *Psychological Bulletin*, *107*, 367–374.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*, 847–855.

Petty, R. E., Cacioppo, J. T., & Haugtvedt, C. (1992). Involvement and persuasion: An appreciative look at the Sherifs' contribution to the study of self-relevance and attitude change. In D. Granberg & G. Sarup (Eds.), *Social judgment and intergroup relations: Essays in honor of Muzafer Sherif* (pp. 147–174). New York: Springer-Verlag.

Petty, R. E., DeMarree, K. G., Briñol, P., & Horcajo, J. (2005). *Chronic accessibility and need for cognition* (Working paper). Columbus: Ohio State University Press.

Petty, R. E., Haugtvedt, C., & Smith, S. M. (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 93–130). Mahwah, NJ: Erlbaum.

Petty, R. E., & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Erlbaum.

Petty, R. E., Ostrom, T. M., & Brock, T. C. (1981). Historical foundations of the cognitive response approach to attitudes and persuasion. In R. Petty, T. Ostrom, & T. Brock (Eds.), *Cognitive responses in persuasion* (pp. 5–29). Hillsdale, NJ: Erlbaum.

Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology*, *90*, 21–41.

Petty, R. E., Tormala, Z. L., & Rucker, D. D. (2004). Resistance to persuasion: An attitude strength perspective. In J. T. Jost, M. R. Banaji, & D. A. Prentice (Eds.), *Perspectivism in social psychology: The yin and yang of scientific progress* (pp. 37–51). Washington, DC: American Psychological Association Press.

Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology*, *29*, 137–165.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 323–390). New York: McGraw-Hill.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y.

Trope (Eds.), *Dual-process theories in social psychology* (pp. 41–72). New York: Guilford Press.

Petty, R. E., Wegener, D. T., & White, P. (1998). Flexible correction processes in social judgment: Implications for persuasion. *Social Cognition*, 16, 93–113.

Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34, 874–884.

Petty, R. E., Wheeler, S. C., & Bizer, G. (2000). Matching effects in persuasion: An elaboration likelihood analysis. In G. Maio & J. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 133–162). Mahwah, NJ: Erlbaum.

Petty, R. E., Wheeler, S. C., & Tormala, Z. L. (2003). Persuasion and attitude change. In T. Millon & M. J. Lerner (Eds.), *Handbook of psychology: Vol. 5. Personality and social psychology* (pp. 353–382). Hoboken, NJ: Wiley.

Pieters, R. G. M., & Verplanken, B. (1995). Intention–behavior consistency: Effects of consideration set size, involvement, and need for cognition. *European Journal of Social Psychology*, 25, 531–543.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75, 811–832.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, 71, 431–449.

Priester, J. R., Wegener, D. T., Petty, R. E., & Fabrigar, L. R. (1999). Examining the psychological processes underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology*, 1, 27–48.

Radecki, C. M., & Jaccard, J. (1995). Perceptions of knowledge, actual knowledge, and information search behavior. *Journal of Experimental Social Psychology*, 31, 107–138.

Raden, D. (1989). Are scores on conventional attitude scales confounded with other measures of attitude strength? Findings from the General Social Survey. *Psychological Reports*, 64, 1247–1252.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8, 364–382.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 45–48.

Rhodewalt, F., & Morf, C. C. (1995). Self and interpersonal correlates of the Narcissistic Personality Inventory: A review and new findings. *Journal of Personality and Social Psychology*, 29, 1–23.

Roese, N. J., & Olson, J. M. (1994). Attitude importance as a function of repeated attitude expression. *Journal of Experimental Social Psychology*, 30, 39–51.

Rothman, A. J., & Hardin, C. (1997). Differential use of the availability heuristic in social judgment. *Personality and Social Psychology Bulletin*, 23, 123–138.

Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, 24, 1053–1064.

Rucker, D. D., & Petty, R. E. (2004). When resistance is futile: Consequences of failed counterarguing for attitude certainty. *Journal of Personality and Social Psychology*, 86, 219–235.

Ruder, M., & Bless, H. (2003). Mood and the reliance on the ease of retrieval heuristic. *Journal of Personality and Social Psychology*, 85, 20–32.

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the big five. *Journal of Research in Personality*, 38, 473–480.

Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology*, 49, 894–903.

Schul, Y., & Goren, H. (1997). When strong evidence has less im-

pact than weak evidence: Bias, adjustment, and instructions to ignore. *Social Cognition*, 15, 133–155.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, 2, 87–99.

Schwarz, N. (2004). Meta-cognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, 14, 332–348.

Schwarz, N., & Bless, H. (1992a). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Schwarz, N., & Bless, H. (1992b). Scandals and the public's trust in politicians: Assimilation and contrast effects. *Personality and Social Psychology Bulletin*, 18, 574–579.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 436–457). Malden, MA: Blackwell.

Schwarz, N., & Clore, G. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Schwarz, N., Strack, F., & Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, 55, 3–23.

Sczesny, S., & Kühnen, U. (2004). Meta-cognition about biological sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence. *Personality and Social Psychology Bulletin*, 30, 13–21.

Sedikides, C. (1995). Central and peripheral self-conceptions are differentially influenced by mood: Tests of the differential sensitivity hypothesis. *Journal of Personality and Social Psychology*, 69, 759–777.

Seta, J. J., Martin, L. L., & Capehart, G. (1979). Effects of contrast and assimilation on the attitude similarity-attraction relationship. *Journal of Personality and Social Psychology*, 37, 462–467.

Setterlund, M. B., & Niedenthal, P. M. (1993). "Who am I? Why am I here?": Self-esteem, self-clarity, and prototype matching. *Journal of Personality and Social Psychology*, 65, 769–780.

Shaw, J. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, 2, 126–146.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31, 713–724.

Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 99–151). Hillsdale, NJ: Erlbaum.

Smith, E. R., Miller, D. A., Maitner, A. T., Crump, S. A., Garcia-Marques, T., & Mackie, D. M. (2006). Familiarity can increase stereotyping. *Journal of Experimental Social Psychology*, 42, 471–478.

Smith, S. M., & Petty, R. E. (1995). Personality moderators of mood congruency effects on cognition: The role of self-esteem and negative mood regulation. *Journal of Personality and Social Psychology*, 68, 1092–1107.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526–537.

Snyder, M., & DeBono, K. G. (1985). Appeals to image and claims about quality: Understanding the psychology of advertising. *Journal of Personality and Social Psychology*, 49, 586–597.

Sorrentino, R. M., Bobocel, D. R., Gitta, M. Z., Olson, J. M., & Hewitt, E. C. (1988). Uncertainty orientation and persuasion: Individual differences in the effects of personal relevance on social judgments. *Journal of Personality and Social Psychology*, *55*, 357–371.

Sorrentino, R. M., & Short, J. C. (1986). Uncertainty orientation, motivation, and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 379–403). New York: Guilford Press.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, *37*, 1660–1672.

Stapel, D. A., Martin, L. L., & Schwarz, N. (1998). The smell of bias: What instigates correction processes in social judgments? *Personality and Social Psychology Bulletin*, *24*, 797–806.

Stone, J. (2003). Self-consistency for low self-esteem in dissonance processes: The role of self-standards. *Personality and Social Psychology Bulletin*, *29*, 846–858.

Story, A. L. (2004). Self-esteem and self-certainty: A mediational analysis. *European Journal of Personality*, *18*, 115–125.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational based strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Förster, J. (1998). Self-reflection and recognition: The role of meta-cognitive knowledge in the attribution of recollective experience. *Personality and Social Psychology Review*, , *2*, 111–123.

Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 579–596). New York: Guilford Press.

Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: The social determinants of information use in judgments of life-satisfaction. *European Journal of Social Psychology*, *18*, 429–442.

Strack, F., & Mussweiler, T. (2001). Resisting influence: Judgmental correction and its goals. In J. Forgas & K. Williams (Eds.), *Social influence processes: Direct and indirect influences* (pp. 199–212). New York: Psychology Press.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Strack, F., Schwarz, N., & Geschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, mood, and mode of thinking. *Journal of Personality and Social Psychology*, *49*, 1460–1469.

Swann, W. B., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, *46*, 1287–1302.

Swann, W. B., Jr., & Gill, M. J. (1997). Confidence and accuracy in person perception: Do we know what we think we know about our relationship partners? *Journal of Personality and Social Psychology*, *73*, 747–757.

Swann, W. B., Pelham, B. W., & Chidester, T. R. (1988). Change through paradox: Using self-verification to alter beliefs. *Journal of Personality and Social Psychology*, *54*, 268–273.

Thompson, W. C., Fong, G. T., & Rosenhan, D. L. (1981). Inadmissible evidence and juror verdicts. *Journal of Personality and Social Psychology*, *40*, 453–463.

Thomsen, C. J., Borgida, E., & Lavine, H. (1995). The causes and consequences of personal involvement. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 191–214). Mahwah, NJ: Erlbaum.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*, 973–988.

Tom, G., Pettersen, P., Lau, T., Burton, T., & Cook, J. (1991). The role of overt head movement in the formation of affect. *Basic and Applied Social Psychology*, *12*, 281–289.

Tormala, Z. L., Briñol, P., & Petty, R. E. (2006). When credibility attacks: The reverse impact of source credibility on persuasion. *Journal of Experimental Social Psychology*, *42*, 684–691.

Tormala, Z. L., Clarkson, J. J., & Petty, R. E. (2006). Resisting persuasion by the skin of one's teeth: The hidden success of resisted persuasive messages. *Journal of Personality and Social Psychology*, *91*(3), 423–435.

Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty. *Journal of Personality and Social Psychology*, *83*, 1298–1313.

Tormala, Z. L., & Petty, R. E. (2004a). Resistance to persuasion and attitude certainty: The moderating role of elaboration. *Personality and Social Psychology Bulletin*, *30*, 1446–1457.

Tormala, Z. L., & Petty, R. E. (2004b). Source credibility and attitude certainty: A metacognitive analysis of resistance to persuasion. *Journal of Consumer Psychology,* *14*, 426–441.

Tormala, Z. L., Petty, R. E., & Briñol, P. (2002). Ease of retrieval effects in persuasion: The roles of elaboration and thought-confidence. *Personality and Social Psychology Bulletin*, *28*, 1700–1712.

Tormala, Z. L., Petty, R. E., & DeSensi, V. L. (in press). Multiple roles for minority sources in persuasion and resistance. In R. Martin & M. Hewstone (Eds.), *Minority influence and innovation: Antecedents, processes, and consequences*. London: Psychology Press.

Tourangeau, R., Rasinski, K. A., Bradburn, N., & D'Andrade, R. (1989). Carryover effects in attitude surveys. *Public Opinion Quarterly*, *53*, 495–524.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

van Harreveld, F., & van der Pligt, J. (2004). Attitudes as stable and transparent constructions. *Journal of Experimental Social Psychology*, *40*, 666–674.

Varhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Visser, P. S., Bizer, G. Y., & Krosnick, J. A. (2006). Exploring the latent structure of strength-related attitude attributes. *Advances in Experimental Social Psychology*, *38*, 1–67.

Visser, P. S., & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology*, *75*, 1388–1409.

Visser, P. S., Krosnick, J. A., & Simmons, J. P. (2003). Distinguishing the cognitive and behavioral consequences of attitude and certainty: A new approach to testing the common-factor hypothesis. *Journal of Experimental Social Psychology*, *39*, 118–141.

Visser, P. S., & Mirabile, R. R. (2004). Attitudes in the social context: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology*, *87*, 779–795.

von Hippel, W., Silver, L. A., & Lynch, M. E. (2000). Stereotyping against your will: The role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin*, *26*, 523–532.

Wänke, M., & Bless, H. (2000). The effects of subjective ease of retrieval on attitudinal judgments: The moderating role of processing motivation. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 143–161). Philadelphia: Taylor & Francis.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin*, *22*, 1105–1113.

Wänke, M., Bohner, G., & Jurkowitsch, A. (1997). There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research*, *24*, 170–177.

Wänke, M., Schwarz, N., & Bless, H. (1995). The availability heuris-

tic revisited: Experienced ease of retrieval in mundane frequency estimates. *Acta Psychologica*, *89*, 83–90.

Weary, G., & Edwards, J. A. (1994). Individual differences in causal uncertainty. *Journal of Personality and Social Psychology*, *67*, 308–318.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*, 1049–1062.

Wedell, D. H., Parducci, A., & Gieselman, R. E. (1987). A formal analysis of ratings of physical attractiveness: Successive contrast and simultaneous assimilation. *Journal of Experimental Social Psychology*, *23*, 230–249.

Wegener, D. T., Clark, J. K., & Petty, R. E. (2006). Not all stereotyping is created equal: Differential consequences of thoughtful versus non-thoughtful stereotyping. *Journal of Personality and Social Psychology*, *90*, 42–59.

Wegener, D. T., Downing, J., Krosnick, J. A., & Petty, R. E. (1995). Measures and manipulations of strength related properties of attitudes: Current practice and future directions. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 455–488). Mahwah, NJ: Erlbaum.

Wegener, D. T., Dunn, M., & Tokusato, D. (2001). The flexible correction model: Phenomenology and the use of naive theories in avoiding or removing bias. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 277–290). Mahwah, NJ: Erlbaum.

Wegener, D. T., Kerr, N., Fleming, M., & Petty, R. E. (2000). Flexible corrections of juror judgments: Implications for jury instructions. *Psychology, Public Policy, and Law*, *6*, 629–654.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). Mahwah, NJ: Erlbaum.

Wegener, D. T., & Petty, R. E. (2001). Understanding effects of mood through the elaboration likelihood and flexible correction models. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 177–210). Mahwah, NJ: Erlbaum.

Wegener, D. T., Petty, R. E., & Dunn, M. (1998). The meta-cognition of bias correction: Naive theories of bias and the flexible correction model. In V. Yzerbyt, G. Lories, & B. Dardenne (Eds.), *Meta-cognition: Cognitive and social dimensions* (pp. 202–227). London: Sage.

Wegener, D. T., Petty, R. E., Smoak, N. D., & Fabrigar, L. R. (2004). Multiple routes to resisting attitude change. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 13–38). Mahwah, NJ: Erlbaum.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Wegner, D. M., & Erber, R. (1992). The hyperaccessibility of suppressed thoughts. *Journal of Personality and Social Psychology*, *63*, 903–912.

Wegner, D. M., Erber, R., & Bowman, R. E. (1993). *On trying not to be sexist*. Unpublished manuscript.

Wegner, D. M., Schneider, D. J., Carter, S., III, & White, L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*, 409–418.

Wells, G. L., & Murray, D. M. (1983). What can psychology say about the *Neil v. Biggers* criteria for judging eyewitness accuracy? *Journal of Applied Psychology*, *68*, 347–362.

Wells, G. L., & Petty, R. E. (1980). The effects of overt head movements on persuasion: Compatibility and incompatibility of responses. *Basic and Applied Social Psychology*, *1*, 219–230.

Werth, L., & Strack, F. (2003). An inferential approach to the knew-it-all-along effect. *Memory*, *11*, 411–419.

Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2005). The roles of the self in priming-to-behavior effects. In A. Tesser, J. V. Wood, & D. A. Stapel (Eds.), *On building, defending and regulating the self: A psychological perspective* (pp. 245–271). New York: Psychology Press.

Whittlesea, B. W. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1235–1253.

Whittlesea, B. W., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, *29*, 716–732.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117–142.

Wilson, T. D., Houston, C. E., & Meyers, J. M. (1998). Choose your poison: Lay beliefs and unintended attitude change. *Social Cognition*, *16*, 114–132.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winkielman, P., & Schwarz, N. (2001). How pleasant was your childhood? Beliefs about memory shape inferences from experienced difficulty of recall. *Psychological Science*, *12*, 176–179.

Witkin, H. A., Lewis, H. G., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. (1954). *Personality through perception*. New York: Harper & Row.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815–827.

Woike, B. A., & Baumgardner, A. H. (1993). Global-specific incongruencies in self-worth and the search for self-knowledge. *Personality and Social Psychology Bulletin*, *19*, 290–295.

Wood, W., Rhodes, N., & Biek, M. (1995). Working knowledge and attitude strength: An information-processing analysis. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 283–313). Mahwah, NJ: Erlbaum.

Wright, P. (1973). The cognitive processes mediating acceptance of advertising. *Journal of Marketing Research*, *10*, 53–62.

Wright, R. (2001). Self-certainty and Self-esteem. In T. J. Owens, S. Stryker, & N. Goodman (Eds.), *Extending self-esteem theory and research: Sociological and psychological currents* (pp. 101–134). New York: Cambridge University Press.

Wyer, R. S., & Budesheim, T. L. (1987). Person memory and judgments: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology*, *53*, 14–29.

Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin*, *26*, 13–25.

Yzerbyt, V. Y., Leyens, J.-P., & Corneille, O. (1998). Social judgeability and the bogus pipeline: The role of naive theories of judgment in impression formation. *Social Cognition*, *16*, 56–77.

Yzerbyt, V. Y., Lories, G., & Dardenne, B. (1998). *Meta-cognition: Cognitive and social dimensions*. Thousand Oaks, CA: Sage.

Yzerbyt, V. Y., Schadron, G., Leyens, J.-P., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, *66*, 48–55.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.

Ziegler, R., Diehl, M., & Ruther, A. (2002). Multiple source characteristics and persuasion: Source inconsistency as a determinant of message scrutiny. *Personality and Social Psychology Bulleting*, *28*, 496–508.

Zimbardo, P. G. (1960). Involvement and communication discrepancy as determinants of opinion conformity. *Journal of Abnormal and Social Psychology*, *60*, 86–94.

CHAPTER 12

Principles of Mental Representation

ROBERT S. WYER, JR.

Judgments and decisions are rarely made in a cognitive vacuum. They are influenced not only by the information we receive in the immediate situation at hand but also by knowledge we have acquired in the past. However, people rarely use all the knowledge they have acquired that is potentially relevant to a judgment they wish to make. Rather, they apply only a subset of this knowledge that comes to mind most quickly and easily (cf. Chaiken, Liberman, & Eagly, 1989; Simon, 1957; Taylor & Fiske, 1978). Many factors can influence the accessibility of concepts and knowledge in memory (for a review, see Higgins, 1996a). A major determinant, of accessibility, however, is the manner in which judgment-relevant knowledge is organized in memory and the associations that have been formed between its features and those of the object being judged.

This chapter reviews the general principles that govern the construction and use of mental representations. Carlston and Smith (1996) described many such principles in the first edition of this book (see also Smith, 1998; Wyer & Carlston, 1994; Wyer & Srull, 1989). This chapter is distinguished from earlier ones in its emphasis on the representation of knowledge of the sort we acquire through experiences outside the laboratory. In particular, it focuses on the representation of familiar activities of the sort we experience as a normal part of daily life, and on the implicit theories we derive from these experiences. I also consider the representation of larger bodies of propositional knowledge that people acquire about the world in general. Before entering into this discussion, however, it is necessary to consider whether a conceptualization of mental representations is at all important in accounting for social psychological phenomena. Is this particular cognitive trip really necessary?

GENERAL CONSIDERATIONS

Theories of mental representation are inherently metaphorical. They must consequently be evaluated on the basis of their utility and not their validity in describing the physiology of the brain. Moreover, it is often difficult to distinguish empirically between the effects of representational processes that occur at the time knowledge is acquired and the effects of retrieval processes that occur at the time it is used. John Anderson (1978) noted that for every theoretical account of memory phenomena in terms of different memory representations and a single retrieval process, a mathematically equivalent account can usually be constructed in terms of a single memory representation and different retrieval processes. Nevertheless, it is difficult to account for many social phenomena without recourse to assumptions about representational processes. For one thing, there is often little correlation between the judgments of an object and the content and implications of information about the object that one can recall (Dreben, Fiske, & Hastie, 1979; Greenwald, 1968; Lichtenstein & Srull, 1987). In some instances, the relation may even be negative (Dreben et al., 1979; Lichtenstein & Srull, 1987). Finally, the type of information recalled is highly contingent upon recipients' goals at the time they receive the information, inde-

pendently of their goals at the time of retrieval (Hamilton, Katz, & Leirer, 1980; Srull, 1981).

The problem noted by Anderson (1978) would be less serious if there were widespread agreement concerning the nature of retrieval processes. In fact, however, several alternative theories of retrieval have been proposed. To provide a framework for discussing the principles that govern the construction and use of mental representations, it is useful to outline briefly the theoretical formulations that have been used to describe memory organization and retrieval and the type of knowledge to which these formulations are normally applied.

THEORIES OF MEMORY RETRIEVAL

Associative Models of Memory

Two classes of theories assume that cognitive elements are organized in memory in the form of an associative network. These elements, which can be either single concepts or more complex bodies of knowledge are represented by nodes, and their associations are represented by pathways connecting them. The identification of the individual elements contained in these representations could proceed either in parallel or sequentially.

Parallel Processes

Parallel search processes are commonly exemplified by spreading activation models of memory (cf. Collins & Loftus, 1975; Higgins, Bargh, & Lombardi, 1985; Wyer & Carlston, 1979). According to these models, the activation of a particular node in the representation (presumably as a result of thinking about the cognitive element to which it refers) elicits excitation that spreads along the pathways connecting it to other nodes. The excitation accumulates at the peripheral nodes, and when the excitation at a node reaches a particular threshold, the element located at the node also becomes activated (thought about). Excitation then spreads from this node to others, and so on. Pathways that reflect stronger associations are assumed to be shorter or easier to traverse. Consequently, elements that are strongly associated with the element one happens to be thinking about are likely to be identified first.

Once an element is deactivated (no longer thought about), the excitation at the node corresponding to it gradually dissipates. However, as long as some residual excitation exists, the element can be more easily reactivated by new excitation that impinges on it from other sources. These and other assumptions permit this model to account for numerous phenomena, including the effects of frequency and recency of thinking about a concept or knowledge unit on its accessibility in memory and, therefore, on the likelihood of identifying and using it in a new situation to which it is potentially applicable (Higgins, 1996a; Higgins et al., 1985; Srull & Wyer, 1979; but see Liberman & Förster, 2000, for a qualification on the effects of actually using a concept on its subsequent accessibility).

Sequential Search Processes

Other conceptualizations of memory retrieval (cf. Anderson & Bower, 1973; Hastie & Kumar, 1979; Srull & Wyer, 1989) assume that search proceeds successively. That is, a search that begins at one node proceeds to a second node along the pathway that connects them, from this node to a third that is connected to the second, and so on. When two or more paths are connected to a given node, the path that reflects the strongest association is selected. If a "dead end" is reached (i.e., an element is identified to which no new paths are connected), the search is reinitiated. Each time a path is traversed, the association between the two elements it connects becomes stronger, and so the likelihood that it is traversed again is increased.

These two versions of an associative network are typically applied to different types of phenomena. The parallel-processing version is particularly useful in conceptualizing the spontaneous activation of knowledge that occurs in the absence of any particular goal that requires it. The sequential-processing version more easily characterizes a deliberate, goal-directed search of memory in pursuit of a specified objective

The cognitive elements that compose an associative network model include single noun and attribute concepts (e.g., Collins & Loftus, 1975) or configurations of features that are stored in memory as a whole (Bower, 1981; Wyer & Carlston, 1979). In one conceptualization of social memory (Carlston, 1994), entire systems of concepts (cognitive, behavioral, visual, affective, etc.) are assumed to be interconnected by associative pathways.

Retrieval Cue Models

A second class of memory models makes few assumptions about the organization of information in memory. These models assume that when information is required for a particular purpose, a set of features is compiled that are relevant to this objective and these features serve as retrieval cues for concepts and knowledge that contain them. The cognitive mechanisms that characterize these retrieval processes, and the knowledge representations to which they are applied, can vary.

Resonance Models

Ratcliff (1978) conceptualized memory search in terms of a tuning-fork metaphor. That is, activating a set of retrieval cues causes these cues to "vibrate" and other units of knowledge that contain these cues then resonate with it. The degree of resonance increases with the similarity of the retrieval cues to those contained in the representation. Therefore, knowledge units whose features are more similar to those of the probe set are activated more quickly. However, the conceptualization makes no specific assumptions about the organization of these knowledge units in memory, each of which is stored independently of one another. (For recent applications of the conceptualization in theories of social memory and judgment, see Luna, 2005; Wyer, 2004.)

Exemplar (Independent-Trace) Models

Hintzman (1986) assumed that each stimulus configuration one encounters leaves its own memory trace that is stored independently of others. When information is sought for a particular purpose, a set of retrieval cues relevant to this objective is compiled, and all memory representations that contain these features are simultaneously activated. This aggregation is experienced as a composite array in which features that are common to the greatest number of representations are weighted most heavily. A new, more abstract representation may sometimes be formed on the basis of these common features, and this representation can then itself be stored in memory. Thus, the knowledge units that are stored in memory can refer not only to specific experiences but also to abstract configurations that result from the cognitive operations that are performed on them.

Distributed Memory Models

Most memory models assume that long-term memory has infinite capacity. Thus, differences in the retrieval of information reflect differences in its accessibility and not its availability. The spontaneous recall of a childhood experience that one has not thought of for nearly 50 years testifies to the apparently validity of the assumption that nothing is ever really forgotten. Nevertheless, the assumption that each experience has its own memory trace seems intuitively implausible. One set of conceptualizations that avoids this assumption is exemplified by parallel distributed processing, or connectionist theories of memory (Humphreys & Kashima, 2002; Smith & deCoster, 1998, 1999). These conceptualizations assume that units of knowledge are not stored in specific memory locations but rather as an array of values along a finite number of discrete dimensions. Suppose for simplicity that 10 dimensions can each take on 10 different values. Then, a total of 10^{10} different patterns can potentially be distinguished. A given experience with values along these dimensions might be represented by a unique combination of these values. To this extent, an enormous amount of information can potentially be stored in a small amount of memory space. A new stimulus may activate an array of these values along with values along other dimensions that have been associated with this array in the past, leading the stimulus to be "recognized" and other associated features to be spontaneously inferred. Although this conceptualization is based on quite different assumptions, many of its implications are similar to those of the independent-trace model proposed by Hintzman (1986).

Retrieval from "Very-Long-Term" Memory

A conceptualization of memory search by Norman and Bobrow (1979) is particularly useful in conceptualizing many memory phenomena that occur outside the laboratory. Suppose a person wishes to recall the names of his high school classmates. To do this, the person might first compile a general descriptor (e.g., "high school") that serves as a retrieval cue for representations in memory that contain its features. These representations may contain the names of some classmates but not all of them. Therefore, a new descriptor might be compiled ("football team," "rock band," "graduation dance," etc.) and the representations retrieved by this descriptor may contain additional names. Descriptors can be at different levels of specificity (e.g., "rock band" or "playing a gig at Mary's New Year's Eve party"). The set of representations that is identified by a general descriptor does not necessarily include those that are identified by a more specific one (Wyer & Srull, 1989).

This conceptualization can easily account for memory phenomena that other conceptualizations find difficult. For instance, it can explain how a chance juxtaposition of cues that have not recently been thought about in combination (e.g., "Mary" and "New Year's Eve parties") can spontaneously stimulate the recall of an experience (Mary's New Year's Eve party) that had not been thought about in years. On the other hand, it can also account for tip-of-the-tongue phenomena (for an elaboration, see Wyer & Srull, 1989).

Reconstructive Memory Models

People do not encode into memory all the features of an experience at the time they encounter it. When they are later asked to recall the experience, therefore, they might often need to fill in gaps by inferring elements that were not contained in the memory representation they formed at the time the experience occurred. This could be done on the basis of other experiences they had in similar situations. Thus, a person who remembers eating a Big Mac during a particular visit to McDonald's might infer that he paid the bill, based on previously acquired knowledge that this event is a normal occurrence in the type of situation at hand. The mental representations that result from these reconstructive processes are presumably stored in memory and can later be retrieved and used as bases for describing the experience at a later point in time. In this case, features that have been added in the course of reconstructing the experience may function in much the same way as the features that had originally been stored.

Hybrid Models

A model suggested by Wyer and Srull (1989) combines features of both a network model and a retrieval cue conceptualization. According to this model, long-term memory comprises a set of "referent bins," each containing knowledge pertaining to a particular referent. Each bin is designated by a header, or set of features, that is strongly associated with the referent and serves to circumscribe its contents. The units of knowledge within a bin can theoretically be of many types and can be coded in difference sense modalities. Thus, they can include visual representations of the referent, trait descriptions or trait–behavior configurations, sequences of observed events, or more general characterizations of the referent

in certain circumscribed types of situations. These knowledge units are theoretically stored in a given bin independently of one another, in the order in which they are constructed.

If information about a referent is necessary for a particular purpose, a set of retrieval cues is compiled that is relevant to this objective and a bin is identified whose header contains these cues. A probabilistic top-down search is then performed for a particular knowledge unit that can be used to attain the objective, and a copy of the representation is retrieved and placed in working memory. After it has been used, this copy (along with any new referent-related new representation that is formed) is returned to the top of the bin from which it was drawn.

Thus, the identification of a goal-relevant body of knowledge is a function of processes similar to those assumed by a retrieval cue model. However, the identification of a specific knowledge unit within this domain can depend on fortuitous past experiences that have led the unit to be represented frequently and, therefore, to be located near the top of the bin. Finally, the search of features *within* a given knowledge unit depends on the nature of the knowledge unit itself but is often assumed to occur in the manner implied by a sequential search, associative network model.

The bin metaphor is useful in conceptualizing many of the phenomena of concern in social psychological research. At the same time, it has difficulty in accounting for the spontaneous comprehension of information about familiar people and events of the sort one encounters outside the laboratory (cf. Wyer, 2004; Wyer & Radvansky, 1999). Its primary advantage may be its eclecticism regarding the different types of mental representations that are formed of social information. In the next section, we first summarize briefly the general types of knowledge representations that are likely to exist in memory, focusing in particular on those that are constructed outside the laboratory. After doing so, we propose a number of principles that govern their formation and use.

TYPES OF MENTAL REPRESENTATION

The types of mental representations that exist in memory reflect the types of knowledge one acquires. Perhaps the most fundamental distinction to be made is between *declarative* knowledge and *procedural* knowledge. Declarative knowledge concerns the referents of everyday life experiences (persons, objects, events, social issues, and relations among these entities). Procedural knowledge refers to the sequence of actions that one performs in pursuit of a particular objective (driving a car, using a word processor, making an inference on the basis of a set of relevant facts, etc.). The two types of knowledge are interrelated. For one thing, declarative knowledge can provide the material on which procedural knowledge operates. Moreover, procedures (e.g., the steps involved in changing a flat tire or obtaining a meal at a restaurant) can be stored as part of declarative knowledge and con-

sulted in deciding how to attain the goal to which they pertain.

On the other hand, this consultation is often not performed. John Anderson (1982, 1983; see also Smith, 1984, 1990, 1994) conceptualizes procedural knowledge as a set of cognitive "productions" in the form of "if [X], then [Y]" rules, where [X] is a configuration of perceptual or cognitive features and [Y] is a sequence of cognitive or motor acts that are elicited automatically when the precondition [X] is met. These productions, which are acquired through learning, may be performed with little conscious cognitive mediation.

A large proportion of social behavior is unquestionably mediated by cognitive productions similar to those that Anderson postulates. As Smith (1990) points out, many phenomena that have typically been attributed to declarative knowledge processing could be governed by procedural knowledge instead. Bargh's (1997) groundbreaking demonstrations of situational influences on overt behavior that occur in the absence of conscious cognitive activity may also be manifestations of the role of procedural knowledge.

Be that as it may, the bulk of research and theory in social psychology has focused on declarative knowledge representations. The work reviewed in this chapter reflects this focus. The representations that have been postulated to exist can vary along many dimensions. They can be specific to a particular instance or can pertain to people and events in general. They can be coded verbally, visually (cf. Kosslyn, 1980; Shepard & Metzler, 1971), or in other sense modalities. They may be based on personal experience or on verbal descriptions of another's experience. Finally, their features may either be temporally or spatially constrained, as in a human face or a series of causally related events, or have no particular relationship except for their common association with their referent. In the latter regard, Mandler (1979, 1984) distinguishes between *categorical* representations, whose features have no a priori relatedness, and *schematic* representations, whose features are organized according to specifiable spatial, causal, or temporal rules. Further distinctions have been made between episodic and semantic knowledge representations (Tulving, 1983), between exemplars and prototypes (Colcombe & Wyer, 2002; but see Barsalou, 1990), and between situation models and generalized representations (Wyer, 2004). The latter distinction is of particular utility in the present context.

Situation Models

Johnson-Laird (1983, 1989; see also Kintsch, 1998) postulates that people construct mental simulations, or models, of the world in which they live. These mental models can pertain to places (e.g., the configuration of rooms in one's house or the United States) or procedures (e.g., how to change a tire). A specific type of mental simulation, proposed by Radvansky and Zacks (1991; see also Zwaan & Radvansky, 1998), is a *situation model*, which pertains to one or more events or states of affairs that are situationally and temporarily constrained (i.e., that occur

in a specifiable, although not necessarily specified, time and place). For example, the representations of "the book is on the table" and "a boy kicked the football" might constitute situation models of conditions that exist at a particular time and place. In contrast, "the boy owns the football" and "the book is boring" do not. Wyer and Radvansky (1999; Wyer, 2004) expanded the construct of a situation model to include *episode models*, or entire sequences of temporally ordered events that are thematically related.

Many features of a situation model are often coded nonverbally, in a mental image that conveys their spatial and temporal relations. Thus, the mental simulation of "the boy kicked the football" requires a mental image of a boy engaged in this activity. If the event is directly experienced or observed, the model of it is likely to be coded in a form that resembles the way it was experienced. Correspondingly, if the event is described orally or in writing, it may be coded metalinguistically. In this case, however, a mental image may be formed spontaneously as well (cf. Wyer & Radvansky, 1999).[1]

The role of situation models has been examined in research on not only language comprehension (Kintsch, 1998; Zwaan & Radvansky, 1998) but also spatial memory (Taylor & Tversky, 1992) and observational learning (Magliano, Dijkstra, & Zwaan, 1996). The construct of an episode model is of particular utility in conceptualizing the mental representation of everyday life experiences or, for that matter, descriptions of others' experiences that one hears or reads about (see, e.g., Wyer, 2004). In this regard, Schank and Abelson (1995) assert that all important social knowledge is stored in memory in the form of stories. Although this claim may be overly zealous (Brewer, 1995; Hastie & Pennington, 1995), it is perhaps not too far off the mark.

Generalized Representations

Social psychological research and theory have focused largely on generalized representations of knowledge that are not directly tied to a specific experience. Unlike situation models, which are formed spontaneously in the course of comprehension, generalized representations are typically constructed in the course of deliberative goal-directed processing (e.g., to form an impression, to make a behavioral decision, or simply to remember the information well enough to communicate about it to others). Some representations can be in the form of propositions that are learned independently of the specific instances to which they apply. Thus, we might learn that honesty is the best policy, or that cigarette smoking is bad for the health, without any direct knowledge of an event that exemplifies these propositions. In other cases, we might abstract commonalities from a number of instances in order to infer a person's attributes on the basis of observations of the individual's behaviors or derive a general conclusion on the basis of propositions that logically imply it. Finally, we might describe an experience in abstract terms in the course of telling others about it. Two types of generalized representations of particular importance pertain to entities and events.

Entity Representations

Generalized entity representations refer either to specific persons, objects, or places or to more general groups or categories. Each representation may contain alternative names of its referent along with a set of features that characterize it. Thus, the representation of a particular person, John Smith, might be characterized by the names "John," "Smitty," and "Mr. Smith," abstract trait descriptors ("honest," "hostile," "intelligent," etc.); and general behaviors that the individual commonly performs ("reads poetry," "swears at slow drivers," etc.). These features could compose an associative network in which the concept is denoted by a central node and the features by peripheral nodes that are associatively linked to it. An attribute (e.g., trait) concept might be represented similarly. The most common type of generalized entity representation investigated in social psychology is a *person stereotype*, consisting of a set of traits and prototypical behaviors that are associated with a general category of individuals (African Americans, women, lawyers, etc.)

Although the features that compose an entity representation are all associated with a concept of their referent, they do not necessarily have any direct relationship to one another. In the example, it makes as much sense to describe John Smith as someone who is "honest, intelligent and shouts and slow drivers" as it does to describe him as someone who "shouts at slow drivers, intelligent, and honest." Thus, in Mandler's (1979, 1984) terms, the representation is typically *categorical*.

Event Representations

A generalized event representation is composed of a sequence of temporally and thematically related events and states of affairs. It is distinguished from an episode model in that the states and events that compose it are coded metalinguistically and may not occur at specific points in time. On the other hand, the individual events that compose the representation are temporally or causally related. Such representations are therefore *schematic* rather than categorical (Mandler, 1979, 1984).

Several types of generalized event representations can be distinguished. *Scripts* are prototypical sequences of events that occur routinely in a particular situational context (Schank & Abelson, 1977). *Histories*, on the other hand, characterize unique sequences of temporally related events that occur over a period of time. They can refer either to events in one's own or another's life or to real-world events involving several different individuals. The time periods to which histories pertain often overlap. Nevertheless, they are presumably stored in memory independently.

IMPLICIT THEORIES

A sequence of temporally related events and states may often compose an implicit theory about the causes and consequences of certain types of actions. A theory can apply to a particular individual (e.g., oneself), a group

(e.g., sorority member) or people in general and might concern either a particular type of situation (e.g., dating, employer-employee relations, a parent–child interactions) or a more general one. The events and states that compose a theory are not localized in time. Instances of a theory can potentially occur in many situations. Consequently, the theory can be applied in comprehending new experiences and construing their implications. It might also be used as a guide in making behavioral decisions.

A theory can be conceptualized as consisting of two or more temporally related propositions, each depicting a different event or state of affairs. In combination, they can exemplify a general principle or generalization. For example, the theory that people get what they deserve might be composed of the segments "P is good (bad)" and "Good (bad) things befall P." A more detailed theory about romantic relationships might consist of the sequence: "P and O meet" P and O fall in love; P and O have a misunderstanding; P and O fight and break up; P and O resolve the misunderstanding; P and O make up and live happily ever after." This more detailed representation might be equivalent to a *story skeleton* of the sort postulated by Schank and Abelson (1995) to guide the interpretation of specific stories that people tell.

THE CONSTRUCTION
OF MENTAL REPRESENTATIONS:
GENERAL CONSIDERATIONS

Current conceptualizations of the mental representation of knowledge have their roots in two more traditional bodies of research and theory. One avenue of inquiry was focused on general characteristics of cognitive organization, such as the way in which the concepts that compose one's cognitive system are differentiated and interrelated (e.g., Harvey, Hunt, & Schroder, 1961; Rokeach, 1960). These characteristics were typically assumed to depend on the personality of the cognizer as well as the domain of knowledge to which the concepts pertained. Numerous measures were proposed to assess individual differences in these structural characteristics (Harvey, Reich, & Wyer, 1968; Scott, 1963, 1969). Alternative measures of the characteristics were often uncorrelated (Wyer, 1964) and their implications for behavior and judgments were not firmly established. Related measures, however, have been used successfully to assess the differentiation and integration of domain-specific sets of cognitions (Linville, 1982; Suedfeld, Tetlock, & Streufert, 1992; Tetlock, 1992). As Gruenfeld (1995) has shown, however, differences in these characteristics can often reflect the effects of communication goals that exist in the social situation in which they are assessed rather than differences in the structure of cognitions on which the communications are based.

The second avenue of early research and theorizing was generally guided by the assumption that people organize their beliefs and attitudes in a manner they consider to be consistent (Abelson et al., 1968; for an analyses of the assumptions underlying representative theories, see Wyer, 1974). Specific theories varied in their assumptions about the criteria for consistency, which in some cases were logical (e.g., Festinger, 1957; McGuire, 1960) and in other cases were hedonic or affective (Heider, 1958; Osgood & Tannenbaum, 1955; Rosenberg, 1956). The theories also placed different emphases on (1) deliberate attempts to maintain cognitive consistency in the face of conflicting information (cf. Abelson, 1959; Festinger, 1957; Lerner, 1970) and (2) the spontaneous formation of consistent sets of cognitions in the course of comprehension (Heider, 1958; McGuire, 1960).

Consistency theories nevertheless agreed in several respects. First, people are motivated to understand the world in which they live and to predict the consequences of their own and others' actions. To attain this objective, they develop a general set of rules they can use to comprehend the information they receive. Furthermore, when this information is incomplete, they often use these rules to infer unmentioned characteristics of the objects or events to which it pertains. Furthermore, if the information conflicts with their prior knowledge about its referents, people may deliberately modify their cognitions about the referents in a way that reestablishes consistency with these cognitions. The implications of cognitive consistency theories are implicit in the principles that have more recently been proposed to govern the construction and use of mental representations. This will become clear in the discussion to follow.

Our discussion begins with two general principles. The first simply recognizes that representations can be formed at two stages of information processing:

> **Principle 1.** Mental representations are formed either (1) spontaneously, in the course of comprehending new information, or (2) deliberately, in the course of attaining a specific objective to which this information is relevant.

Because different processes occur at each stage of cognitive activity, the representations that are formed can also differ. In each case, it is important to note that the construction of a mental representation is usually not itself a goal of information processing. Rather, representations are typically by-products of processing that occur for other reasons. Nevertheless, principle 2 applies.

> **Principle 2 (Parsimony).** The representations that are constructed of information and stored in memory are no more complex than necessary to attain the processing objective to which they are relevant.

Likewise, the processing that leads a representation to be constructed is itself no more extensive than is necessary to attain the goal at hand. Thus, for example, people often base their comprehension of information on the concepts and knowledge that come most easily to mind, without searching for alternative criteria that are equally or more applicable (Higgins, 1996a). Taylor and Fiske (1978) and Chaiken (1987; Chaiken et al., 1989) make a similar assumption in conceptualizing judgment processes. Although individ-

ual differences may exist in the motivation to think extensively about the implications of information (cf. Cacioppo & Petty, 1982), the role of parsimony is pervasive. In fact, several other principles we propose are more specific versions of principle 2.

We first consider the processes that occur at the comprehension stage. Then, we turn to those that occur in the course of more specific goal-direct processing.

COMPREHENSION-BASED MENTAL REPRESENTATIONS

The Role of Mental Imagery in Comprehending Social Information

Visual Imagery in the Comprehension of Verbal Information

To comprehend events of the sort they observe or read about, people often construct a mental simulation of these events that has nonverbal components. Groundbreaking studies by Bransford and his colleagues (Bransford, Barclay, & Franks, 1972; Bransford & Johnson, 1972) are illustrative. For example, apparently anomalous sentences (e.g., "The haystack was important because the cloth would rip") that are normally hard to remember are easily recalled when they are preceded by a single word ("parachute"). The added word presumably activates a mental image of an event that gives the statement meaning (e.g., a parachutist landing in the hay rather than in a tree), and this image facilitates its later recall. Similarly, ostensibly anomalous paragraphs can be given meaning, and thus are recalled better, when they are accompanied by a picture or, in some cases, a title, that facilitates the formation of a mental image of the events being described (Bransford & Johnson, 1972; Bransford & Stein, 1984).

Although the construct of a visual image is metaphorical,[2] several more recent studies of text comprehension (e.g., Black, Turner, & Bower, 1979; Garnham, 1981; Glenberg, Meyer, & Lindem, 1987) provide evidence of its utility. Glenberg and colleagues (1987), for example, asked participants to read a description of a man's activities that began with either "John put on his sweatshirt and went jogging . . . " or "John took off his sweatshirt and went jogging. . . . " The sweatshirt was never again mentioned in either story. Nevertheless, participants identified the word more quickly in a later recognition memory task in the first condition than in the second. In the first condition, a mental image of John wearing the sweatshirt apparently persisted throughout the story, and so the feature was more accessible in memory at the time the recognition task was performed.

Garnham (1981) found that people found it easier to distinguish between "The hostess bought a mink coat from the furrier" and "The hostess bought a mink coat at the furrier's" than to distinguish between the structurally identical statements, "The hostess received a telegram from the furrier" and "The hostess received a telegram at the furrier's." This is presumably because the mental images elicited by the first two statements are very similar, whereas the images elicited by the last two statements are quite different.

Although these examples pertain only to visual images, images in other sense modalities could also be involved. To this extent, the following two principles are tenable:

Principle 3. When people directly observe an event, the representation they form in the course of comprehending it is typically coded in the same modalities in which the event is experienced.

Principle 4. When people receive a verbal description of an event, the representation they form in the course of comprehending it consists in part of a nonverbal image of the event as well as a metalinguistic coding of it.

Observations versus Verbal Descriptions

Although the mental simulations that people form from verbal descriptions of a situation are often analogous to those they would form from direct observation, they are not identical. Most obviously, the visual images that are formed from written descriptions may be less detailed (see note 1). On the other hand, several features that are not specified in the verbal description of an event may often need to be added in order to construct a simulation of it. Barsalou (1993) postulates the existence of "perceptual symbols" that people can use to construct mental pictures of objects and events they have never observed. Thus, we can comprehend the statement "a tiger walked into the classroom and sat down to take notes" without ever seeing such an event, based on previously formed symbols of the elements to which the statement refers. Similarly, we can imagine an acquaintance engaged in activities that we have never observed by applying a perceptual symbol that we formed of the individual in other situations. (For a theoretical explication of these processes, see Wyer, 2004.)

Although the representations formed from verbal descriptions can differ in content from those that are formed through observation, they are normally similar in structure. Verbal descriptions often refer to discrete events (e.g., "John reached for an hors d'oeuvre. Someone bumped his arm, and he spilled his Bloody Mary over Mary's white dress."). In contrast, an observation of these events consists of a continuous stream of activity. It is unlikely, however, that this continuous activity is stored in memory. Newtson (1973, 1976) postulated that people who observe a continuous stream of behavior extracted static frames that denote the transitions ("breakpoints") between one meaningful act and the next, and that these frames, stored in memory, permit the action sequence to be reconstructed. Ebbesen (1980), on the other hand, proposed that a single frame pertaining to each meaningful action is extracted that is typical of the action as a whole (e.g., a static picture of John in the act of reaching for an hors d'oeuvre). To this latter extent, the nonverbal frames that compose the representation of an observed sequence of behaviors might correspond to the images that are formed from verbal descriptions of these behaviors.

An important difference between the representations is nevertheless implied by principles 3 and 4. That is, nonverbal descriptions of events may not be spontaneously coded linguistically, as this coding is not necessary in order to construct a mental simulation of it (principle 3). In contrast, verbal descriptions of events and states are likely to be coded into memory linguistically as well as nonverbally (principle 4). This is partly because the referents of the verbal description must be identified before a mental simulation can be formed.

Wyer, Adaval, and Colcombe (2002; see also Wyer, 2004) evaluated implications of this coding asymmetry. Participants who had been shown pictures of a sequence of temporally related events (e.g., a day in school) took longer to verify the occurrence of these events when they were described verbally than when they were portrayed in similar (but not identical) pictures. This suggests that the pictures were not spontaneously coded verbally at the time they were presented, and so a verbal recoding was necessary at the time of judgment in order to verify them. In contrast, participants who had initially received verbal descriptions of the events verified them just as quickly on the basis of a picture as they did on the basis of a semantically equivalent verbal description. In this case, a mental image of the events had apparently been formed at the time the verbal descriptions of them were presented, and so a recoding at the time of judgment was not required.

Constraints on the Construction of Nonverbal Representations

The representations that are formed in the course of comprehending situation-specific states and events are equivalent to situation models of the sort described earlier in this chapter. However, a question arises concerning the number of such representations that are constructed from a given body of information. The answer to this question is implied by the general rule of parsimony (principle 2). In the present context, a more specific principle is suggested.

> **Principle 5.** People will construct no more mental representations of information than are necessary in order to comprehend it.

Radvansky and his colleagues (Radvansky, Spieler, & Zacks, 1993; Radvansky & Zacks, 1991) confirmed this principle. Radvansky and Zacks (1991) asked participants to learn sets of one, two, or three sentences. In some cases, the sentences described different objects in the same location ("The book is on the table"; "The cup is on the table," etc.) In other cases, they described the same object in different locations ("The book is on the table"; "The book is on the chair," etc.). It is easy to imagine several things in a single location. However, it is impossible to imagine the same thing in several different locations simultaneously. Therefore, according to principle 5, people should construct a single mental representation in the course of comprehending the first set of sentences but different representations when comprehending the second set.

Radvansky and Zacks (1991) reasoned that if people construct a different representation of each statement, the presence of one representation is likely to interfere with the identification of others. Consequently, the time to verify a given statement should increase with the number of statements presented (i.e., the number of representations that were formed and, therefore, the amount of inference). This was true in single-object, multiple-location conditions. In multiple-object, single-location conditions, however, participants are able to form a single representation from the statements in combination, and so the interference in retrieving the representation should be minimal. This was also the case; the statements in this condition were verified quickly regardless of the number presented.

A second constraint on the construction of situation models should also be noted. Mental images are necessarily restricted to events and states that occur at a specific time and place. If the situations described by verbal information are not so constrained, an image (and, therefore, a situation model) is unlikely to be constructed. That is:

> **Principle 6.** When people encounter verbal descriptions of events or states of affairs that are not constrained to a specific time and place, they comprehend these descriptions in terms of semantic concepts alone, and do not form mental images of them.

Using a similar procedure, Radvansky, Wyer, Curiel, and Lutz (1997) demonstrated the applicability of principle 6 to the representation of events. Some participants received information that a person bought either one, two, or three objects that could all be found in a drugstore (i.e., a toothbrush, shaving cream, and aspirin). Others received information that the person bought one, two, or three objects that were found in different stores (e.g., a toothbrush, a diamond ring, and a bicycle). Response times in the multiple-location condition increased with the number of items presented, whereas recognition times in the single-location condition were independent of the number of statements presented.

In other conditions, however, the statements pertained to ownership rather than purchase behavior. Thus, they described states of affairs that were not temporally constrained, and so mental images were unlikely to be formed in the course of comprehending them. Rather, the statements are likely to be coded into memory propositionally, perhaps forming a generalized entity representation of the sort described earlier. Consistent with this assumption, the time to verify a statement about ownership increased with the number of other statements presented regardless of relatedness of the objects described.

A further determinant of whether events are integrated into a single representation is their thematic relatedness. People undoubtedly do not construct a single mental representation of their life experiences. Rather, they

form different representations of events that are thematically unrelated.

> **Principle 7.** The mental representations that people form in the course of comprehending information are domain and theme specific. If information that pertains to different persons or to situations that are not thematically related, a different representation is formed of each.

Perhaps because the validity of this proposition seems almost self-evident, there is surprisingly little research to support it. In a study by Wyer and Bodenhausen (1985), however, people read a story about events that occurred during the course of a cocktail party. Some of the events described were thematically related but were mentioned at different points in the story. Participants recalled these events together and in chronological order regardless of when they were described. However, they recalled unrelated events that occurred near the end of the story before events that were mentioned earlier. Thus, participants apparently integrated thematically related events into a single representation in the course of comprehending them and stored the representation in memory as a single unit of knowledge. However, they stored each knowledge unit independently in the order they were formed, in a manner implied by a "bin" conception of memory (Wyer & Srull, 1989). Consequently, they later recalled more quickly the ones they had stored more recently (those near the top of the bin).

The Impact of Prior Knowledge on the Representation of New Experiences

The comprehension of new information necessarily requires the activation and use of previously acquired concepts and knowledge. This knowledge can potentially influence the mental representations that are formed from the information in two, apparently opposite ways. On one hand, the image component of a situation model that is formed in the course of comprehending information may require features that were not specified, and these features might be added to the representation that is formed. On the other hand, when aspects of new information are redundant with prior knowledge and can be reconstructed on the basis of it, these aspects may omitted from the new representation.

Comprehension-Based Intrusions

To reiterate, the construction of a mental image sometimes requires the addition of features that were not specified in the information on which it is based. The addition of this feature to the representation may be reflected in memory later on, appearing as an intrusion error. This possibility is summarized in the following principle:

> **Principle 8.** People who comprehend information may spontaneously make inferences about features

that are necessary in order to interpret it. These inferences become part of the representation that is formed and, therefore, are later recalled as actually having been mentioned.

Barclay (1973; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974) found evidence of these intrusions in memory for simple sentences, and a more detailed analysis of their occurrence is provided by Graesser, Singer, and Trabasso (1994) in conceptualizing prose comprehension phenomena. Intrusions can also result from attempts to comprehend more complex sets of information. Spiro (1977), for example, asked participants to read an account of a heated argument between an engaged couple over having children. After reading the story, the experimenter incidentally remarked that the couple was happily married. Participants in the course of comprehending this latter information apparently made inferences about what might have happened in the interim. These inferences became part of the mental representation they formed of the story and stored in memory and their later use of this representation as a basis for recalling the sequence produced errors. (For example, they recalled that the woman found out she couldn't have children, or that the man changed his mind.)

The Role of Generalized Representations in the Comprehension of New Information

Many features of the events we experience in daily life are similar to those we have encountered in the past. It seems unlikely that we encode and store these redundant experiences in detail (but see Hintzman, 1986). This possibility, which is consistent with the rule of parsimony noted earlier, is captured by the following more specific principle:

> **Principle 9.** If new information can be comprehended in terms of a previously formed knowledge representation, features that can be reconstructed on the basis of this representation will only be retained if they are necessary in order to reconstruct other information that is presented.

Graesser, Gordon, and Sawyer (1979) proposed a theory of prose comprehension that makes use of this principle. They assumed that when people read about events that exemplify those that compose a previously formed, generalized event representation (e.g., an event prototype or a script), these events are not retained in memory. Rather, people retain only a "pointer" to the generic representation along with a set of translation rules that permit the features of the preexisting representation to be instantiated in terms of the more specific features of the exemplar. Events that occur but are not specified in the generic representation are simply appended as "tags." Thus, suppose someone learns that John went to Yeng Ching's, decided to order chicken fried rice, said "hi" to Mary as she passed by his table, paid $15.80 for the meal, and left. This sequence could be represented

by (1) a pointer to a generic representation of a restaurant visit along with translation rules (customer = John, restaurant = Yen Ching's; meal = chicken fried rice, etc.) and (2) a tag describing the interaction with Mary.

A problem with Graesser and colleagues' (1979) conceptualization lies in its failure to specify how the location of unexpected events that occur is reconstructed. However, an extension of this conceptualization by Trafimow and Wyer (1993) solves this problem. They assumed that events that are redundant with a more general representation are not retained *unless* their retention is necessary in order to localize the point at which the unexpected events occur. Thus, in the preceding example, the decision to have chicken fried rice is redundant with a generic event, "ordered the meal." Nevertheless, it might be retained because it is necessary in order to localize the interaction with Mary. This conceptualization implies that the greater the number of unexpected events that occur, the more theme-related events are likely to be retained. This prediction contrasts with the effect of additional events when the events are treated as independent. In the latter case, the likelihood of recalling a particular item typically decreases with the number presented (Anderson & Bower, 1973; Rundus, 1971). Trafimow and Wyer obtained support for these predictions. That is, they found that the recall of prototype-related events increased with the number of unrelated events contained in a prose passage when participants were aware at the outset of the prototype to which they pertained. When this prototype could not immediately be identified, however, the reverse was true.

Trafimow and Wyer's results suggest a criterion for determining when people use a generalized representation as a basis for comprehending information and when they do not. By applying this criterion, Colcombe and Wyer (2002) showed that people are unlikely to form and use a generalized representation spontaneously. This only occurs when they are exposed to several exemplars of the generalized representation in close temporal contiguity and have an implicit objective of thinking about them in relation to one another. (For a similar conclusion, see Gick & Holyoak, 1983.)

More important, we found that although people use generalized representations to comprehend unfamiliar persons' behavior, they do *not* use them to comprehend events that they imagine themselves or a well-known other performing. In the latter case, people appear to comprehend experiences in terms of an *exemplar* of their own or the other's experience that they have previously formed and stored in memory. This makes intuitive sense. That is, people rarely need to think about different instances of their own behavior in relation to one another and, therefore, may rarely be motivated to extract a generalized representation of it. Rather, they may simply use a recent exemplar of the routine to understand and predict the events that occur. These considerations give rise to the following principle:

> **Principle 10.** People use a generalized representation to comprehend new information if and only if an ex-

emplar (e.g., a situation model) that can be used for this purpose is not easily accessible in memory.

Principle 10 implies that most daily life experiences that we encounter with no particular goal in mind are likely to be represented in memory as situation (e.g., episode) models rather than as abstractions of these experiences. However, the principle may only apply to conditions in which (1) comprehension is the primary objective and (2) the information is of the form to which a situation model is relevant (events that are situationally and temporally constrained). When people have other, more specific objectives, they are likely to construct generalized representations that are relevant to the attainment of these objectives. We now turn to this possibility.

THE FORMATION OF GENERALIZED REPRESENTATIONS

Many mental representations are constructed in the course of more deliberative goal-directed processing. For example, people often wish to form an impression of a person or object, to explain an event that has occurred, to predict a future occurrence, or to make a behavioral decision. In these cases, they might intentionally abstract commonalities among specific experiences that will be useful in attaining their objective, using learned rules of inductive or deductive reasoning. In still other cases, people wish to describe an experience they have had to someone else. Then, they may recode the experience in linguistic terms that they can communicate verbally. In each case, the mental representation that is formed in the course of this goal-directed activity is stored in memory and, therefore, can later be retrieved for use in attaining other objectives to which it is relevant.

Many conceptualizations of goal-directed representational processes assume the following principle:

> **Principle 11.** When people consciously think about two cognitive elements in relation to one another, an association is established between them. The strength of this association increases with the frequency and duration of this cognitive activity. Once the association is formed, thoughts about one of the associated elements are likely to stimulate thoughts about the second.

This principle is implicitly and explicitly applied in conceptualizing not only the content and structure of human thought systems in general (McGuire & McGuire, 1991) but also the representations formed in more specific knowledge domains, including the organization of beliefs about causally related events (McGuire, 1960, 1981; see also Shah & Kruglanski, 2000, 2003), the construction of implicit theories, and the formation of person impressions. These conceptualizations are described in the pages that follow.

The Content and Structure
of Human Thought Systems

The most comprehensive attempt to conceptualize the mental representation of real-world knowledge was proposed by McGuire and McGuire (1991). The McGuires were concerned with the manner in which people construct representations of both real and hypothetical events in the course of thinking about them. Based on the type and implications of the thoughts that individuals spontaneously generated in response to a proposition about a real or hypothetical event, they were able to draw conclusions about the associations that are typically formed among the propositions that compose general world knowledge. Moreover, by identifying participants' cognitive responses to events that made target propositions more or less likely, they drew conclusions about the motivational and cognitive processes that underlie the formation of these associations. They proposed four basic postulates:

1. *Utility maximization.* Events stimulate thoughts about consequences that are similar to the events in desirability. That is, desirable (undesirable) events stimulate thoughts about possible consequences that are also desirable (undesirable).
2. *Congruent origins.* Events stimulate thoughts about antecedents that are similar to them in desirability. That is, desirable (undesirable) events stimulate thoughts about desirable (undesirable) causes.
3. *Wishful thinking.* Desirable events stimulate thoughts about why the events are likely to occur, whereas undesirable events stimulate thoughts about why they will not occur.
4. *Rationalization.* Events that appear likely to occur stimulate thoughts about desirable consequences, whereas events that are considered unlikely stimulate thoughts about undesirable consequences.

These principles are consistent with those assumed in several other areas of research. Kahneman and Miller (1986) note that the contemplation of social information (e.g., "Mr. X is a 26-year-old assistant professor who earns $50,000 a year") activates a personal or social norm that is used as a basis for comparison and, if the information's implications deviate from this norm, to consider possible explanations for it. The spontaneous generation of counterfactuals may be a reflection of these processes.

Principle 11 may also apply to both the McGuires' conceptualization and Kahneman and Miller's. According to this principle, the cognitive activity that results from thinking about the antecedents and consequences of an event leads propositions pertaining to the events to become associated in memory. To this extent, thinking about one proposition should increase the likelihood of recalling the other and, therefore, to use it as a basis for judgment at a later point in time. Wyer and Hartwick (1984) obtained evidence consistent with the utility-maximization and congruent-origins postulates. They showed, for example, that asking participants to report

their belief in a proposition stimulated them to search for an antecedent (a reason why it might either be true or false) and, if a proposition describing such an antecedent happened to be accessible in memory at the time, to identify and use it as a basis for this belief. Correspondingly, asking persons to report the desirability of the event specified in a proposition stimulated them to search for a possible consequence and to use an easily accessible proposition describing such a consequence as a basis for this judgment. The associations that participants formed on the basis of this cognitive activity increased the likelihood that calling their attention to one proposition would cue their recall of the other. Thus, participants were more likely to form an association between a conclusion, C (e.g., "Global warming will make life unbearable by the end of the decade"), and an antecedent, A (e.g., "Air pollution controls are not being enforced") if they had previously reported their belief in C when A was easily accessible in memory, or if they had reported the desirability of A when C was easily accessible. In contrast, reporting their belief in A or the desirability of C under these conditions did not have these facilitating effects.

The Representation of Causal Relationships

People are undoubtedly more likely to base their belief in a conclusion (C) on their belief in a particular antecedent (A) when the two propositions are associated in memory. However, the actual impact of A depends on the extent to which people believe that C is more likely to be true if A is true than if it is not. A model of belief organization proposed by McGuire (1960) and refined by Wyer (1974; Wyer & Goldberg, 1970) captures these possibilities. McGuire argued that people organize their beliefs syllogistically and that, consequently, beliefs in a proposition should be predictable from beliefs in the propositions that logically imply it. Thus, for example, the proposition, C, might be viewed as the conclusion of two mutually exclusive sets of premises, (1) A; if A, then C, and (2) not-A, if not-A, then C. Wyer and Goldberg (1970) postulated that if beliefs were in units of subjective probability, the belief in C, P(C) could be an additive function of beliefs the two sets of premises that imply it, or

$$P(C) = P(A)P(C/A) + P(\sim A)P(C/\sim A) \qquad [1]$$

where P(A) and P($\sim$A) [=1 − P(A)] are beliefs that A is and is not true, respectively, and P(C/A) and P(C/$\sim$A) are beliefs that C is true if A is and is not true, respectively. Numerous empirical studies confirmed the validity of this equation in describing belief formation and change (for a review, see Wyer & Hartwick, 1980).

McGuire (1960) noted that inconsistencies might exist among people's beliefs because they have not recently thought about the beliefs in relation to one another. If this is so, however, calling people's attention to the beliefs in temporal proximity should stimulate them to reduce these inconsistencies. Evidence of this "Socratic effect" was obtained by Rosen and Wyer (1972). That is, when people report their beliefs in propositions of the

sort to which equation 1 pertains in two sessions 1 week apart, their beliefs are more consistent, as defined by the equation, in the second session than they were at first. This effect is less apparent when people report their beliefs in the premises of syllogistically related propositions before reporting their belief in the conclusion during the first session of the experiment (Henninger & Wyer, 1976). Perhaps people find it easy to change their belief in a conclusion to make it consistent with their beliefs in premises, and so they make this change immediately upon encountering it. However, changing beliefs in premises to make them consistent with a previously reported belief in a conclusion requires more cognitive work, and thus it takes more time to occur.

A more interesting contingency was observed in 2000 by Norenzayan and Kim (unpublished data). They found that the Socratic effect observed by Rosen and Wyer (1972) was restricted to North Americans. That is, Asian participants did not change the consistency of their beliefs appreciably over time. This suggests that the motivation to eliminate logical inconsistencies among beliefs is a Western cultural phenomenon and may not generalize to Asian cultures. (For other evidence that Asians are less inclined than Westerners to engage in analytic thinking of the sort implied by a syllogistic model of belief organization, see Choi, Nisbett, & Norenzayan, 1999.)

Related Considerations

The formation of associations between antecedents and consequences has recently been elaborated in a conceptualization of goal structure proposed by Shah and Kruglanski (2000, 2003; Shak, Kruglanski & Friedman, 2003). This conceptualization has many features in common with that of McGuire and McGuire (1991). They propose that calling people's attention to a goal activates thoughts about the means of attaining it. Correspondingly, calling attention to means can activate thoughts about the goals that may result from it. In support of this possibility, activating a goal increased the ease of identifying means-related concepts in a word-recognition task (Shah & Kruglanski, 2000), whereas activating a means to an end increased the ease of identifying words that exemplified the end (Shah & Kruglanski, 2003).

The Construction of Implicit Theories

General knowledge about the causes and consequences of events can function as *implicit theories* about oneself, other persons, or the world in general. As noted earlier in this chapter, an implicit theory consists of two or more temporally or causally related propositions. The simplest form of such a theory might be an *implicational molecule* similar to that postulated by Abelson and Reich (1969; see also Wyer, 2004; Wyer & Carlston, 1979) (e.g., the theory that people make about themselves, others, or real-world events). For example, the generalization that people get what they deserve might be captured by a molecule of the form

$$[\text{P is good (bad). Good (bad) things befall P}] \quad (2)$$

The theory that people who share interests and values like one another might be conveyed in a "similarity-attraction" molecule:

$$[\text{P likes (dislikes) X; O likes (dislikes) X; P likes O}] \quad (3)$$

Molecules about the world in general can also exist. For example, the theory that things that occur frequently are easy to remember might be represented by an "ease-of-retrieval" molecule of the form

$$[\text{Instances of X occur frequently. Instances of X are easy to recall}] \quad (4)$$

Implicit theories may be formed in the course of understanding the causal relations among the events that one encounters in daily life. With frequent repetitions, however, the propositions that compose the representation may become so strongly linked that thoughts about one of them spontaneously activate the others. To this extent, the set of propositions acquires a schematic quality, being used to interpret information that exemplifies them in much the same way as the schema of a human face or a situation model.

The evidence of this possibility, although limited, is compelling. A study by Sentis and Burnstein (1979), in the context of research on cognitive balance (Heider, 1958), is illustrative. Participants were asked to learn liking relations among sets of three persons that were either consistent with the similarity–attraction molecule (or, in Heider's terms, were balanced) or were inconstant with this molecule (imbalanced). Later, they were asked to verify sets composed of one, two, or three of the relations they had seen. The time to verify the relations contained in molecule-inconsistent sets increased with the number to be verified. This suggests that the relations were stored separately from one another and that each had to be verified individually. When the relations were molecule consistent, however, participants took *less* time to verify three relations in combination than to verify only one. Participants apparently formed a schematic representation of molecule-consistent sets of relations at the time they encountered them. Therefore, they found it easier to verify information that matched the configuration as a whole, whereas they were required to unpack the representation to verify its individual features.

Implicit theories are widely applied in explaining and predicting events in one's social and physical world. We return to this possibility later in this chapter.

Person Impression Formation

Preliminary Considerations

Applications of principle 11 are particularly evident in research on person impression formation (Hastie et al., 1980; Srull & Wyer, 1989). In this research, people are typically asked to form an impression of a hypothetical

person on the basis of a set of favorable or unfavorable behaviors. The behaviors are sometimes preceded by favorable or unfavorable trait descriptions. By applying principle 11 to the content, type, and order in which the behaviors are later recalled in this paradigm, Srull and Wyer (1989) drew the following conclusions:

1. People with an impression formation objective interpret the behaviors in terms of traits, thus forming associations between the trait concepts and the behaviors they exemplify (Gordon & Wyer, 1987; Hamilton et al., 1980; but see Klein & Loftus, 1990).
2. People extract an evaluative concept of the person as likeable or dislikeable on the basis of the initial information they receive.
3. Once this concept is formed, they encode the behaviors of the person as favorable or unfavorable and think about them with reference to the person concept, thereby forming associations between the behaviors and the concept.
4. When people encounter a behavior that is eval-

uatively inconsistent with the general concept of the person, they respond in two ways.
a. First, they think about the behavior in relation to others in an attempt to understand why it might have occurred (e.g., why a good person might do a bad thing, or why a bad person might do something admirable). This activity leads associations to be formed between the inconsistent behavior and others.
b. Second, they review the behaviors that are consistent with the central person concept in an attempt to confirm the validity of this concept. This strengthens the behaviors' associations with the concept.

Thus, applying an associative network metaphor, the top half of Figure 12.1 shows the trait-based representations that might be formed from information that a person has engaged in three honest and three intelligent behaviors. If prior information about the person has led a favorable evaluative concept to be formed of him, and if
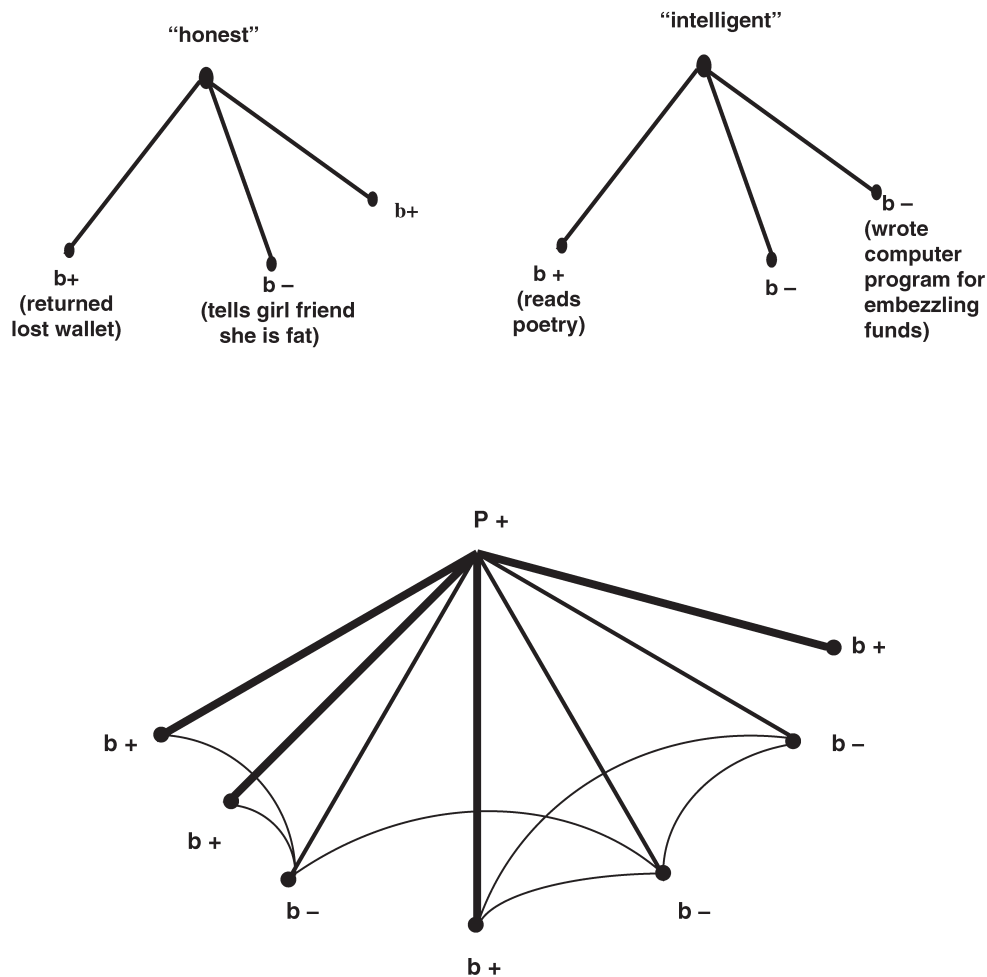
FIGURE 12.1. Metaphorical trait-based and person-based representations of a person characterized by a favorable person concept (P+) and by favorable (b+) and unfavorable (b–) behaviors pertaining to honesty and intelligence. Behaviors that are inconsistent with the person concept are assumed to be thought about in relation to the two behaviors that precede them.

these behaviors are interpreted as favorable (e.g., "returned a lost wallet" or "reads poetry") and unfavorable ("tells his girl friend she is fat" or "developed a computer-based system for embezzling funds from his employer"), the person-based representation might be constructed of the form shown at the bottom of Figure 12.1.

However, it is important to note that the content and structure of the mental representation assumed by the model is closely tied to the *processes* that theoretically underlie impression formation under the conditions of concern. If the processes do not occur, the associations of the sort described in Figure 12.1 would not be formed.

Other qualifications on the processes assumed by Srull and Wyer (1989) to underlie person impressions come from investigations of (1) the formation of impressions of groups, (2) the formation of impressions based on belief and opinions of social relevance, and (3) the impressions formed from behaviors described in a social context (e.g., a conversation). The mental representations that are formed are qualified correspondingly.

Forming Impressions of Groups

Suppose people are asked to form impressions of a group of individuals, each described by a different behavior. Some members' behaviors may be evaluatively inconsistent with their concept of the group as a whole. If the group is cohesive and its members are likely to interact, an understanding of other members' behavior might be thought about in an attempt to explain these inconsistencies. In many cases, however, the group may simply be a collective (African Americans, Nazis, etc.), whose members may not even know one another. Then, this cognitive activity is unlikely to occur, and so the interbehavior associations that facilitate the recall of expectancy-inconsistent behaviors of individual persons may not be formed. Consequently, persons in these conditions often recall expectancy-consistent behaviors better than inconsistent ones (Hamilton & Sherman, 1996; Srull, 1981; Wyer & Srull, 1989). (But see Hamilton & Gifford, 1976, for a qualification on the latter conclusion.)

These considerations have implications for the effects of stereotyping on memory. If people have formed a stereotype-based impression of an ethnic group on the basis of behaviors of several different members, they may have better recall of a person whose behavior confirms their stereotype than of a person whose behavior disconfirms it. At the same time, if people encounter several behaviors of a particular group member, they are more likely to recall those behaviors that conflict with their stereotype-based expectations than behaviors that confirm these expectations.

Impressions Based on Beliefs and Opinions

In much research on person impression formation, people are asked to form impressions of someone on the basis of behaviors that are of little personal or social relevance ("got confused by television sitcoms," "shouted and honked at slow drivers," etc.). Outside the laboratory, however, impressions are often based on expres-

sions of opinions about social issues. In this case, people may not think about the consistency of these opinions with a concept of the actor. Rather, they may compare the opinions to their *own* positions on the issues. In doing so, they are likely to expend more cognitive effort on expressions of opinion with which they disagree. As a result of this cognitive activity, they may have better recall of these behaviors and opinion statements than of those with which they agree, independently of their consistency with the general attitudes and orientation of the person who expressed them (Wänke & Wyer, 1996; Wyer, Lambert, Budesheim, & Gruenfeld, 1992).

Impression Formation in Social Contexts

Outside the laboratory, we rarely encounter a listing of someone's behavior out of the context in which the behaviors occur. Rather, we hear stories about the person, or anecdotes about his or her activities in specific situations. Moreover, the information is itself conveyed in a social context (e.g., a conversation in which the individual is a topic of discussion. In these conditions, the impression formation processes may differ in several ways from those that Srull and Wyer (1989) assumed.

First, when trait and behavior descriptions of a person are conveyed in a conversation, speakers' trait descriptions of the target appear to influence participants' impressions of the speakers rather than the person the speakers are discussing (Wyer, Budesheim, & Lambert, 1990). Furthermore, listeners use the behaviors mentioned by one speaker as a basis for confirming the validity of their impression of the other speaker. (Thus, for example, if Arthur mentions a favorable behavior of a target that is inconsistent with Bob's unfavorable trait description of the individual, it confirms the suspicion that Bob is simply being nasty rather than conveying a true description of the target's personality.) Consequently, one speaker's behaviors become associated with the other speaker as well and are better recalled for this reason (Wyer et al., 1990).

Second, when participants overhear a conversation about one of the speakers, their reactions to the behaviors described are guided by normative expectations for speakers to be both polite and modest. Thus, they think more extensively about behaviors that deviate from these expectations. For example, they think more about a speaker's unfavorable descriptions of the person to whom they are speaking (which violate a norm to be polite) but think more about favorable descriptions that speakers provide about themselves (which violate a norm to be modest). Thus, these behaviors become more strongly associated with the speaker and are better recalled later (Wyer, Budesheim, Lambert, & Swan, 1994).

These contingencies exemplify a general point. That is, the impression-related mental representations that result from the associative processes implied by principle 11 depend substantially on the social context in which the information is presented as well as on the information itself. Conclusions about the content and structure of the representations cannot be divorced from these processes.

Spontaneous Construction of Generalized Representations

In the absence of a specific goal that requires it, the spontaneous abstraction of a generalized representation from specific instances occurs less often than one might intuitively expect. This is particularly true of event representations. Gick and Holyoak (1983; see also Holyoak & Koh, 1987) found that participants' solutions to a problem were typically unaffected by their exposure to analogous problems that could be solved using the same general principle. In fact, several specific instances of the problem were required before the principle was recognized. Colcombe and Wyer (2002; Wyer et al., 2002) drew similar conclusions in a different research paradigm.

The formation of generalized entity representations may also not be spontaneous. Klein and his colleagues (Klein & Loftus, 1993b; Klein, Loftus, Trafton, & Fuhrman, 1992; Sherman & Klein, 1994) showed that participants' trait judgments of a person was independent of their knowledge about the person's behaviors when they had had substantial past experience with the person being judged. Although this suggests that trait representations of well-known persons exist in memory independently of behaviors, this does not necessarily mean that the representations were formed spontaneously. In some studies (Sherman & Klein, 1994), participants' experience with the target person was experimentally manipulated. In these studies, however, participants were given an implicit impression formation objective. Therefore, these results also do not indicate that new generalized entity representations are constructed in the absence of specific goals that require them.

THE MENTAL REPRESENTATION OF SELF-KNOWLEDGE

Virtually all the knowledge we acquire implicates ourselves as either an actor, an observer, or an experiencing agent. Even when the information we receive pertains to matters of little personal concern, we are often aware of the act of acquiring it and the situation in which it occurs. To this extent, self is part of the knowledge representation we form.

On the other hand, many features of a mental representation decay over time. Furthermore, not all features of an existing representation are included in the new representation that we construct in the course of goal-directed activity. For these reasons, the knowledge we acquire often becomes dissociated from the experience of acquiring it. Thus, for example, we may have been aware of the experience of hearing George W. Bush make assertions about Iraq's stockpiling of nuclear weapons at the time we first encountered them in a speech and may have thoughts that his assertions were false and irresponsible. After a period of time, however, we may still recall his statements but be unable to remember the particular time and place in which the statements were uttered. A detailed discussion of self-knowledge and its impact on memory and judgments is beyond the scope of this chapter (for reviews, see Greenwald & Pratkanis, 1984; Kihlstrom & Klein, 1994; Markus & Nurius, 1986). Certain observations are nevertheless worth mention.

The content and structure of knowledge about oneself are presumably similar to that of knowledge about others. Procedural knowledge concerns one's own behavioral reactions to the stimuli we encounter in a situation. Declarative knowledge can be of any or all of the forms mentioned earlier in this chapter. For example, it can consist of both situation models of one's experiences in specific situations and generalized event representations that have been formed in the course of thinking about an experiences or communicating about it to others. One might also form generalized entity representations of oneself, consisting of traits and prototypical behaviors.

In this regard, many theories assume that people construct a *self-concept*, or self-schema, that is evaluative in nature, and to which more specific trait and behavioral descriptions are attached. In fact, several independent self-representations could exist that are domain and situation specific (Markus & Nurius, 1986; Markus & Wurf, 1987). Other distinctions are implicit in Higgins's (1987, 1996b, 1998) theory of self-regulatory processes. Higgins (1996b) distinguishes between "instrumental" self-representations (analogous to procedural self-knowledge, as noted earlier in this chapter) and "dispositional" self-representations (composed of traits and competencies). These latter representations can be subdivided into "actual" selves, "ideal" selves, and "ought" selves (characteristics of oneself that significant others consider desirable). Calling attention to differences in the content of these latter self-representations can have important emotional and behavioral consequences (Higgins, 1996b, 1998).

The different self-representations postulated by Markus, Higgins, and their colleagues might be structurally similar to an associative network of the sort described earlier. On the other hand, Klein and his colleagues (for a review, see Klein & Loftus, 1993b) have reported evidence that trait self-descriptions and behaviors that exemplify them are stored separately in memory. In a typical study (Klein & Loftus, 1993a), participants were asked, in counterbalanced order, both (1) to verify that they had a particular trait and (2) to generate a behavior they had performed that exemplified the trait. Asking participants to recall a trait-related behavior had little effect on the time they took to verify that they possessed the trait it exemplified. Correspondingly, making a trait judgment did not influence the time required to generate a trait-related behavior. These and other findings suggest that people's trait representations of themselves are stored separately from the representation of behaviors they perform that exemplify them. Consequently, neither is typically used as a basis for judgments of the other.

An important component of self-knowledge could concern one's subjective reactions to the objects and events that one encounters. These feelings are often used as bases for judgments and decisions about their referents (Schwarz, 2004; Schwarz & Clore, 1983, 1996; Strack & Deutsch, 2004). However, these feelings are not them-

selves part of the self-knowledge that is stored in memory. Rather, they may be conditioned *responses* to descriptive concepts and knowledge that occur when this knowledge is activated (Wyer, Clore, & Isbell, 1999).

THE USE OF MENTAL REPRESENTATIONS AS BASES FOR JUDGMENT

As noted earlier, the nature of mental representations is of limited interest in the absence of how these representations come into play in judgment and behavioral decision making. That is, one must ultimately be able to specify both *when* one or another representation is likely to be applied and *how* the representation is used, once it is called into play. This chapter concludes with a brief discussion of each question.

Which Representations Are Used?

More than one mental representation can have implications for a judgment or decision, and these implications can differ. In such cases, people are unlikely to conduct an exhaustive review and analysis of the alternative criteria they might potentially apply. Rather, they may use the first knowledge representation that comes to mind, provided it is deemed applicable for attaining the goal one is pursuing. The factors that influence the activation and use of previously acquired knowledge have been reviewed in detail elsewhere (Higgins, 1996a) and need not be reiterated in detail. Two factors, recency and frequency, are particularly important.

Recency Effects on Memory and Judgment

The more recently a unit of knowledge has been formed and/or used, the more accessible it is in memory. Therefore, the more likely it is to be brought to bear on judgments and decisions to which it is potentially applicable. This gives rise to the following principle:

> **Principle 12.** Once a new mental representation has been formed, it is likely to be used as a basis for judgments and decisions to which it applies independently of the information that led to its construction.

Numerous studies confirm this principle (for reviews, see Higgins, 1996a; Wyer, 2004). Impression formation research, for example, shows that once people form an initial trait concept of a person on the basis of a subset of his or her behaviors, this concept is later used as a basis for later judgments of the person without consulting the behavior that led the concept to be formed (Carlston, 1980; Carlston & Skowronski, 1986; Lingle & Ostrom, 1979; Srull & Wyer, 1989). The verbal descriptions that are assigned to stimuli can be influenced by the context in which the stimuli are encountered. However, these descriptions are later recalled out of their original context and used as bases for judgments and decisions without considering the conditions that gave rise to their construction (Higgins & Lurie, 1983; Sherman, Ahlm, Berman, & Lynn, 1978).

Similar effects can occur in memory. People who have formed a mental representation of an experience at the time the experience occurred might later recall and use the representation for a purpose they did not anticipate. In the course of pursuing this objective, they may form a new representation, the content and structure of which differ from the one they had formed earlier, and may store this representation in memory independently of the original. When they are later asked to reconstruct the experience, they may use this more recently formed representation rather than the original as a basis for their reconstruction. In this case, memory errors are likely to occur.

This phenomenon may be particularly evident when the original representation is coded nonverbally. In a study by Schooler and Engstler-Schooler (1990), for example, participants were shown a number of faces and asked to describe one of them verbally. They were later less able to recognize this face than faces that had not been described. Thus, the verbally coded representation they had constructed, which was less detailed than the nonverbal representation they had formed, was used as a basis for recognition, decreasing accuracy.

This may only occur if the verbally coded representation that is accessible in memory is assumed to be a sufficient basis for judgment (Chaiken et al., 1989). If the representation is deemed irrelevant, participants may revert to a previously formed representation that they believe to be more appropriate. A study by Adaval and Wyer (2004) provides evidence of this contingency. Participants watched a 10-minute segment of a movie depicting the interaction between a husband and wife. The protagonists' verbal exchanges during the interaction were particularly relevant to impressions of their personalities, whereas their nonverbal behaviors were not. Participants were assumed to construct a situation model of the events they observed in the course of watching them. After observing the interaction, some participants wrote down their impressions of the protagonists whereas others described the sequence of events that occurred. Participants in the course of communicating their impressions were expected to construct a new representation of the protagonists on the basis of things the protagonists had said to one another. However, participants who communicated the sequence of events that occurred were expected to include the protagonists' nonverbal behaviors in their representation as well.

After performing the writing task, participants were given a recognition memory test. We expected that participants would use the representation they had formed in performing the writing task as a basis for identifying recognition items to which it was potentially relevant, and that their use of this representation would decrease accuracy. However, they should use their more detailed situation model to verify events to which the abstract representation was irrelevant, and so their recognition accuracy should not be affected. Consistent with expectations, describing the sequence of events that occurred

decreased participants' recognition of both statements and nonverbal behaviors relative to control conditions. In contrast, communicating impressions of the protagonists decreased recognition of protagonists' statements but *not* their nonverbal behaviors. In other words, the abstract representations that participants constructed decreased recognition of only those events to which the representations were relevant.

The construction of a new representation can influence not only what is remembered but also how it is remembered. In a study by Higgins and Rholes (1978), participants who had received information about a target person's behavior were asked to communicate a description of the individual to someone who either liked the person or disliked the person. Participants' written communications were biased toward the ostensible attitudes of the intended recipient. However, their later recall of the original information was similarly biased. This memory difference was not evident, however, when participants expected to prepare a message but did not actually do so. Thus, participants presumably constructed a new representation of the target in the course of preparing their written communication and used this representation, once constructed, as a basis for recalling the information on which it was based.

The influence of communicating about an experience or memory for the original experience has important implications. In an early study by Loftus and Palmer (1974), participants who had seen a picture of an automobile accident were asked either how fast the car had been going when it "smashed into" the tree or how fast it had been going when it "hit" the tree. Not surprisingly, participants' estimates were greater in the first condition than the second. When later describing the scene of the accident, however, the former participants were more likely to report seeing broken glass, although this was not actually shown in the picture. Participants when answering the first question apparently constructed a mental image of the accident that was consistent with implications of the question, and features of this image provided the basis for their memory of what they had seen.

It is easy to imagine how similar effects could occur in actual situations. For example, the prior questioning of witnesses before they appear on the stand could influence their reconstructive memory for the events they had witnessed, thus potentially invalidating their "eyewitness" testimony (Loftus, 1995). Adults' post hoc "memories" of the sexual abuse they experienced as a child could be similarly affected (Loftus, 2000). In this regard, Loftus and Palmer (1974) suggest that new representations can actually replace the originals rather than simply being more easily accessible. However, the two possibilities have not been distinguished empirically.

Frequency Effects

The relative accessibility of a representation in memory, and, therefore, the likelihood of using it, is also a function of the frequency with which it has been applied in the past. When an abstract representation has been fre-

quently applied, it may acquire the status of an implicit theory that is brought to bear on not only judgments of future events but memory for past ones. Furthermore, these theories, if applied frequently, can become chronically accessible in memory (Bargh, Bond, Lombardi, & Tota, 1986; Higgins, King, & Mavin, 1982). Consequently, these theories not only can influence what aspects of new information are retained in memory (Higgins et al., 1982) but also can be used as a basis for judgment in their own right. Ross (1989) cites several examples of these effects in conceptualizing the role of implicit theories in reconstruction of the past. For example, women's retrospective accounts of their emotional states during their most recent menstrual period are more similar to their implicit theories about their typical reactions during their menstrual cycle than to the actual emotions they experienced, as indicated by a daily diary they kept during the period in question (Ross, 1989). Similarly, persons who had an implicit theory that a training program would improve their study skills tended to recall their pretraining performance level as less than their posttraining level regardless of whether they had actually improved or not (Conway & Ross, 1984).

How Are Mental Representations Used?

Research on the use of different types of representations are limited. Three types of mental representation described in this chapter—entity (e.g., person) representations, event representations, and implicit theories (implicational molecules)—are particularly worth considering.

The Use of Entity Representations

The most common type of entity representation of concern in social psychological research pertains to persons (e.g., a specific individual, a prototypical person, or a group). The representation is typically categorical, consisting of a central concept with which a number of specific features are attached (see Figure 12.1). Once such a representation is formed, participants typically use its central concept as a basis for judgment without performing an extensive review and analysis of its contents (principle 12). This is true even if the behaviors associated with the concept have different implications. Thus, for example, if a favorable or unfavorable concept of someone has been constructed, this concept is used as a basis for evaluating the person despite the fact that behaviors that are inconsistent with this concept are better recalled (Srull, Lichtenstein, & Rothbart, 1985). This is true even when subsequent information reveals that the basis for forming the central concept is invalid (Lingle & Ostrom, 1979; Wyer & Budesheim, 1987; Wyer & Unverzagt, 1985).

Similar considerations surround the use of categorical representations as bases for judgment. Fiske and Pavelchak (1986; see also Fiske & Neuberg, 1990) postulate that when people receive a trait or behavioral description of a person who belongs to a specified social

category, they first decide whether the person is a typical member of the category, based on a global (schematic) comparison of the individual's features to those that characterize the category in general. If the individual is judged to be typical, the favorableness of the category as a whole is used as a basis for their evaluation. Only if the individual is judged to be atypical do people perform a detailed, "piecemeal" analysis of the individual's specific features.

To the extent individual features of an entity representation are considered, the question arises as to how they are used to compute a judgment. This could be done using mechanistic procedures similar to those produced by Fishbein (1963; Fishbein & Hunter, 1964) or Anderson (1971, 1981). To this extent, the diagnosticity and extremity of the individual features are likely to be important. This may not be the case, however, when other computational procedures are used, as suggested in the next section.

The Use of Event Representations

Event representations are composed of a sequence of causally and temporally related events that are conceptualized as a single unit. Evaluations that are made on the basis of these representations are unlikely to involve a piecemeal computation of specific features. Rather, they may be based on a more global assessment of the implications of the event sequence as a whole.

Pennington and Hastie (1986, 1988, 1992) first identified the effects of these representations in research on jury decision making. Their studies suggest that the sort of representation that is formed from information depends on the format in which it is presented, and that the judgments that are made depend on the type of representation that is formed. In a typical study (Pennington & Hastie, 1988), participants received courtroom testimony about a crime in which the evidence was conveyed in one of two ways. In *witness-order* conditions, the testimony was organized according to the witness who presented it. In *story-order* conditions, it was presented in the order it became relevant in constructing a narrative of the events that led up to the crime, the crime itself, and its aftermath. When the prosecution testimony and the defense testimony were conveyed in different orders, 70% of the participants favored the side whose evidence was conveyed in story order. When the evidence for both sides was conveyed in the same order, participants were equally likely to favor each side, but they were more confident of their decision in story-order conditions.

In Pennington and Hastie's studies, an understanding of the temporal order of the events was particularly important in assessing the circumstances of the crime and the likelihood of the defendant's guilt. When this is not the case, the representation that is formed, and the strategy that is employed, can depend on the format in which the information is presented. This possibility was examined in a series of studies by Adaval and her colleagues. In some experiments (Adaval & Wyer, 1998), participants evaluated a vacation trip on the basis of information about the events that would occur. In others

(Adaval, Isbell, & Wyer, in press), they evaluated a politician on basis of events that took place in the course of his career. In *narrative-format* conditions, the events described in both sets of studies were conveyed in a narrative that indicated the order they occurred. In *list-format* conditions, they were conveyed in a bullet-point format and no indication was given of their temporal order. The results of the studies were quite consistent. When all the information was evaluatively similar (all favorable or all unfavorable), evaluations of the target (the vacation or the politician) were nonsignificantly more extreme when the information was conveyed in a list, suggesting that a holistic processing strategy was not inherently any more effective than a piecemeal strategy.

However, accompanying the verbal information with pictures increased the evaluations of the target when the information was conveyed in a narrative but decreased the extremity of these evaluations when the same information was conveyed in a list. As noted earlier, visual images are likely to be components of an event representation, and pictures may facilitate the construction of these images. Consequently, they increase the impact of a narrative representation of the sequence of events as a whole. On the other hand, pictures may interfere with the piecemeal integration of the evaluations of individual events, which is primarily a semantic process. In short, the effects of pictures on evaluations depend on the computational strategy that participants employed, and this strategy, in turn, depends on the sort of representation that participants formed from the information they receive.

Implicit Theories and Implicational Molecules

The influence of implicit theories is evident in virtually all domains of social judgment and decision making. For example, they can guide reactions to success and failure (Dweck, 1991; Dweck & Leggett, 1988), perceptions of social support (Mankowski & Wyer, 1996), perceptions of social reality (Shrum, 2002), person perception (Hong, Chiu, Dweck, & Sacks, 1997), and hindsight bias (Fischoff, 1982; Sanna & Schwarz, 2003). Cultural differences in judgments and behavior can also be traced to differences in the implicit theories that are applied (cf. Hong & Chiu, 2001; Hong, Morris, Chiu, & Benet-Martinez, 2000). A review of research on the use of implicit theories is beyond the scope of this chapter (see Wyer, 2004). A general conceptualization of the rules that govern the use of implicit theories is worth noting briefly.

The simplest implicit theories can be conceptualized as implicational molecules of the form described earlier in this chapter. The use of these molecules in explanation and prediction is postulated by Abelson and Reich (1969) to be governed by a *completion principle*:

> **Principle 13.** If a specific experience or set of experiences instantiates all but one proposition in an implicational molecule, an instantiation of the remaining proposition is inferred to be true as well.

The power of this molecule derives from the assumption that its validity does not depend on which components of a molecule are instantiated by the information available and which are to be inferred. For example, reconsider the "just deserts" molecule noted earlier in this chapter; that is, [P is good (bad). Good (bad) things befall P]. To the extent this molecule is applied, people are likely to infer that a particular person will get what he or she deserves (e.g., that if the person is bad, he or she will be punished or will otherwise come to harm). However, they are also likely to infer that a person deserves what he or she gets (e.g., if the person has experienced misfortune, he or she is likely to be bad or to have done a bad thing). The tendency for persons to make this latter inference has been demonstrated in numerous studies on the disposition to maintain beliefs in a just world (Lerner, Miller, & Holmes, 1976; Lerner & Simmons, 1966; Walster, 1966; Wyer, Bodenhausen, & Gorman, 1985).

The use of numerous cognitive heuristics can also be conceptualized as applications of implicational molecules (Wyer, 2004). For example, consider the molecules [Instances of X occur frequently. Instances of X come to mind easily] and [P is a member of category X. P has attributes x, y, z . . .]. These molecules have implications similar to the availability and representativeness heuristics identified by Tversky and Kahneman (1973, 1982). That is, people not only infer that things come to mind easily if they occur frequently but also infer that things occur frequently if they come easily to mind. (For numerous examples of the impact of ease of retrieval on judgments, see Schwarz, 1998, 2004.) Similarly, people may infer not only that someone who is a Chinese studies professor is likely to be short, have dark hair, and read poetry but also that a person who is short, has dark hair, and reads poetry is likely to be a Chinese studies professor (Kahneman & Tversky, 1971).

Other molecules can be specific to certain types of persons, objects and events. For example, "liking" and "loving" reflect similar affective relations. Yet, [P likes X. O likes X. P likes O] and [P loves X. O loves X. P dislikes O] may coexist and may be applied in different situations, depending on the nature of X. (For example, the first molecule may be applied when X is an inanimate object, a personal interest, or a basic value. The second may be applied when X is a specific job to which both P and O aspire, or a woman whom both would like to marry.) However, the use of the molecules can also vary with their relative accessibility in memory, as noted in the previous section.

CONCLUDING REMARKS

This chapter has attempted to review the general principles that govern the representation of social knowledge and to provide examples of their application. Although the research cited has generally been conducted in the laboratory, I have tried to focus on the representation of knowledge of the sort people acquire in every day life, from conversation, movies, and television or from direct experience. Although the research I have covered is

broad, it is by no means complete. Several areas, including the mental representation of affect and emotion (Ortony, Clore, & Collins, 1988) and the motivational determinants of mental representation (cf. Kunda, 1987, 1990), have barely been covered at all. Nevertheless, this chapter ideally provides a general indication of the state of existing knowledge on the determinants and consequences of mental representations and, in doing so, suggests directions for future research and theorizing.

NOTES

1. A distinction should be made between the mental image that is formed of a verbally described event and a picture. For one thing, many features that might be conveyed in a picture may be absent. (For example, the image of "a boy kicked the football" may not contain an indication of his hair color.) At the same time, features in other sense modalities (e.g., an acoustic representations of a speaker's tone of voice) may be part of the image as well.
2. We will not enter into the debate as to whether visual images are actually formed (cf. Anderson, 1983; Pylyshyn, 1973). However, research by Kosslyn (1980) and others (cf. Shepard & Metzler, 1971) testifies to the utility of a visual image metaphor in conceptualizing information processing and the representations that result from it. Its utility is further demonstrated in the discussion to follow.

REFERENCES

Abelson, R. P. (1959). Modes of resolution of belief dilemmas. *Conflict Resolution*, *3*, 343–352.

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. (Eds.). (1968). *Theories of cognitive consistency: A sourcebook*. Chicago: Rand-McNally.

Abelson, R. P., & Reich, C. M. (1969). Implicational molecules: A method for extracting meaning from input sentences. In D. E. Walker & L. M. Norton (Eds.), *Proceedings of the International Joint Conference on Artificial Intelligence* (pp. 641–647). Washington, DC.

Adaval, R., Isbell, L. M., & Wyer, R. S., Jr. (in press). The impact of pictures on narrative-based impression formation: A process interference model. *Journal of Experimental Social Psychology*.

Adaval, R., & Wyer, R. S., Jr. (1998). The role of narratives in consumer information processing. *Journal of Consumer Psychology*, 7, 207–245.

Adaval, R., & Wyer, R. S., Jr. (2004). Communicating about a social interaction: Effects on memory for protagonists' statements and nonverbal behaviors. *Journal of Experimental Social Psychology*.

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, *85*, 249–277.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston.

Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, *78*, 171–206.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Barclay, J. R. (1973). The role of comprehension in remembering sentences. *Cognitive Psychology*, *4*, 229–254.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, *13*, 471–481.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–62). Mahwah, NJ: Erlbaum.

Bargh, J. A., Bond, R. N., Lombardi, W., & Tota, M. E. (1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology*, *50*, 869–878.

Barsalou, L. W. (1990). On the indistinguishabiity of exemplar memory and abstraction in category representations. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Content and process specificity in the effects of prior experience: Advances in social cognition* (Vol. 3, pp. 61–88). Hillsdale, NJ: Erlbaum.

Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 29–102). Hillsdale, NJ: Erlbaum.

Black, J. B., Turner, T., & Bower, G. H. (1979). Point of view in narrative comprehension, memory, and production. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Bower, G. H. (1991). Mood congruity of social judgments. In J. P. Forgas (Ed.), *Emotion and social judgments* (pp. 31–53). Oxford, UK: Pergamon Press.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretative approach. *Cognitive Psychology*, *3*, 193–209.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.

Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver: A guide to improving thinking, learning, and creativity*. New York: Freeman.

Brewer, W. F. (1995). To assert that essentially all human knowledge and memory is represented in terms of stories is certainly wrong. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 8, pp. 109–119). Hillsdale, NJ: Erlbaum.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.

Carlston, D. E. (1980). Events, inferences and impression formation. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, & D. Carlston (Eds.). *Person memory: The cognitive basis of social perception* (pp. 89–119). Hillsdale, NJ: Erlbaum.

Carlston, D. E. (1994). Associated Systems Theory: A systematic approach to cognitive representations of persons. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 7, pp. 1–78). Hillsdale, NJ: Erlbaum.

Carlston, D. E., & Skowronski, J. J. (1986). Trait memory and behavior memory: The effects of alternative pathways on impression judgment response times. *Journal of Personality and Social Psychology*, *50*, 5–13.

Carlston, D. E., & Smith, E. R. (1996). Principles of mental representation. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 184–210). New York: Guilford Press.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63.

Colcombe, S. J., & Wyer, R. S., Jr. (2002). The role of prototypes in the mental representation of temporally-related events. *Cognitive Psychology*, *44*, 67–103.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology*, *47*, 738–748.

Dreben, E. K., Fiske, S. T., & Hastie, R. (1979). The independence of item and evaluative information: Impression and recall order effects in behavior-based impression formation. *Journal of Personality and Social Psychology*, *37*, 1758–1768.

Dweck, C. S. (1991). Self-theories and goals: Their role in motivation, personality, and development. In R. Dienstbier (Ed.), *Nebraska Symposium on Motivation: Vol. 38. Perspectives on motivation* (pp. 199–235). Lincoln: University of Nebraska Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *25*, 109–116.

Ebbesen, E. B. (1980). Cognitive processes in understanding ongoing behavior. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, & D. Carlston (Eds.), *Person memory: Cognitive basis of social perception* (pp. 179–226). Hillsdale, NJ: Erlbaum.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Fischoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 332–351). New York: Cambridge University Press.

Fishbein, M. (1963). An investigation of the relationships between beliefs about an object and attitude toward that object. *Human Relations*, *16*, 233–239.

Fishbein, M., & Hunter, R. (1964). Summation versus balance in attitude organization and change. *Journal of Abnormal and Social Psychology*, *69*, 505–510.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 167–203). New York: Guilford Press.

Garnham, A. (1981). Mental models as representations of text. *Memory and Cognition*, *9*, 560–565.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*, 69–83.

Gordon, S. E., & Wyer, R. S., Jr. (1987). Person memory: Category-set-size effects on the recall of a person's behaviors. *Journal of Personality and Social Psychology*, *53*, 648–662.

Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Memory for typical and atypical actions in scripted activities: Test of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *18*, 319–322.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101,* 371–395.

Greenwald, A. G. (1968). Cognitive learning, cognitive responses to persuasion and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147–170). New York: Academic Press.

Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer,

Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol.3, pp. 129–178). Hillsdale, NJ: Erlbaum.

Gruenfeld, D. H. (1995). Status, ideology, and integrative complexity on the U.S. Supreme Court: Rethinking the politics of political decision making. *Journal of Personality and Social Psychology*, *68*, 5–20.

Hamilton, D. H., Katz, L. B., & Leirer, V. O. (1980). Organizational processes in impression formation. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, & D. Carlston (Eds.), *Person memory: The cognitive basis of social perception* (pp. 121–153). Hillsdale, NJ: Erlbaum.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgment. *Journal of Experimental Social Psychology*, *13*, 392–407.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, *103*, 336–355.

Harvey, O. J., Hunt, D., & Schroder, H. (1961). *Conceptual systems and personality organization.* New York: Wiley.

Harvey, O. J., Reich, J., & Wyer, R. S., Jr. (1968). Effects of attitude direction, attitude intensity and structure of beliefs upon differentiation. *Journal of Personality and Social Psychology*, *10*, 472–478.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*, 25–38.

Hastie, R., Ostrom, T. M., Ebbesen, E. B., Wyer, R. S., Jr., Hamilton, D. L., & Carlston, D. E. (Eds.). (1980). *Person memory: Cognitive basis of social perception.* Hilldale, NJ: Erlbaum.

Hastie, R., & Pennington, N. (1995). The big picture: Is it a story? In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (vol. 8, pp. 133–138). Mahwah, NJ: Erlbaum.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Henninger, M., & Wyer, R. S., Jr. (1976). The recognition and elimination of inconsistencies among syllogistically-related beliefs: Some new light on the "Socratic effect." *Journal of Personality and Social Psychology*, *34*, 680–693.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1996a). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1996b). The "self digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*, 1062–1083.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). San Diego, CA: Academic Press.

Higgins, E. T., Bargh, J. A., & Lombardi, W. (1985). The nature of priming effects on categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 59–69.

Higgins, E. T., King, G. A., & Mavin, H. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, *43*, 35–47.

Higgins, E. T., & Lurie, L. (1983). Context, categorization and recall: The "change-of-standard" effect. *Cognitive Psychology*, *15*, 525–547.

Higgins, E. T., & Rholes, W. S. (1978). "Saying is believing": Effects of message modification on memory and liking for the person described. *Journal of Experimental Social Psychology*, *14*, 363–378.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace model. *Psychological Review*, *93*, 411–428.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition*, *15*, 332–340.

Hong, Y., & Chiu, C. (2001). Toward a paradigm shift: From cross-cultural differences in social cognition to social-cognitive mediation of cultural differences. *Social Cognition*, *19*, 181–196.

Hong, Y., Chiu, C., Dweck, C. S., & Sacks, R. (1997). Implicit theories and evaluative processes in person cognition. *Journal of Experimental Social Psychology*, *33*, 296–303.

Hong, Y., Morris, M. W., Chiu, C., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*, 709–720.

Humphreys, M. S., & Kashima, Y. (2002). Connectionism and self: Distributed representational systems and their implications for self and identity. In Yoshihisa, M. Foddy, & M. Platow (Eds.), *Self and identity: Personal, social, and symbolic* (pp. 27–54). Mahwah, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469–500). Cambridge, MA: MIT Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Kahneman, D., & Tversky, A. (1971). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.

Kihlstrom, J. F., & Klein, S. B. (1994) The self as a knowledge structure. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 2, pp. 153–208). Mahwah, NJ: Erlbaum.

Kintsch, W. (1998).*Comprehension: A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Klein, S. B., & Loftus, J. (1990). Rethinking the role of organization in person memory: An independent trace storage model.*Journal of Personality and Social Psychology*, *53*, 400–410.

Klein, S. B., & Loftus, J. (1993a). Behavioral experience and trait judgments about the self. *Personality and Social Psychology Bulletin*, *19*, 740–745.

Klein, S. B., & Loftus, J. (1993b). The mental representation of trait and autobiographical knowledge about the self. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 5, pp. 1–49). Mahwah, NJ: Erlbaum.

Klein, S. B., Loftus, J., Trafton, J. G., & Fuhrman, R. W. (1992). Use of exemplars and abstractions in trait judgments: A model of trait knowledge about the self and others. *Journal of Personality and Social Psychology*, *63*, 739–753.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories.*Journal of Personality and Social Psychology*, *53*, 636–647.

Kunda, Z. (1990). The case for motvaated reasoning. *Psychological Bulletin*, *108*, 480–498.

Lerner, M. J. (1970). The desire for justice and reactions to victims. In J. Macaulay & L. Berkowitz (Eds.), *Altruism and helping behavior* (pp. 205–229). New York: Academic Press.

Lerner, M. J., Miller, D. T., & Holmes, J. G. (1976). Deserving and the emergence of forms of justice. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 9, pp. 133–162). New York: Academic Press.

Lerner, M. J., & Simmons, C. H. (1966). Observer's reaction to the "innocent victim: Compassion or rejection? *Journal of Personality and Social Psychology*, *4*, 203–210.

Liberman, N., & Förster, J. (2000). Expression after suppression: A motivational explanation of postsuppressional rebound. *Journal of Personality and Social Psychology*, *79*, 190–203.

Lichtenstein, M., & Srull, T. K. (1987). Processing objectives as a determinant of the relationship between recall and judgment. *Journal of Experimental Social Psychology*, *23*, 93–118.

Lingle, J. H., & Ostrom, T. M. (1979). Retrieval selectivity in memory-based impression judgments. *Journal of Personality and Social Psychology*, *37*, 180–194.

Linville, P. W. (1982). Affective consequences of complexity regarding the self and others. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition.* Hillsdale, NJ: Erlbaum.

Loftus, E. F. (1995). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560–572.

Loftus, E. F. (2000). Remembering what never happened. In E. Tulving (Ed.), *Memory, consciousness and the brain: The Tallinn conference* (pp. 106–118). Philadelphia: Taylor & Francis.

Loftus, E. F., & Palmer, J. (1974). Reconstruction of automobile destruction. *Journal of Verbal Learning and Verbal Behavior*, *2*, 467–471.

Luna, D. (2005). Integrating ad information: A text-processing approach. *Journal of Consumer Psychology*, *15*, 38–51.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes*, *22*, 199–224.

Mandler, J. (1979). Categorical and schematic organization in memory. In C. R. Puff (Ed.), *Memory, organization and structure*. New York: Academic Press.

Mandler, J. (1984). *Stories, scripts anda scenes: Aspects of schema theory.* Hillsdale, NJ: Erlbaum.

Mankowski, E. S., & Wyer, R. S., Jr. (1996). Cognitive processes in perceptions of social support. *Personality and Social Psychology Bulletin*, *22*, 894–905.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Markus, H., & Wurf, E. (1987). The dynamic self-concept. *Annual Review of Psychology*, *38*, 299–337.

McGuire, W. J. (1960). A syllogistic analysis of cognitive relationships. In M. J. Rosenberg & C. I. Hovland (Eds.), *Attitude organization and change* (pp. 140–162). New Haven, CT: Yale University Press.

McGuire, W. J. (1981). The probabilogical model of cognitive structure and attitude change. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 291–307). Hillsdale, NJ: Erlbaum.

McGuire, W. J., & McGuire, C. V. (1991). The content, structure and operation of thought systems. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 4, pp. 1–78). Hillsdale, NJ: Erlbaum.

Newtson, D. A. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*, 28–38.

Newtson, D. A. (1976). Foundations of attribution: The perception of ongoing behavior. In J. Harvey, W. Ickes, & R. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 223–247). Hillsdale, NJ: Erlbaum.

Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, *11*, 107–123.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* New York: Cambridge University Press.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*, 42–55.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, *51*, 242–258.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14*, 521–533.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the story model for juror decision making. *Journal of Personality and Social Psychology*, *62*, 189–206.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*, 1–24.

Radvansky, G. A., Spieler, D. H., & Zacks, R. T. (1993). Mental model organization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 94–114.

Radvansky, G. A., Wyer, R. S., Jr., Curiel, J. M., & Lutz, M. F. (1997). Mental models and abstract relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1233–1246.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*, 940–953.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Rokeach, M. (1960). *The open and closed mind*. New York: Basic Books.

Rosen, N. A., & Wyer, R. S., Jr. (1972). Some further evidence for the "Socratic effect" using a subjective probability model of cognitive organization. *Journal of Personality and Social Psychology*, *24*, 420–424.

Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *Journal of Abnormal and Social Psychology*, *53*, 367–372.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63–77.

Sanna, L. J., & Schwarz, N. (2003). Debiasing the hindsight bias: The role of accessibility experiences and (mis)attributions. *Journal of Experimental Social Psychology*, *39*, 287–295.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: The real story. In R S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 8, pp. 1–85). Hillsdale, NJ: Erlbaum.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*, 36–71.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, *2*, 87–99.

Schwarz, N. (2004). Meta-cognitive experiences in consumer judgment and decision-making. *Journal of Consumer Psychology*, *14*, 332–348.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: A handbook of basic principles* (pp. 433–465). New York: Guilford Press.

Scott, W. A. (1963). Conceptualizing and measuring structural properties of cognition. In O. J. Harvey (Ed.), *Motivation and social interaction.* New York: Ronald Press.

Scott, W. A. (1969). Structure of natural cognitions. *Journal of Personality and Social Psychology*, *12*, 261–278.

Sentis, K. P., & Burnstein, E. (1979). Remembering schema-consistent information: Effects of a balance schema on recognition memory. *Journal of Personality and Social Psychology*, *37*, 2200–2212.

Shah, J. Y., & Kruglanski, A. W. (2000). The structure and substance of intrinsic motivation. In C. Sandone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 105–127). San Diego, CA: Academic Press.

Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: Bottom-up priming of goals by means and its effects on self-regulation. *Journal of Personality and Social Psychology*, *84*, 1109–1122.

Shah, J. Y., Kruglanski, A. W., & Friedman, R. (2003). Goal systems theory: Integrating the cognitive and motivational aspects of self-regulation. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 247–275). Mahwah, NJ: Erlbaum.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Sherman, J. W., & Klein, S. B. (1994). Development and representation of personality impressions. *Journal of Personality and Social Psychology*, *67*, 972–983.

Sherman, S. J., Ahlm, K., Berman, L., & Lynn, S. (1978). Contrast effects and the relationship to subsequent behavior. *Journal of Experimental Social Psychology*, *14*, 340–350.

Shrum, L. J. (2002). Media consumption and perceptions of social reality: Effects and underlying processes. In J. Bryant & D. Zillmann (Eds.), *Media effects: Advances in theory and research* (pp. 69–95). Mahwah, NJ: Erlbaum.

Simon, H. A. (1957). *Models of man.* New York: Wiley.

Smith, E. R. (1984). Model of social inference processes. *Psychological Review, 91,* 392–413.

Smith, E. R. (1990). Content and process specificity I the effects of prior experiences. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 3, pp. 1–59). Hillsdale, NJ: Erlbaum.

Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 99–151). Hillsdale, NJ: Erlbaum.

Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindsey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 391–445). New York: McGraw-Hill.

Smith, E. R., & deCoster, J. (1998). Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology, 74,* 21–35.

Smith, E. R., & deCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-processing models. In S. Chaiken & Y. Trope (Ed.), *Dual-process theories in social psychology* (pp. 323–336). New York: Guilford Press.

Spiro, R. J. (1977). Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp.). Hillsdale, NJ: Erlbaum.

Srull, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 440–463.

Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11,* 316–345.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660–1672.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review, 96,* 58–63.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). Conceptual/integrative complexity. In C. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 401–418). Cambridge, MA: Cambridge University Press.

Taylor, H. A., & Tversky, B. (1992). Spatial situation models derived from survey and route descriptions. *Journal of Memory and Language, 31,* 261–292.

Taylor, S. E., & Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249–288). New York: Academic Press.

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331–376). New York: Academic Press.

Trafimow, D. A., & Wyer, R. S., Jr. (1993). Cognitive representation of mundane social events. *Journal of Personality and Social Psychology, 64,* 365–376.

Tulving, E. (1983). *Elements of episodic memory.* Oxford, UK: Clarendon Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.

Tversky, A., & Kahneman, D. (1982). Causal schemas in judgments under uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117–128). New York: Cambridge University Press.

Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology, 3,* 73–79.

Wänke, M., & Wyer, R. S., Jr. (1996). Individual differences in person memory: The role of sociopolitical ideology and in-group versus out-group membership in responses to socially relevant behavior. *Personality and Social Psychology Bulletin, 22,* 742–754.

Wyer, R. S., Jr. (1964). Assessment and correlates of cognitive differentiation and integration. *Journal of Personality, 32,* 495–509.

Wyer, R. S., Jr. (1974). *Cognitive organization and change: An information-processing approach.* Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr. (2004). *Social comprehension and judgment: The role of situation models, narratives and implicit theories.* Mahwah, NJ: Erlbaum.

Wyer, R. S., Jr., Adaval, R., & Colcombe, S. J. (2002). Narrative-based representations of social knowledge: Their construction and use in comprehension, memory and judgment. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp.131–197). San Diego, CA: Academic Press.

Wyer, R. S., Jr., & Bodenhausen, G. V. (1985). Event memory: The effects of processing objectives and time delay on memory for action sequences. *Journal of Personality and Social Psychology, 49,* 304–316.

Wyer, R. S., Jr., Bodenhausen, G. V., & Gorman, T. F. (1985). Cognitive mediators of reactions to rape. *Journal of Personality and Social Psychology, 48,* 324–338.

Wyer, R. S., Jr., & Budesheim, T. L. (1987). Person memory and judgment: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology, 53,* 14–29.

Wyer, R. S., Jr., Budesheim, T. L., & Lambert, A. J. (1990). Person memory and judgment: The cognitive representation of informal conversations. *Journal of Personality and Social Psychology, 58,* 218–238.

Wyer, R. S., Jr., Budesheim, T. L., Lambert, A. J., & Swan, S. (1994). Person memory and judgment: Pragmatic influences on impressions formed in a social context. *Journal of Personality and Social Psychology, 66,* 254–267.

Wyer, R. S., Jr., & Carlston, D. E. (1979). *Social cognition, inference and attribution.* Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., & Carlston, D. E. (1994). The cognitive representation of persons and events. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 41–98). Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., Clore, G. L., & Isbell, L. M. (1999). Affect and information processing. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 31, pp. 1–77). San Diego, CA: Academic Press.

Wyer, R. S., Jr., & Goldberg, L. (1970). A probabilistic analysis of the relationships among beliefs and attitudes. *Psychological Review, 77,* 100–120.

Wyer, R. S., Jr., & Hartwick, J. (1980). The role of information retrieval and conditional inference processes in belief formation and change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 241–284). New York: Academic Press.

Wyer, R. S., Jr., & Hartwick, J. (1984). The recall and use of belief statements as bases for judgments: Some determinants and implications. *Journal of Experimental Social Psychology, 20,* 65–85.

Wyer, R. S., Jr., Lambert, A. J., Budesheim, T. L., & Gruenfeld, D. (1992). Theory and research on person impression formation: A look to the future. In L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 3–36). Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychological Review, 106,* 89–118.

Wyer, R. S., Jr., & Srull, T. K. (1989). *Memory and cognition in its social context.* Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., & Unverzagt, W. H. (1985). The effects of instructions to disregard information on its subsequent recall and use in making judgments. *Journal of Personality and Social Psychology, 48,* 533–549.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

CHAPTER 13

Standards

MONICA BIERNAT
SCOTT EIDELMAN

Standards are relevant to nearly every field of inquiry, to every domain of life. There are standards for admission to schools and to professions, for promotions and other evaluations in schools and the workplace, for diagnosing illnesses and assessing recovery, and for determining guilt in the legal system and progress in society. Indeed, it seems difficult—if not impossible—to imagine a world without points of reference against which we might judge and evaluate. In this chapter we focus our attention on *social* standards—individual people or groups—as they affect judgments or evaluations of others and the self. Despite this restricted focus, we hope to highlight general principles by which standards exert influence on judgment and behavior.

Higgins (1990) provides a useful definition of the key construct with which we are concerned: "a standard is a criterion or rule established by experience, desires, or authority for the measure of quantity and extent, or quality and value" (p. 302). A *social standard* in particular may be thought of as "any attribute of a person or of a collection of people that serves as a point of comparison for an individual" (Miller & Prentice, 1996, p. 800). This chapter briefly examines how standards are selected but primarily considers the consequences of the comparison process—that is, the judgment outcome of comparison to some standard. This focus on comparison means that we are concerned with the *relativity* of judgment. Beginning with research on such diverse areas as psychophysics (Helson, 1947, 1964; Parducci, 1956; Postman & Miller, 1945; Stevens, 1957; Volkman, 1951; Wever & Zener,

1928) and the self (James, 1890/1948), psychologists have long emphasized the relativity of intra- and interpersonal experience. That judgment is relative or comparative appears in research on attitudes (Sherif & Hovland, 1961) and decision making (Houston & Sherman, 1995; Kahneman & Miller, 1986), as well as in two domains that are given attention in this chapter: person perception (e.g., Higgins, Rholes, & Jones, 1977) and self-evaluation (Festinger, 1954a, 1954b).

One central theme pervades this chapter: that the outcome of a comparative process can be conceptualized in terms of either *assimilation* or *contrast*. Assimilation occurs when the target of evaluation (e.g., another person or the self) is pulled toward or judged consistently with the standard or expectation, and contrast occurs when the target is differentiated from (judged in a direction opposite) the comparative frame. By "judgment," we typically refer to a connotative evaluation (e.g., a rating on some trait dimension) but also to outcomes that are affective (such as mood or self-esteem) or behavioral (e.g., task performance) in nature as well. Delineating when assimilation versus contrast occurs represents an important goal of this review.

We begin with a brief overview of how standards might be selected in a given judgment situation or, alternatively, thrust on the perceiver. We then review some recent models of standard use and its consequences, and finally we turn to illustrations of how standards are used in self- and other judgment. Here we consider the role of self as standard in judging others, the role of stereotypes

as standards in judging others, and the role of internalized guides and other people to judge the self. Though selective, we hope this review illustrates some basic principles about how standards influence judgment in everyday life.

SELECTION OF STANDARDS

Though the use of standards is ubiquitous, relatively little research attention has been paid to how standards are selected from among the myriad possibilities that exist (cf. Miller & Prentice, 1996). One reason for this lack of attention is simple: Standards are often encountered rather than chosen (Wood, 1989). For example, hopeful undergraduates are not afforded the luxury of choosing what criteria will determine admission into graduate school. And social standards may be imposed (or impose themselves) on individuals. The job candidates who caught a glimpse of the competition ("Mr. Clean" or "Mr. Dirty") in Morse and Gergen's (1970) classic study did not ask to do so, nor did they postpone drawing inferences about their own likelihood of success as a result of this encounter.

As this last example suggests, comparison to standards is often a relatively automatic process (Dunning & Hayes, 1996; Gilbert, Giesler, & Morris, 1995; Stapel & Blanton, 2004), which is to say that standards are often used without intent or awareness (Bargh, 1994). For example, Herr (1986) found that presentation of stimuli outside participants' awareness affected their judgments of an ambiguous target: Moderate stimuli produced assimilation to the standard, and extreme stimuli produced contrast. Mussweiler, Rüter, and Epstude (2004) found similar results in the domain of self-judgment, with participants rating themselves as more or less athletic when subliminally exposed to moderate and extreme standards, respectively. Research by Stapel and Blanton (2004) went even further by showing that subliminal exposure to standards affected *implicit* self-evaluation. When primed with a low standard, the size of participants' signatures (an implicit measure of self-evaluation) became larger; a high standard prompted smaller signatures. In other words, standards—in this research, Albert Einstein or a clown—affected self-evaluation in the absence of any intent to form a judgment. Even when judgment is deliberate, standards used may be unknown or inaccessible (cf. Nisbett & Wilson, 1977; see Parducci, 1983, for a discussion in the realm of psychophysics).

Which standards are used in judgment is often based on the preceding and current contexts in which we find ourselves. Helson's (1964) adaptation-level theory is a case in point. According to Helson, judgment is made with reference to a psychological neutral point (the adaptation level), essentially the average of one's past experiences along the dimension in question. Of course, not all past experience carries equal weight. Factors such as recency and frequency may determine the impact of particular standards on judgment (Avant & Helson, 1973), and judgments may be based on some combination of both the most recent context and the normative standard established by the average of one's past experiences (e.g., see Higgins & Lurie, 1983; Higgins & Stangor, 1988). Judgments may also be influenced by some combination of both local and general norms (Giladi & Klar, 2002; Klar & Giladi, 1997).

Not all perspectives, however, emphasize precomputed points of reference. Norm theory (Kahneman & Miller, 1986), for example, emphasizes the construction of standards *after* events have occurred (see also Goethals, Messick, & Allison, 1991; Taylor, Wood, & Lichtman, 1983). According to norm theory, stimuli evoke their own frame of reference by either recruiting specific exemplars from memory or constraining mental simulation. In each case, counterfactual alternatives are evoked in response to instigating stimuli, and the stimuli are then compared against these postcomputed alternatives.

To this point, we have considered the selection of standards to be a fairly passive process. Yet people have goals, and standards are often chosen with these goals in mind. For example, when selecting one of two possible partners for an upcoming problem-solving task, participants in a study by Ditto and Lopez (1992, study 1) required more evidence of intelligence when an unlikable target was the obvious candidate compared to when the obvious candidate was likable. Because the two possible partners differed in their desirability (irrespective of their intelligence), differential standards were set based on preferred outcomes (see also Sanitioso & Kunda, 1991). Similarly, when deciphering the results of a self-administered medical test, participants allowed more time to pass (i.e., used a higher standard) when led to believe that no change in the color of a test strip indicated the presence of an enzyme deficiency than when led to believe this lack of color change indicated normal enzyme presence (Ditto & Lopez, 1992, study 2). Standards for judging others have also been linked to current needs. Esteem threats, for example, prompt egocentric contrast; Beauregard and Dunning (1998) found that participants' scores on an aptitude test were negatively related to their judgments of a target's intelligence, but only when self-esteem first was challenged. By setting more stringent intelligence standards, high-scoring participants protected their accomplishments by making their performance more unique and harder to obtain. In contrast, the less stringent standards set by low-scoring participants meant that they could also "qualify" for the label "intelligent."

These examples point to *quantitative* differences in standard setting based on particular goals. These same goals may also be reflected in *qualitative* differences in the selection of standards. The vast literature on social comparison theory offers a number of examples. It is widely known that Festinger (1954a, 1954b) posited a basic drive to assess one's abilities and opinions, and that in the absence of objective standards such needs would be best met by comparison with similar others. Thus, social standards were thought to be chosen based on their diagnosticity (see also Wheeler, Martin, & Suls, 1997).

Though Festinger was somewhat vague as to what constituted similarity, others specified that people should select for comparison those who are similar along attributes related to the dimension being judged (Goethals & Darley, 1977). This "related attributes" hypothesis has received empirical support (e.g., Gastorf & Suls, 1978; Miller, 1982; Wheeler, Koestner, & Driver, 1982; Zanna, Goethals, & Hill, 1975).[1]

Motives besides self-assessment have also been shown to drive social comparison (Helgeson & Mickelson, 1995; Taylor & Lobel, 1989), with selection of relevant standards following suit. For example, needs for self-enhancement may motivate comparison with those less well off (Hakmiller, 1966; Wills, 1981; Wood, Taylor, & Lichtman, 1985). Perhaps counterintuitively, comparison with superior others may also foster self-enhancement, provided that we see ourselves as similar to these outperforming others (Collins, 1996; Thornton & Arrowood, 1966; Wheeler, 1966). Self-improvement goals are also associated with the selection of upward comparison targets (Lockwood & Kunda, 1997; Taylor & Lobel, 1989; Wood, 1989).

With so many standards available, one frequent motive may be efficiency, such that standards routinely used in judgment continue to be selected. When evaluating others, participants seem to consistently invoke one well-known and well-used standard; specifically, the self (Dunning & Hayes, 1996; Holyoak & Gordon, 1983). Close friends also become more accessible following self-evaluation, implying their role as a frequent source of comparison (Mussweiler & Rüter, 2003, study 1). Interestingly, routine standards may even be used in spite of the fact that they are often uninformative (e.g., when they are dissimilar along the critical judgment dimension) (Mussweiler & Ruter, 2003, study 3; cf. Festinger, 1954a, 1954b).

People also seem inclined to choose specific *dimensions* on which to be judged (Wood, 1989), as well as contexts in which particular standards may be more or less relevant. For example, participants often express favor for dimensions on which they excel (Tesser & Campbell, 1980; Van Knippenberg, Wilke, & de Vries, 1981; see also Van Knippenberg, 1978). Consider research by Lemaine (1974), who asked two groups of boys at a summer camp to build huts. One group was intentionally disadvantaged with inadequate building supplies. Not to be outdone, these boys fought to gain recognition along a new dimension—the garden they grew next to their second-rate hut. Research with women suffering from breast cancer also indicates intent to change or narrow dimensions of comparison so that negative mood or self-assessment can be deflected (Taylor et al., 1983). Because local contexts instantiate standards, these too may be strategically chosen (Crocker & Major, 1989; Davis, 1966; Frank, 1985). As Gilbert and colleagues (1995) note, though comparisons often arise spontaneously, people often have the choice over which standards they encounter via the contexts they enter. In short, though standards may often be thrust on us, we also play an active role in selecting and construing those standards.

MODELS OF STANDARD USE AND CONSEQUENCES

In this section, we highlight relatively recent models of standard use that have been developed in the literature on social judgment. But first, a quick mention of some major contributions to understanding *psychophysical* and *attitudinal* judgments is warranted. Helson's (1947, 1964) notion of adaptation level, described earlier, notes that the standard in a judgment setting is the "point of perceived neutrality" or, technically, the weighted logarithmic mean of previous stimulation on some dimension. In other words, the context (past experience) provides the standard. Essentially Helson's model is about *contrast*– that is, judgment (and perception) shifts away from the adaptation level, as when a "moderately bright" room appears *very bright* to one who has been adapted to the dark, and *very dim* to one adapted to the bright sunlight.

Volkmann's (1951) "rubberband model" introduced the notion of a frame of reference: Judges set endpoints of response scales to capture the stimulus range they expect to encounter. Anchoring and reanchoring is done as the stimulus range extends and restricts (see Parducci, 1956; Postman & Miller, 1945). The context again provides the standard, but not in the perceptual manner described by Helson. Instead, *cognitions* determine the range of stimuli perceivers *expect* to encounter, and these are translated into judgmental frames of reference. Parducci (1963) incorporated much of the rubberband model in the "range" principle of his model but also invoked a "frequency" principle as well, which suggested that judges use judgments to reflect the frequency—the rank-order position—of items in the stimulus range. For both Volkmann and Parducci, judgments of stimuli are generally contrasted from extreme anchors: With the extension of a range in one direction, judgments move toward the other.

In attitudinal judgment, the idea of own attitude as an anchor was introduced by Hovland and Sherif (1952). That is, one's own position on an issue serves as an anchor so that all other attitudinal positions are judged relative to it. This ultimately led to the development of the social judgment or "assimilation-contrast" model (Sherif & Hovland, 1961), which posited that stimuli *close to* the anchor (i.e., within the "latitude of acceptance") are assimilated, whereas those further away from the anchor (within the "latitude of rejection") are contrasted.[2] Upshaw's (1962, 1969) variable perspective model also took some of Volkmann's ideas and extended them to the domain of attitudinal judgment. Response scales are anchored to the expected range of stimuli on a dimension; this is the judge's "perspective." But if the judge's own position on the dimension is outside the range, the perspective gets extended to encompass his or her position, and the result is enhanced contrast effects.

These early models—along with work by Fechner (1860), Stevens (1957), Thurstone (1928), and Anderson (1974), among others—set the stage for a number of the modern judgment perspectives reviewed below (see Eiser, 1990, for more detail on early judgment

models). These modern models again ask the central question—when assimilation and when contrast?—and they focus on social judgment—judgments of others and the self.

The Set–Reset Model

The set–reset model (Martin, 1986; Martin & Achee, 1992; Martin, Seta, & Crelia, 1990), like many to be reviewed in this section, focuses on context as standard and considers the overlap that exists between an ambiguous target and this context. In a process termed "setting," a judge's representation of an ambiguous target can be pulled toward that overlapping context. In essence, assimilation occurs because reactions to the context can be mistaken for reactions to the target. But if the judge realizes he or she is affected by context, an attempt may be made to avoid using the context—to partial out or subtract its effects. However, because of the overlap between the context and target, this correction may involve subtracting out some of the "true" reaction to the target. The result of this resetting process is reduced assimilation or *contrast* from the context. Contrast effects based on resetting are posited to require more cognitive effort than assimilation effects (Martin et al., 1990).

This model suggests several features that determine the use of context for setting versus resetting. For setting to happen, the target must be ambiguous and the context must be relatively broad and indistinct, such that overlap between context and target can occur. Resetting requires some sort of cue that the effects of the context should be partialed out of the evaluation of the target. For example, Martin (1986) found that an ambiguous target was assimilated to valenced primes, but only when participants were led to believe the priming procedure was interrupted—that is, the effect of the prime was maintained over the interruption and continued to "set" the target. But when the task was completed—signaling separation of prime and target—resetting was triggered, resulting in more favorable ratings in the negative prime condition and less favorable ratings in the positive prime condition (contrast). Other cues to reset include heightened awareness of the context and its influence (Lombardi, Higgins, & Bargh, 1987; Strack, Schwarz, Bless, Kübler, & Wänke, 1993) as well as direct instruction or conversational norms (e.g., "Don't let today's weather affect your mood estimate"; Schwarz & Clore, 1983).

The Inclusion–Exclusion Model

The basic prediction of the inclusion–exclusion model (Schwarz & Bless, 1992a, 1992b) is simple: that "assimilation effects are likely to emerge when the target stimulus and the context stimuli are assigned to the same category, whereas contrast effects may emerge when they are assigned to different categories" (Schwarz & Bless, 1992a, p. 218). The former process—jointly categorizing context and target—reflects *inclusion*, and the latter reflects *exclusion*. In this model, inclusion is the default

mode, as long as the context includes information that is potentially relevant to the representation of the target (obviously irrelevant information is ignored). But if features of the task or situation at hand suggest that contextual information should not be used, "exclusion" or subtraction occurs, resulting in contrast. Inclusion–exclusion is highly similar to setting–resetting (Martin, 1986), and indeed both of these models can be thought of as partialing models, which involve a relatively automatic assimilation mode and a more effortful correction process.

Schwarz and Bless's model is more specific, however, in its emphasis on categorization processes as the key contributor to assimilation and contrast effects (see also Tajfel & Wilkes, 1963). It also offers the elegant organizing principle that inclusion = assimilation and exclusion = contrast. Factors that contribute to exclusion (vs. inclusion) include *extreme*, *distinct*, and *narrow* contexts. The inclusion–exclusion model also adds the proposition that excluded information can serve as a standard of comparison. And because excluded information is likely to constitute an *extreme* standard, this use of the context as standard can contribute to a contrast effect. Thus, two forms of contrast are posited: *exclusion* (correction) or *comparison*, and which occurs, depends on whether the context stimulus is directly *linked* to the dimension of judgment. If so, it will serve as a standard of comparison. If not, it is merely subtracted from the representation but is unlikely to be used as a standard (Schwarz, Münkel, & Hippler, 1990).

The Flexible Correction Model

The previous two models assume that assimilation is the default judgment outcome, and that it reflects a less effortful mode of processing. In contrast, the flexible correction model (Petty & Wegener, 1993; Wegener & Petty, 1995, 1997) suggests that *either* assimilation or contrast may be the default; one is not more effortful or more likely than the other. This approach suggests that perceivers have naive theories about how contexts might affect their subsequent judgments and that they then engage in steps to correct these biases (preemptively, online, or after the fact). If a theory suggests that a context may produce assimilation, those who are motivated and able to do so will correct their judgments away from this assimilative bias. But if a theory suggests a context will produce contrast, correction will occur away from this contrastive bias. For example, most people assume that first considering a dream vacation spot will likely bias judgments of "average" locations in a contrastive fashion (Indianapolis seems less appealing after thoughts of Paris), and indeed judgments show this pattern (Wegener & Petty, 1995). But when cued to *not* let prior judgments affect later judgments, the contrastive tendency is reduced and slightly reversed. This model, then, argues for either assimilation or comparison contrast as a basic outcome of judgment, and for *corrective* contrast or assimilation when motivation and ability to correct are high.

The Interpretation–Comparison Model

Stapel and colleagues have conducted an array of studies pointing to features that contribute to assimilation versus contrast effects, and their general approach can be characterized in terms of an interpretation–comparison model (ICM; Stapel & Koomen, 1998, 2000; Stapel & Winkielman, 1998). According to their model, accessible and applicable contextual cues are likely to be used as an interpretation frame and produce assimilation effects, whereas distinct contextual cues function as comparison standards and produce contrast. Thus, interpretation = assimilation, comparison = contrast (see Higgins, 1996; Stapel & Koomen, 1998).

In general, this model requires some kind of similarity between the target and the context or standard for any type of context effect to emerge. This idea in itself is not novel—in Brown's (1953) classic study of weight estimates, judges showed contrast from extreme weight anchors of the same type as the targets, but not from an equally heavy *tray* that was also lifted. Brown (1953) stated: "The anchor, to be effective, must be perceived as a member of the same class" as the target (p. 210; see also Parducci, Knobel, & Thomas, 1976; Suls & Wills, 1991; Upshaw & Ostrom, 1984). Stapel and colleagues note that given similarity, another feature that determines whether interpretation (assimilation) or comparison (contrast) occurs is the *distinctness* of the standard or context (Stapel, Koomen, & van der Plight, 1996, 1997; see also Helson, 1964). *Distinct* contexts serve as comparison standards, as do *extreme* contexts (e.g., Herr, 1986). On the other hand, the use of primed/contextual information as an interpretation frame is enhanced by *nondistinct*, *moderate*, and/or *broad* contextual cues.

Research examining the impact of *distinctness* versus *breadth* has typically involved manipulations in which either broad, abstract behavior labels are primed or specific actor–trait links or *exemplars* are primed. In general, exemplar primes produce *contrast*, whereas broad or abstract primes produce assimilation (e.g., Stapel & Koomen, 1996; Stapel & Schwarz, 1998). Comparable effects have also emerged in research examining behavioral responses to primes or context. In a study that might be called a modern classic, researchers found that priming of an elderly stereotype (through exposure to broad terms associated with that category) led individuals to walk more slowly as they left an experiment—a behavioral assimilation effect (Bargh, Chen, & Burrows, 1996). In follow-up research, however, others found that priming of *exemplars* (e.g., Albert Einstein) rather than broad categories (professor) produced behavioral contrast effects (Dijksterhuis et al., 1998; see also Dijksterhuis, Spears, & Lépinasse, 2001).

Central to the ICM is the fact that narrow primes are more likely to be used as comparative standards of judgment that exclude the target, whereas wide or abstract primes serve as interpretive frames in which the target can be included (Schwarz & Bless, 1992a; Stapel & Koomen, 1996, 1997). Indeed, it is worth noting that the concept of an interpretation frame is quite similar to the inclusion principle in Schwarz and Bless's (1992a)

inclusion–exclusion model. That is, the same principles that lead a context to serve as an interpretation frame also lead the target to be included in the context, with assimilation as the likely result. However, the comparison standard portion of the ICM is *not* analogous to Schwarz and Bless's exclusion principle (which implies more active correction) but does resemble the separate comparison process posited by Schwarz and Bless (see above). The ICM also does not make any claims about the default context effect, or the relative ease/effort involved in these outcomes. The key contribution of the ICM is its broad characterization of two functions that context can serve—the interpretation/framing function versus the comparison function.

The Selective Accessibility Model

The selective accessibility model (SAM), recently formalized by Mussweiler (2003a, 2003b), highlights how knowledge made accessible through comparison processes influences judgments of targets. Specifically, the model suggests that there are three phases involved in rendering a comparative judgment. First, a standard must be selected; second, the comparison must occur—one assesses the similarity/dissimilarity of the target object to the standard—and third, knowledge made accessible by the earlier phases must be integrated to produce an evaluation of the target.

Although the standard selection phase is important, the key contribution of the SAM lies in its articulation of what occurs in subsequent steps. The comparison step involves an initial, holistic assessment of similarity between the target and the standard, and based on that decision, the judge then engages in either similarity or dissimilarity testing. This testing then leads to the selective accessibility of evidence consistent with the hypothesis. That is, if one is testing the hypothesis that the target = the standard, standard-consistent knowledge of the target will become accessible. And conversely, when one tests the hypothesis that the target ≠ the standard, standard-*in*consistent knowledge about the target becomes accessible.

In the final judgment stage, the SAM predicts that similarity testing and the resultant accessibility of standard-consistent knowledge about the target leads to *assimilation*, whereas dissimilarity testing and the resultant accessibility of standard-inconsistent knowledge leads to *contrast*. In the SAM, then, the direction of context effects depend on whether accessible information makes the target seem like or unlike the standard.

Mussweiler's model also acknowledges that the comparison process, in addition to producing selective accessibility effects, provides a reference point—a standard—against which the accessible knowledge can be compared. Thus, considering whether one is adjusting well to college compared to a very well-adjusted student might activate knowledge of one's own good adjustment (and therefore prompt assimilation) at the same time that a high standard is invoked (facilitating contrast). Mussweiler (2003b) notes that "reference point use may well constitute an additional mechanism that influences

the evaluative consequences of comparison independently of selective accessibility . . . the same comparison may involve both the assimilative tendencies of selective accessibility and the contrastive tendencies of reference point use" (p. 483). This theme will be taken up in the next judgment model to be considered.

The Expectation and Contrast Model

A number of social judgment models suggest that assimilation is a relatively effortless process, whereas contrast requires both motivation and capacity to proceed (e.g., Martin et al., 1990; see similar themes in Gilbert, Pelham, & Krull's, 1988, stage model of attributional judgment). In the flexible correction model, *either* assimilation or contrast is the default, depending on naive theories about the impact of context. In contrast to these perspectives, Manis and Paskewitz (1984a, 1984b) suggest that contrast and assimilation operate in parallel, with no exclusive processing advantage for either: Judgments are affected by both *expectational* influences and *comparison* strategies. For example, a priming experience or past exposure to relevant stimuli can create an expectation about what future experiences will be like, promoting assimilation; at the same time, the prime creates an anchor against which subsequent stimuli are judged, promoting contrast (see Manis, Biernat, & Nelson, 1991). The net judgment result depends on which tendency is strongest.

In an illustrative study, participants were first exposed to a series of word definitions that indicated either high or low pathology and were asked to indicate which had been produced by schizophrenic patients (Manis & Paskewitz, 1984a). This induction series was designed to create *expectations* about the kinds of definitions likely to be seen later in the study, and indeed, measures indicated that those in the high pathological induction condition expected more psychopathology in a subsequent group of targets. When participants then judged a test series of moderately psychopathological definitions, a marked contrast effect emerged, with the moderate definitions judged more pathological by those originally exposed to the nonpathological induction set. Furthermore, expectations were positively correlated with judgments. Thus, the induction phase produced corresponding expectations that positively predicted judgments, at the same time these judgments were *contrasted* from that initial context.

The expectation and contrast model does not speak directly to specific factors that affect assimilation and contrast effects. However, implicit in the model is that context effects in general require that an unambiguous—perhaps *extreme*–expectation be induced, and contrast effects require that the target be *ambiguous or moderate* on the dimension of interest, or discrepant from the context. In addition, Manis and Paskewitz (1984a) propose that the time course of the expectation and contrast paths operate differently; namely, the contrastive tendency may decline more rapidly than the assimilative one. Indeed, with a time delay introduced between a contextual induction of expectations and target judgments, assimilation rather than contrast is a typical result (Manis

& Blake, 1963; Manis & Moore, 1978). Perhaps the most significant aspect of this model is its premise that assimilative and contrastive tendencies work in tandem; that they may "derive from the same eliciting experience" (Manis et al., 1991, p. 210).

The Reflection and Evaluation Model

Another general model of comparative thinking, the reflection and evaluation model (REM), was recently proposed by Markman and McMullen (2003). We describe this model last because it shares much in common with the inclusion–exclusion model (Schwarz & Bless, 1992a) and Mussweiler's (2003b) SAM. It also picks up the theme of simultaneity in comparative processes articulated by Manis and Paskewitz (1984a, 1984b). However, the heart of this model is the notion of *mental simulation*— "the consideration of alternatives to present reality" (p. 244; see Kahneman & Miller, 1986), and the theory is concerned primarily with *self*-judgments (counterfactual comparison research providing the main empirical base). In general, the model proposes two modes of mental simulation that operate in parallel. The *reflection* mode involves "as if" thinking, in which one "simulates that information about the comparison standard is true of, or part of, the self" (Markman & McMullen, 2003, p. 245). Note that this mode is quite similar to Mussweiler's notion of similarity testing, and that the likely outcome of this process is also *assimilation*. The second mode of simulation is *evaluation*, in which the standard operates as a point of reference against which the self is evaluated, likely leading to *contrast*. Markman and McMullen note that "reflection occurs when information about the standard is included in one's self-construal, and evaluation occurs when such information is excluded" (p. 245), a statement that highlights the strong connection to Schwarz and Bless's (1992a) inclusion–exclusion model.

One distinction between "reflection" in the REM and "selective accessibility" in the SAM is that the information one comes up with in reflection may be *imaginary* rather than selected from an assortment of facts about the self. Nonetheless, imagination brings standard-consistent cognitions to the fore of thought. And rather than the biased hypothesis-testing mechanism posited by Mussweiler, the imagination effect is similar to findings in the imagination–explanation literature, whereby coming up with explanations for possible future events increases estimates of the probability those events will occur (Koehler, 1991; Ross, Lepper, Strack, & Steinmetz, 1977).

Also distinguishing the REM from most other models is the notion of parallel yet independent processes of reflection and evaluation. The ultimate outcome of these processes (i.e., whether assimilation or contrast occurs) depends on the relative strength of these processes, which, in turn, is based on "the extent to which contextual features encourage one to think about the self and the standard together, as a single unit or entity (i.e., inclusion), or the extent to which one thinks about the self and the standard separately, as two distinct entities (i.e., exclusion)" (Markman & McMullen, 2003, p. 249). That is,

context will determine whether a mode of inclusion or exclusion operates. What are the contextual features that matter here? Many of the features described earlier: Similarity between standard and target or temporal closeness will prompt the reflection process and assimilation; distinctness of the standard or immutability of the self will prompt evaluation and contrast. Markman and McMullen (2003) also note that attentional focus is key: To the extent that context focuses one on the standard itself, reflection is likely, but to the extent the focus is on the *comparison* between the self and standard, evaluation is likely (see McMullen, 1997).

In general, then, the REM can be thought of as a hybrid of the inclusion–exclusion model (Schwarz & Bless, 1992a), the ICM (Stapel & Koomen, 1998, 2000), and the SAM (Mussweiler, 2003a, 2003b), with a hint of expectation–contrast thrown in (Manis & Paskewitz, 1984a, 1984b). The authors suggest that contextual factors affecting inclusion–exclusion set the processes of reflection and evaluation in motion. These processes make different kinds of information accessible and result in assimilation or contrast, in judgment and behavior.

Summary

The models reviewed above have in common the idea that a contextual cue can function as either an interpretive framework or a reference point. Factors that make the context less distinct from the target (i.e., allow for inclusion of the target in the context) facilitate the *interpretation* framework, whereas those that make the context distinct (allow for exclusion) generally facilitate the use of context as a *reference point* that prompts contrast. The models are largely cognitive in nature and largely silent on the issue of how motivational variables may contribute to judgment outcomes (cf. Higgins, Chapter 19, this volume). They also differ in important ways; in terms of the assumed default process (assimilation or contrast), the sequential versus simultaneous nature of the process, the assumed effort involved in contrast (i.e., is it a basic perceptual effect or part of an effortful correction process?), and the extent to which process variables are articulated (e.g., similarity–dissimilarity testing in the SAM; as-if thinking in the REM). Assimilation and contrast effects both occur under different conditions, and the models described here have been able to predict and explain a large array of findings. Integrating the models fully would be an arduous task, and no doubt each model may have its own special sphere of relevance. What seems more important is that these models offer a general framework for predicting and understanding the when and why of assimilation and contrast.

ILLUSTRATIONS OF STANDARD USE IN SELF- AND OTHER JUDGMENT

The previous section reviewed general models of social judgment. But precisely how is standard use reflected in everyday evaluation of self and others? In this section, we provide a selective sampling of specific ways in which standard use affects judgments of others and the self. We consider the *self* and *stereotypes* as standards for judging others and a variety of standards used to judge the self.

The Self As a Standard in Judging Others

Psychologists have long accorded the self a prominent role in the judgment of others (e.g., Freud, 1924/1956; James, 1890/1948), often specifying that the self can serve as a standard to judge others (Combs, 1959; Heider, 1958; Krech & Crutchfield, 1948). For example, attitude theorists have argued that one's own attitudes serve as anchors in the evaluation and perception of other attitude stances (Eiser & Stroebe, 1972; Hovland, Harvey, & Sherif, 1957; Judd & Harackiewicz, 1980; Sherif & Hovland, 1961; Upshaw, 1962). Even Helson (1964), whose work emphasized sensory contributions to judgment, suggested that an individual's attitudes and beliefs can be included in his or her general "adaptation level," affecting subsequent experience and judgment. In this section, we review findings from two literatures, each indicating that the self is influential as a standard for judging others. However, as we will see, the form this judgment takes—either assimilation to or contrast from the self—depends on a host of variables, many reviewed in the models described earlier.

Using Our Own Attributes in Judgments of Other Individuals

In this section, we consider how the self's attributes may be used to judge others in situations in which there is no explicit call to compare. Research on attributive projection indicates a positive relationship between one's standing on a given trait and judgment of that trait in others (Dunning, 2003; Holmes, 1968). For example, self-relevant attributes appear in the free descriptions of others more often than do non-self-relevant attributes (e.g., Dornbush, Hastrof, Richardson, Muzzy, & Vreeland, 1965; Lemon & Warren, 1976; Shrauger & Patterson, 1976), and individuals who judge themselves extremely on attributes tend to rate ambiguous or attribute-consistent targets more highly on those traits than do individuals for whom the attribute is less relevant (e.g., Carpenter, 1988; Catrambone & Markus, 1987; Lewicki, 1983; Markus, Crane, Bernstein, & Siladi, 1982; Markus & Smith, 1981; Markus, Smith, & Moreland, 1985). Recently, it has been shown that we may also assume that others have goals similar to our own (Kawada, Oettingen, Gollwitzer, & Bargh, 2004).

Lambert and Wedell (1991) documented that this positive relationship between own and ambiguous other's standing on a trait holds even after controlling for dimension importance and affective response to the target. For example, self-described "sociable" participants judged an ambiguously sociable actor (who "installed a telephone answering machine" and "decided to join the Peace Corps") to be more sociable than did nonsociable participants, even when the importance of sociability and affective extremity in reactions to the target behaviors were controlled (Lambert & Weddel, 1991, study 1).

These findings demonstrate assimilation: Judgments of others are drawn toward the self. As suggested previously, this assimilative pattern is most pronounced when targets are ambiguous with respect to the trait in question (Green & Sedikides, 2001). Similarly, the type of judgment task also may affect the propensity to assimilate targets to the self. Factual tasks leave little room for the self to influence judgment of others. In contrast, more "conjectural tasks"—those that require inferential judgment—allow perceivers to go "beyond the information given" (Catrambone & Markus, 1987). Both target ambiguity and judgment tasks that are open to construal will reappear in this chapter as factors that make judgmental assimilation more likely.

However, some of this same literature points to circumstances under which differentiation or contrast of others from the self is more likely to occur. One condition seems to be that the target other is unambiguously low on the self-relevant trait. For example, Markus and Fong (1979, cited in Markus & Smith, 1981) reported that "independent schematics" judged an actor who never behaved independently as more dependent than did aschematic participants. Somewhat different interpretations have been advanced for these sorts of contrast effects. Markus and Smith (1981) interpret them in social judgment theory terms (Sherif & Hovland, 1961). To the extent that a target's actions fall outside an individual's latitude of acceptance (as do very dependent actions when one is an independent person), those actions are seen as distinctive and therefore contrasted from one's own position. However, Lambert and Wedell (1991) suggest that participants' judgments of such *un*ambiguous behaviors are not directly mediated by self-standing on the trait in question but by the extremity in one's evaluative reactions to the target behaviors. This explanation is more consistent with accentuation theory, which posits that reactions to stimuli (in this case, targets who behave dependently vs. not) are polarized in accordance with value connotations of those behaviors (Eiser, 1990; Eiser & Stroebe, 1972; Judd & Harackiewicz, 1980; Tajfel, 1957).

A somewhat different form of contrast effect is also evident in research on "egocentric definitions" of trait concepts (Dunning, 1993; Dunning & Cohen, 1992; Dunning, Meyerowitz, & Holzberg, 1989). In a series of studies, Dunning and his colleagues have consistently reported *negative* relationships between participants' objective standing on an attribute (e.g., SAT scores) and their subjective judgments of a target person (whose SAT scores are known) on that same attribute (e.g., intelligence; Beauregard & Dunning, 1998; Dunning & Cohen, 1992; Dunning & Hayes, 1996; Dunning & McElwee, 1995; Dunning et al., 1989; Dunning, Perie, & Story, 1991; see also Felson, 1990). For example, Dunning and Cohen (1992, study 4) found that the more hours of athletic activity participants engaged in per week, the less athletic they judged targets described as engaging in various hours of athletic activity themselves. This general contrast effect was strongest for low-athletic targets (those who spent just 1 hour per week in athletic pursuits). That is, low-athletic participants tended to judge these targets favorably on the athletic dimension, whereas high-athletic participants judged them unfavorably.

The type of contrast effect demonstrated by Dunning and his colleagues appears to be self-serving in nature: "Individuals will define traits in ways that reflect back favorably on themselves" (Beauregard & Dunning, 1998, p. 608). The high-performing individual—for example, one who scores 780 on the math SAT—will set high standards for math performance, thereby distinguishing among targets with lower scores and successively denigrating these targets as their scores decrease. The result is that one's own performance and underlying ability are favorably distinguished from others, presumably increasing one's sense of self-worth. The low-performing individual (e.g., one who scores 500 on the math SAT) will set lower standards for math performance such that he or she can be included among those "qualifying" for the trait (i.e., having math ability), with some regard for "reality constraints" (Kunda, 1990). Additional research has demonstrated that this "egocentric effect" is strongest among those participants who explicitly mention comparing the target to themselves (Dunning & Hayes, 1996), among those dispositionally high in self-esteem (who presumably have a positive self-image to protect; Dunning & Beauregard, 2001), and under conditions when self-esteem is threatened (e.g., as a result of failure on a task; Beauregard & Dunning, 1998).

It is worth noting an important distinction between the work of Dunning and his colleagues and that of Markus, Lambert, and Wedell and others. As indicated earlier, these latter researchers have primarily documented assimilative effects of self on judgments of others, whereas Dunning and his colleagues report contrastive results. A key difference seems to be that the self-attributes Dunning assesses are participants' specific performances or behaviors on the dimension of interest (e.g., SAT scores, hours of athletic activity, and times per week being late to class) rather than the more typically used global trait ratings (e.g., intelligence, athleticism, and punctuality). Consistent with the notion that *distinct* contexts trigger contrast, others may more easily be contrasted from *specific* self-referential standards but assimilated to *abstract* self-views. Indeed, in one study from Dunning's lab that relied on less distinct self-ratings (e.g., to what extent do you "make friends easily"?), judgments of others on the related trait (e.g., sociability) were assimilated to self-ratings (Beauregard & Dunning, 2001).

False Consensus and False Uniqueness

This discussion of self-as-standard for judging others has thus far skirted around two other relevant literatures: One on the false consensus effect—the tendency to overestimate the likelihood that others will act or feel as we do (e.g., Marks & Miller, 1987; Mullen et al., 1985; Ross, Greene, & House, 1977; Sherman, Pressin, & Chassin, 1984)—and one on the false uniqueness effect—the tendency to underestimate the commonality of our self-attributes (Campbell, 1986; Goethals et al., 1991; Marks, 1984; Suls & Wan, 1987). In general terms, these two

phenomena represent assimilation and contrast effects, respectively. Although clearly relevant to our previous discussion of the role of self in judging others, they have one distinction: False consensus and false uniqueness effects involve making predictions regarding the attitudes and behaviors of a target population or group rather than judging a specific individual about whom some descriptive information has been provided. Thus, these effects are more predictive than evaluative, and more general (group-related) than particular (person-related).

Operationally, the false consensus effect is said to have occurred "when a person engaging in a given behavior estimates that behavior to be shared by a larger proportion of some reference group than would be estimated by a person engaging in an alternative behavior" (Mullen, Driskell, & Smith, 1989, p. 84). In other words, false consensus is not about deviations from *actual* consensus (see Gross & Miller, 1997, for some discussion of this type of error) but, rather, about deviations between individuals who differ in their self-reported attributes in terms of their predictions and expectations regarding others' standing on those attributes. The effect has been reported in such domains as behaviors, traits, preferences, beliefs, and personal problems (see Marks & Miller, 1987; see Wallen, 1943, for an earlier attitudinal example). For example, people who indicate that they would wear a sandwich board sign saying "Eat at Joe's" assume that 65% of others would do the same (compared to a 31% estimate among those who would not wear the sign themselves; Ross et al., 1977); similarly, people who take showers during a shower ban assume that 63% of others do as well (compared to a 39% estimate by nonbathers; Monin & Norton, 2003). The false consensus effect—again, an assimilative tendency—is conceptually linked to the phenomena of social projection (Allport, 1924), attributive or assimilative projection (Berkowitz, 1960; Holmes, 1968, 1978; Murstein & Pryer, 1959), and assumed similarity (Cronbach, 1955). One meta-analysis based on 23 studies indicated a reliable, moderate false consensus effect size ($r = .31$; Mullen et al., 1985; see also Krueger, 2000).

As is true of other self-as-standard phenomena, explanations for the false consensus effect have focused on both cognitive and motivational factors. The most common cognitive accounts emphasize selective exposure and/or cognitive availability and accessibility—people normally associate with (and therefore have available in memory) others who are similar to themselves (Bosveld, Koomen, & van der Pligt, 1994; Goethals, Allison, & Frost, 1979; Manstead, 1982; Mullen et al., 1985; Ross et al., 1977; Sherman, Presson, Chassin, Corty, & Olshavsky, 1983). The fact that false consensus effects are stronger when projecting to ingroups than outgroups also supports the "assumption of similarity" account (Mullen, Dovidio, Johnson, & Copper, 1992; Spears & Manstead, 1990). Dawes's (1989) cognitive/statistical induction account further suggests that one's own behavior may be the only available information regarding prevalence; if the self is akin to a randomly drawn individual from the population of interest, it may be reasonable to assume that own behavior is the majority position.[3]

In addition to these cognitive factors, false consensus may also be motivated by the need to validate the correctness of a position, maintain or protect self-esteem, promote smooth interaction, and so on (Agostinelli, Sherman, Presson, & Chassin, 1992; Sherman et al., 1984; for reviews, see also Dunning, 2003; Marks & Miller, 1987). In one set of studies, participants given failure feedback (about their ability to discriminate among fake and real suicide notes) showed a strong false consensus effect in their estimates of the performance of other undergraduates on the same task (Sherman et al., 1984). Importantly, in the success conditions, participants were just as likely to "project" from their own behavior as from an available other individual's behavior in their estimates of group performance, but in the failure condition, projection was greatest when the self rather than another available individual was involved. The authors write, "when self is threatened, perceptions of consensus may be increased by a motivation to seek normalization and support for one's own behavior" (Sherman et al., 1984, p. 127).

The false uniqueness effect is a complementary pattern indicating contrast of others from the self. With false uniqueness, the prevalence of one's own attributes in a given population is underestimated. This result tends to appear in judgments of talent, ability, or other desirable characteristics (or internal traits; see McFarland & Miller, 1990), such that one's virtues are perceived as rare (Suls, Wan, Barlow, & Heimberg, 1990). In general, false uniqueness effects are less commonly reported than false consensus effects, as a count of relevant published articles attests (at this writing, PsycInfo retrieves 154 "false consensus" hits and only 17 on "false uniqueness"). Researchers creating situations favorable to both phenomena seem more likely to find evidence for consensus (e.g., Agostinelli et al., 1992; Campbell, 1986; Sherman et al., 1984). Perhaps this indicates the cumulative effect of cognitive and motivational forces in the false consensus effect.[4] That is, the dominant cognitive account of false consensus effects—self-standing as highly accessible—cannot explain false uniqueness effects, leaving self-enhancement as the primary contributor to false uniqueness.

Still, a number of factors have been proposed as determinants of whether assimilation (false consensus) or contrast (false uniqueness) occurs. For example, studies have documented assimilative effects on some dimensions (e.g., attitudes) and contrast on others (e.g., abilities; Marks, 1984). Valence of the dimension may play a role, such that we underestimate the percentage of others who perform "moral" acts such as giving blood, but overestimate the prevalence of "selfish" acts (Allison, Messick, & Goethals, 1989; Goethals, 1986; Mullen & Goethals, 1990). Furthermore, a meta-analysis indicated that independent of this valence effect, false consensus reduces as the *actual* consensus regarding an attribute increases. That is, when perceivers are actually in the majority, the tendency to overestimate this consensus is reduced (Mullen & Goethals, 1990; see also Krueger, 1998).

Additional studies have demonstrated that which effect is observed depends on an individual's standing

on the dimension (Campbell, 1986; Goethals, 1986; Sigelman, 1991). For example, Campbell (1986) found that those low in self-rated ability showed stronger false consensus effects (especially when the abilities were self-relevant) than those *high* in ability (though the latter did not show reliable false uniqueness effects). And in a study focusing on own attitudes toward stigmatized groups and perceptions of the attitudes of others, those *low in tolerance* showed false consensus effects, whereas those *high in tolerance* showed strong false uniqueness effects (Sigelman, 1991).

Still others have found evidence for both effects in a single judgment context, but on different sets or types of questions. For example, in one study, false consensus was found on measure of choice between two unpleasant situations, whereas false uniqueness was found on a measure of emotional reactions in the least preferred choice situation (McFarland & Miller, 1990). Biernat, Manis, and Kobrynowicz (1997) further suggested that the form of measurement (objective or subjective) as well as judgment order determines whether others are contrasted from or assimilated to the self.

False consensus and false uniqueness effects may also depend on the extent to which one's construal of an attribute being judged (e.g., in positive or negative terms, or in broad or narrow terms) matches beliefs about the construal of the population whose opinions are being estimated (Bosveld, Koomen, & van der Plight, 1996; Bosveld, Koomen, van der Pligt, & Plaisier, 1995). For example, in one study, highly involved Christians (compared to less involved Christians and "nonbelievers") showed a false uniqueness effect when estimating the percentage of Christians in the Netherlands, and this effect was based in part on their construal of the category "Christian" in narrow terms (Bosveld et al., 1996). Gilovich (1990) has argued that false consensus effects are stronger on items or issues that are open to "subjective construal." For example, false consensus is greater when individuals make self–other judgments of "competitiveness" (a highly subjective attribute) compared to self–other judgments of being a "first-born child" (an objective attribute, not open to differential construal).

In general, this literature points to assimilation as the more likely outcome in projections from the self to other groups. However, motivational goals of self-enhancement, as well as narrow construals of self-relevant categories, may prompt differentiation of self from others—or contrast. Interestingly, Karniol (2003) has recently pointed out an apparent paradox in the false consensus literature: False consensus effects are *weaker* (though not indicative of false uniqueness) when self-standing is salient (e.g., when participants first indicate their choices or attitudes and then estimate those of others; see meta-analyses by Fabrigar & Krosnick, 1995; Mullen et al., 1985; Mullen & Hu, 1988). This suggests that individuals do not assume similarity between self and other; in fact, the self is typically viewed as fairly distinct from others (Holyoak & Gordon, 1983; Kihlstrom et al., 1988), a feature that typically gives rise to *contrast* (Stapel & Koomen, 1998). Karniol (2003) suggests a "protocentric" alternative to the "egocentric" view re-viewed in this section: Predictions about others (typically, assimilative predictions) are made relative to generic representations of prototypical others. The next section considers a particular type of social representation—stereotypes—as they affect other judgments.

Stereotypes as Standards in Judging Others

Most prominent models and studies of stereotyping highlight *assimilation* effects. Indeed, the very definition of a "stereotyping effect" invokes assimilation—an individual member of a group is judged in line with a group stereotype (see Brewer, 1996). From this perspective, a stereotype serves as an interpretive framework that guides encoding of an individual target's behavior (von Hippel, Sekaquaptewa, & Vargas, 1995) and judgment of that target in a manner consistent with the stereotype (Bodenhausen & Macrae, 1998; Brewer, 1988; Fiske & Neuberg, 1990). Hundreds of studies suggest that stereotypes often operate in this fashion (for a review, see Schneider, 2004).

Another view suggests that in addition to serving as interpretive frameworks, stereotypes also function as *comparison standards* for judging individual group members on stereotype-relevant dimensions (Biernat & Kobrynowicz, 1997; Biernat & Manis, 1994; Biernat, Manis, & Nelson, 1991). For example, the stereotype that women are more verbally able than men may not only lead perceivers to expect that a given woman will have higher verbal skill than a given man but may also prompt them to compare the woman's skill to the expected (high) skill level of women, and the man's skill to the expected (lower) skill level of men. Because women are compared to women and men to men, standards "shift" with category membership. The result in this example is that women are held to a higher standard than men, with the potentially paradoxical consequence that a given woman may be judged *less* verbally skilled than a comparable man—a contrast effect (see Biernat & Kobrynowicz, 1997; Biernat & Manis, 1994).

Biernat and colleagues have noted, however, that such contrastive patterns are more likely when judges use *subjective* evaluative language (labels such as "high" or "low," "good" or "bad") rather than "objective" or common-rule assessments (e.g., estimates of verbal SAT scores). This is the case because subjective language is slippery—there is no inherent, stable, agreed-on meaning of the label "good"—and therefore it can be used and interpreted in a within-category fashion. "Good verbal skill" for a woman may mean something quite different (objectively better) than "good verbal skill" for a man. By definition, common-rule judgments cannot shift in meaning in this way (e.g., an SAT score means the same thing regardless of the target's category membership). Thus, for example, objective estimates of verbal skill (SAT scores, grades) tend to be higher for female than male targets, even when subjective evaluations for the sexes are identical (Biernat & Manis, 1994). Indeed, the signature "shifting standards" finding is that stereotype-assimilative findings tend to emerge when target judgments are made in objective units, but reductions or reversals (contrast) ap-

pear when subjective evaluations are rendered (for reviews, see Biernat, 2003; Biernat, Vescio, & Manis, 1998).

Stereotype-contrastive judgment patterns are also described in the literature on expectancy violation theory (Bettencourt, Dill, Greathouse, Charlton, & Mulholland, 1997; Biernat, Vescio, & Billings, 1999; Jackson, Sullivan, & Hodge, 1993; Jussim, Coleman, & Lerch, 1987). According to this perspective, when information about a target violates stereotype-based expectations, evaluations are extremitized in the direction of that violation. Thus, Blacks with strong academic credentials may be judged more competent than Whites with comparable credentials (Jackson et al., 1993; Linville & Jones, 1980), Whites who speak nonstandard English may be viewed more negatively than Blacks who do the same (Jussim et al., 1987), and baby-faced children who commit serious misbehavior may receive harsher punishment than their mature-faced peers (Zebrowitz, Kendall-Tackett, & Fafel, 1991; see also Zebrowitz & Lee, 1999).

An important distinguishing characteristic of the shifting standards model, however, is its contention that these contrastive patterns may be more apparent than real, in that they tend to diminish when targets are judged along a common continuum (i.e., in objective units). Research bears this prediction out. In an applicant evaluation study, Biernat and Kobrynowicz (1997) found relatively high subjective evaluations of women relative to men. However, the same study revealed that when other judgments and decisions were made about the applicants (e.g., objective performance estimates and decision rules to be certain of the applicant's ability) women were judged more negatively than men (lower objective performance appraisals were given; harsher decision rules for diagnosing ability were applied). In other words, assimilation to stereotypes persisted (see also Foddy & Smithson, 1989; Foschi, 1992, 1998). A very similar pattern emerged in a study of expectations and stereotypes based on baby-facedness. Berry and Zebrowitz-McArthur (1988) found that baby-faced defendants who committed intentional offenses were sentenced more severely than analogous mature-faced defendants. However, this contrast effect appeared only on the subjective sentencing decision (rated from "minimum" to "maximum"). In forced-choice judgments of *guilt* (yes or no), baby-faced defendants received the benefits of the benign stereotype associated with their group; they were less likely than mature-faced defendants to be found guilty of intentional crimes.

More generally, research on expectancy violation highlights the dual role that stereotypes and expectations may play in social judgment: They tell us what we are likely to see (and thereby influence our perceptions in the process), and they serve as benchmarks against which deviations from expectation can be noted. This echoes themes introduced in the expectation–contrast model of Manis and Paskewitz (1984a, 1984b). The outcomes of these two processes seem to depend, in part, on whether judgments are rendered in subjective or objective units.

Abele and Petzold (1998) also considered the role of stereotypes as expectations versus reference points or "perspectives." Expectations produce assimilation effects, in that assumptions about the "typical case" are integrated into judgments of the target (Anderson, 1981), whereas reference points or "perspectives" produce contrast in that the target is judged relative to the category range or boundaries (Eiser & van der Pligt, 1982; Ostrom & Upshaw, 1968; Parducci, 1965; Upshaw, 1962). The consequence of using stereotypes as expectations is that the distinction *between* categories is enhanced, and differentiation within categories is reduced (Tajfel, 1969; Tajfel & Wilkes, 1963). When stereotypes are used as reference points, however, greater within-category distinction relative to between-category differentiation results. Thus, this perspective suggests that stereotype-assimilative effects will be observed whenever conditions are present that emphasize the need for between-category differentiation. These may include motivational factors such as social identity concerns (Hogg & Abrams, 1988; Tajfel, 1957; Turner, Oakes, Haslam, & McGarty, 1994) or position in a high-power role (Fiske, 1993), as well as explicitly stated task purposes or "metainformational cues" (Leyens, Yzerbyt, & Schadron, 1994; Yzerbyt, Schadron, Leyens, & Rocher, 1994).

In their own research, Abele and Petzold (1998) documented that when judges evaluated the helpfulness of both nurses and stockbrokers in an intermixed format, this cued them that differentiation between categories was important, and assimilation to stereotypes occurred (nurses were judged more helpful than stockbrokers; see also Manis, Paskewitz, & Cotler, 1986; Martin & Seta, 1983; Seta, Martin, & Capehart, 1979; Tajfel & Wilkes, 1963). When targets were presented in blocks (nurses separated from stockbrokers), this apparently cued perceivers that within-category differentiation was critical—in blocked conditions, stockbrokers were judged more helpful than nurses (a contrast effect).

When assimilation to and when contrast from stereotypes? Few would deny that stereotypes play an important role in the judgments we make about other people. And most of the literature on stereotyping concerns itself with *assimilation* effects—the tendency for individual members of stereotyped groups to be judged consistently with group stereotypes. However, assimilation is less likely to occur when target attributes are *in*consistent with the category label, and more generally when clearly diagnostic individuating information is available (see Kunda & Thagard, 1996, for a review). As noted previously, assimilation to stereotypes is also less likely when judgments are assessed in subjective as opposed to common rule units (Biernat et al., 1991)

A number of these features have been discussed elsewhere in this chapter as factors affecting assimilation and contrast. For example, ambiguous features of targets are likely to prompt assimilation, as the information can readily be drawn to and interpreted in light of the stereotype (an "inclusion" or "interpretation" effect; Schwarz & Bless, 1992a; Stapel & Koomen, 1998). But extreme targets will not be assimilated and may in fact be contrasted from the stereotype (an "exclusion" effect). Furthermore, "fit" between category and individuating informa-

tion is important—when fit occurs, judgments of others are likely to be category based (assimilated).

It is also worth noting that to the extent that stereotypes provide a context for interpreting or evaluating a target, they can be conceptualized as *wide* or *indistinct*, as well as *entitative*—all features that promote assimilation. And major models of stereotyping, such as Fiske and Neuberg's (1990) continuum model, Brewer's (1988) dual-process model, and Bodenhausen and Macrae's (1998) model of facilitation and inhibition, highlight the importance of *processing goals*: Motivational factors such as accuracy and interdependence, or the desire not to use stereotypes, may make assimilation to stereotypes *less* likely.

Standards Used to Judge the Self

In judging others, standards and expectancies tend to function as reference points, providing information useful for interpreting and evaluating data about a given target. In judging the *self*, this informational function is also relevant, but standards can additionally serve as regulatory criteria (see Carver & Scheier, 1981, 2002; Duval & Wicklund, 1972). That is, standards provide information about "valued end states" according to which we appraise ourselves and our progress toward goals (Higgins, 1990, 1996; Higgins, Strauman, & Klein, 1986). In this section, we focus on two general sources of standards according to which we evaluate, appraise, and regulate the self: (1) internally represented "guides" and (2) other people.

Internalized Guides as Standards

That we evaluate ourselves relative to personal standards is clear in much theoretical work on the self (e.g., Cooley, 1902/1964; Kelley, 1952; Lewin, 1935). For example, Lewin (1951) noted that the experience and judgments of success and failure are based on some frame of reference, such as internalized norms and values of the culture or one's own level of aspiration (see also Sherif, 1936). That such standards are important for self-evaluation can be seen in the case of an objectively "successful" person who, by virtue of setting high standards for performance, may have a relatively lower self-appraisal than someone with low standards or with worse objective performance (see Felson, 1993). The use of internal evaluative standards may "increase the number of failures who are happy with themselves and successful people who are unhappy" (Felson, 1993, p. 16; see Brickman, Coates, & Janoff-Bulman, 1978).

OBJECTIVE SELF-AWARENESS

One theoretical perspective that has highlighted the tendency to compare the self to these internal standards is objective self-awareness (OSA) theory (Duval & Wicklund, 1972). OSA occurs "when attention is directed inward and the individual . . . is the object of his (sic) own consciousness" (p. 2). OSA is hypothesized to initiate, automatically or spontaneously, a comparison of the self against standards, where standards are defined as

"mental representations of correct behavior, attitudes and traits" (Duval & Wicklund, 1972, p. 3). To the extent that comparison to standards reveals discrepancies, negative affect should result, along with motivation to restore consistency—either by *changing* the discrepant attribute (typically, the behavior, attitude or trait, or perhaps even the standard; Batson, Thompson, Seuferling, Whitney, & Strongman, 1999; Duval & Lalwani, 1999) or by *avoiding* self-focus—getting out of the business of comparison to standards altogether (Greenberg & Musham, 1981).

In general, OSA theory suggests that comparison to standards will produce *contrast* effects when discrepancies are noted—typically negative self-evaluation relative to a (positive) standard. But assimilation (favorable self-views, consistent with the standard) can also result when the standard is met (and relative to conditions when the comparison did not occur; e.g., when one was not self-aware). Furthermore, extensions of the theory have suggested that *behaviorally*, self-awareness may prompt movement toward the standard, another kind of assimilation effect. In their test-operate-test-exit (TOTE) feedback model of OSA, Carver, Blaney, and Scheier (1979) posited that self-standard discrepancies also produce an implicit appraisal of the likelihood of reducing the discrepancy. To the extent that expectations of reducing the discrepancy are high, individuals work to meet the standard. But if expectations of meeting the standard are low, avoidance of self-focus is the result. Others have suggested that it is not expectations per se that are key but, rather, the individual's sense that his or her rate of progress in meeting the standard is sufficient relative to the size of the discrepancy (Duval, Duval, & Mulilis, 1992; Silvia & Duval, 2001). In any case, movement toward the standard reflects an assimilative process. Interestingly, the *assimilative* behavioral tendency may be prompted by the initial *contrastive* emotional reaction. That is, the negative affect produced by becoming aware of a self-standard discrepancy has a motivational quality geared toward assimilative change—typically, moving the self toward the standard (see Silvia & Duval, 2001).

SELF-DISCREPANCY

The phenomenon of self-evaluation and -regulation relative to internalized standards can also be seen in theoretical work by Higgins (1987, 1990, 1996). In his perspective on the "self as digest," an important component of self-knowledge is the *monitored self*—a representation of "how a person is doing in relation to some desired (or undesired) end state" (Higgins, 1996, p. 1070). According to this model, the self is monitored and therefore regulated by means of comparison between various *domains* of and *standpoints* on the self. It is by virtue of comparing these various selves (e.g., comparing the actual/own self to the ought/own self) and noting discrepancies between them that we experience affect (e.g., sadness in the case of actual/ideal discrepancies or guilt in the case of actual/ought discrepancies) and, potentially, engage or redirect our behavioral efforts. Higgins (1999) outlined four factors that enhance self-discrepancy effects: (1) magni-

tude/extremity of discrepancies, (2) accessibility of discrepancies, (3) applicability/relevance, and (4) importance.

Most of these moderators will sound similar to those mentioned elsewhere in this chapter. In various judgment models, extremity, applicability, and accessibility are described as factors determining whether assimilation or contrast effects emerge in social judgment. To what extent, then, can self-discrepancy effects be conceptualized in assimilation-contrast terms? As was the case with objective self-awareness theory, the emotion patterns described above reflect a type of *contrast* effect: By comparing the actual self to ideal or ought standards, we may accentuate our differences from the desired state and feel worse than we would had the comparison not occurred. Extreme, accessible, and applicable discrepancies prompt the most emotional contrast.[5]

STANDARDS OF NONPREJUDICE

Research on prejudice and prejudice expression represents an important area in which the issue of self-guides and regulation has been considered (Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, 1996; Monteith, Devine, & Zuwerink, 1993; Plant & Devine, 2001). In general, this research focuses on the likelihood that intergroup settings will reveal to individuals that they fail meet their internalized standards of nonprejudice. For example, when Whites are asked to consider the scenario "a Black person sits next to you on the bus," they may notice discrepancies between how they "would" react, and how they "should" react. As might be expected, those high in prejudice report both more prejudiced "would" reactions in these settings and more lenient standards (for how they "should" react). Furthermore, should–would discrepancies are associated with self-dissatisfaction, guilt, and discomfort among low-prejudice individuals more so than among high-prejudice individuals, though guilt can be intensified for both groups if standards are initially made salient (Monteith, 1996). As noted earlier, this general pattern can be viewed as a type of contrast effect—comparison to standards typically reveals discrepancies, and these prompt more extreme emotional reactions than occur in the absence of the comparison.

And again, comparison to prejudice-relevant internalized standards may produce a kind of *behavioral* assimilation effect—moving toward the standard—similar to that posited in the theory of objective self-awareness. For example, Monteith (1993) found that among low-prejudiced individuals, should–would discrepancies activated self-regulatory processes (such as slowing down and self-focusing) that could lead to increased control of prejudicial responses in the future (Monteith, Ashburn-Nardo, Voils, & Czopp, 2002). These responses are ultimately geared toward establishing "cues for control" that can aid the individual in inhibiting prejudiced responses. Again, the consequence of noticing a discrepancy (and experiencing negative affect as a result) prompts behavior designed to move the individual toward the (nonprejudiced) standard.

POSSIBLE SELVES

The notion of "possible selves" as goal states further suggests that by comparing the current self to a desired future self, activity directed toward that state becomes more energized and organized, with resulting positive effects on performance (Markus & Nurius, 1986; Markus & Wurf, 1987; Ruvolo & Markus, 1992; see also James's 1890/1948, notion of "potential" selves). Possible selves are "the cognitive manifestation of enduring goals, aspirations, motives, fears, and threats" and they "provide the essential link between the self-concept and motivation" (Markus & Nurius, 1986, p. 954).

Affectively, positive possible selves may create optimism and belief in the mutability of the current self (Markus & Nurius, 1986), thus fulfilling basic needs for self-enhancement. On the behavioral side, possible selves have been predicted to improve self-regulation and enhance performance precisely because they focus attention on goal-related activity and strategies necessary to achieve goals (Gollwitzer, 1996; Oyserman, 2001). In the absence of these strategies, possible selves may not produce positive outcomes. For example, in a study of low-income eighth graders, academic possible selves predicted improved grades, more time doing homework, more class participation, and less referral to remedial summer school, but only when they were accompanied by "roadmaps" or strategies for achieving positive outcomes (Oyserman, Bybee, Terry, & Hart-Johnson, 2004). Thus, a moderator of the behavioral assimilative tendencies produced by possible selves is the extent to which concrete strategies are made accessible by those future selves.

SELF-REGULATION

All the concepts described previously are ultimately about self-regulation. As the foregoing descriptions attest, standards (such as goal states) prompt positive self-evaluation when met and negative self-evaluation when not met (e.g., Bandura, 1997). Effortful action to meet standards or goals is prompted when one notices a discrepancy between reality and the standard. Again, this suggests that affective *contrast* is geared toward behavioral *assimilation*. For Bandura (1997), the behavioral movement toward the goal will only occur, however, if the individual feels self-efficacious with respect to the action required for goal attainment.

In their model of cybernetic feedback control, Carver and Scheier (2002) also posit that a goal is a standard toward which the individual is oriented. Moving toward the goal requires "identifying one's location on the relevant variable . . . , determining where that location is with respect to the goal (by means of a comparison between input and goal), and making changes to diminish the gap between present location and goal" (p. 305). This "discrepancy-reducing feedback loop" operates in an assimilative fashion. That is, discovering the discrepancy ultimately moves the individual toward the goal.

Interestingly, these researchers have also discussed *antigoals* such as feared possible selves that garner self-

regulation geared toward *enhancing* discrepancies or creating behavioral contrast (Carver, 2001; Carver, Lawrence, & Scheier, 1999; Carver & Scheier, 1998). For example, a teenager may want to avoid being like his parents (see Carver, 2001). But, such "discrepancy-enlarging" loops often get taken up in the "orbit" of a discrepancy-reducing loop. Carver (2001) writes, "the rebellious adolescent, trying to be different from his parents, soon finds a group of other adolescents to conform to, all of whom are remaining different from their parents" (p. 310). Thus, the contrastive inclination may be subsumed by an assimilative or discrepancy-reducing mechanism.

A second *affective* loop also operates and "checks on how well the behavior loop is reducing its discrepancies" (Carver, 2001, p. 311). If a discrepancy is noted (such that progress is too slow), negative affect results (e.g., depression) *and* the behavioral response is to change something—go faster, work harder, choose a different course of action. If no discrepancy is noted (progress is as expected or better than expected), affect is positive (e.g., elation or joy is experienced) and resources may be pulled back. Again, this affective system is ultimately in service of *behavioral* assimilation toward goals. However, as was briefly discussed in the context of self-awareness theory, one can also decide to change one's standards, either downward, such that progress toward goals seems better and positive affect results, or upward, such that a greater challenge is experienced (though negative affect may also result). That is, the standard may assimilate toward the behavior (Carver, 2001).

In a related vein, and building on his theory of self-discrepancy, Higgins (1997, 1998) posited two broad types of self-regulatory goals designed to meet needs for nurturance and accomplishment on the one hand and for security and safety on the other: *promotion* and *prevention*. A promotion orientation is focused on advancement or accomplishment—meeting one's goals, becoming one's "ideal self." This is an "eager *approach* strategy," in which one aims for desirable outcomes. Prevention orientation, on the other hand, is strategically failure avoidant. The individual is focused on "ought selves" and pursues strategies designed to avoid undesirable outcomes. How is this relevant to assimilation and contrast? At a very general level, promotion goals can be conceptualized in terms of assimilation: One wants to become the hoped-for self, and this organizes action in pursuit of that goal. When the goal is reached, happiness is the result, a kind of affective assimilation (Higgins, Shah, & Friedman, 1997; Idson, Liberman, & Higgins, 2000). Prevention focus, on the other hand, can be conceptualized in terms of contrast: One hopes to remain distant from the undesired self, and thus behavior is geared toward contrasting the self from that undesired outcome. When the outcome is successfully avoided, quiescent emotions result, perhaps also a contrast from the to-be-avoided state.

SUMMARY

A desired self (whether in "ought," "ideal," or "possible" form) "functions as both a goal to be attained and a standard for self-evaluation" (Higgins, 1996, p. 1071). Thus, the more important a desired self, the greater its goal strength, and the greater motivation, effort, and energy the individual is likely to direct toward the goal. The result is an increased likelihood that the goal will actually be achieved (assimilation to the standard). At the same time, the more important a desired self, the more significant it is as a standard for self-evaluation, and the more likely that standing relative to the standard will be closely monitored. The result is enhanced attention to discrepancies and intensified distress reactions when discrepancies are noted (a type of contrast; see Newman, Higgins, & Vookles, 1992). In this sense, internalized standards are like many other constructs discussed in this chapter, in that they both draw in and push away the target of evaluation under different circumstances and on different outcomes.

Others as Standards for Judging the Self: Social Comparison

One major source of standards according to which we evaluate, appraise, and regulate the self is other people. Early theorists in psychology described the importance of *reference groups* for self-definition (Hyman, 1942; Kelley, 1952; Merton & Kitt, 1950; Newcomb, 1942; Sherif, 1948). For example, Hyman (1942) noted that one's status on relevant dimensions may be determined by an understanding of oneself within the context of one's reference group. In his well-known Bennington study, Newcomb (1943) described how conservative first-year college students adopted the liberal senior class as their frame of reference and attitudinally moved toward (assimilated to) this reference group.

And, of course, social comparison theory (Festinger, 1954a, 1954b) suggests that an important source of self-assessment or -evaluation is comparisons with other people, particularly when objective information is lacking (cf. Klein, 1997). That is, others may serve as standards for evaluating the self. As discussed earlier, motives besides accuracy may drive social comparison (e.g., self-enhancement and self-improvement; Wood, 1989), leading to different *choices* of comparison targets and different *outcomes* of the comparison process.

Perhaps the most common outcomes examined in social comparison research are affective (as in measures of mood or self-esteem) or cognitive/evaluative (as in evaluative trait ratings, particularly on the dimension of the comparison) in nature. Earlier we mentioned Morse and Gergen's (1970) "Mr. Clean/Mr. Dirty" study. In this study, prospective job candidates' self-esteem was harmed by upward comparison to the well-dressed competitor, Mr. Clean, but improved by downward comparison to the sloppy competitor, Mr. Dirty. This basic pattern—downward comparisons feel good, upward comparisons feel bad—has been noted in many studies (for reviews, see Buunk, Collins, Taylor, VanYperen, & Dakof, 1990; Lockwood, 2002; Major, Testa, & Bylsma, 1991; Smith, 2000). These are *contrast* effects, and they reflect the basic notion of social comparison standards as points of reference against which the self is evaluated.

However, contrast is not an inevitable outcome of social comparison. Assimilation to an *upward* target was demonstrated very early in the history of social comparison research (Wheeler, 1966). Noting that people are often motivated to compare upward even though such comparisons may lead to negative affect, Wheeler (1966) suggested that the comparer, in such situations, assumes *similarity* with the superior other and will experience negative comparison consequences only to the extent that he or she is *not* as similar to the target as expected. Wheeler's results were consistent with this prediction: In his research, a majority (75%) of the participants who compared upward perceived themselves as similar to the comparison target, but only a minority (36%) of those who compared downward assumed they were similar to the upward targets.

More recently, Collins (1996, 2000) directly invoked the processes of assimilation and contrast to explain the consequences of comparing upward. She suggests that if perceivers expect to be different from an upward target, they will pay particular attention to these differences and contrast the self from the target. This could produce a perception of the self as even more different from (worse than) the target than is actually the case. On the other hand, an expectation that one is *similar* to an upward target results in the perceptual "blurring" of actual differences between self and target (inclusion), leading the comparer to assimilate the self to the target. It is through this assimilative process that an upward comparison target may actually be used to self-enhance.

These ideas are similar to those articulated in Mussweiler's (2003a, 2003b) SAM and Markman and McMullen's (2003) REM. Both suggest that the outcome of a comparison process depends on the information that is accessible when the comparison is made. Assimilation will result if similarity testing has made information about target–self similarity accessible (Mussweiler, 2003a), or if a reflection process invokes mental simulation of the self as like the target (Markman & McMullen, 2003a). Contrast results from dissimilarity testing for Mussweiler (2003), and from an evaluation process—in which self is distinguished from the target—for Markman and McMullen (2003).

Evidence does suggest that upward comparisons can be self-enhancing and inspiring, and that similarity between self and the upward target is important to this process (Wood, 1989; Wood & van der Zee, 1997; see also Lockwood & Kunda, 1999). For example, Lockwood and Kunda's (1997) research on comparison with "superstars" found that role models whose success was perceived as attainable produced assimilation effects, whereas dissimilar upward comparison targets—those whose accomplishments could not possibly be attained by the self—produced a contrast effect (negative self-evaluations on traits related to career success). Wood (1989) argued that upward comparisons can have *negative* consequences if the comparison target is perceived as a competitor—again, someone *dissimilar* from the self. In another clever study, women's self-assessments of attractiveness were generally *contrasted* from highly attractive standards (a photograph of a very attractive woman).

However, when the attractive other was similar to the self (because she shared one's attitudes or one's birthday), *assimilation* occurred (Brown, Novick, Lord, & Richards, 1992). Mussweiler and Bodenhausen (2002) focused on ingroup–outgroup status as an important similarity factor. They found that comparison to ingroup members prompted similarity testing and judgmental assimilation. For example, when men compared themselves with a "highly caring" male standard, self-assessments of caringness were high—an assimilation effect. But a comparison with a highly caring *female* standard—an outgroup target—prompted dissimilarity testing and contrast (see also Blanton, Crocker, & Miller, 2000; Brewer & Weber, 1994; Buunk, Ybema, van der Zee, Schaufeli, & Gibbons, 2001; McFarland, Buehler, & MacKay, 2001).

Also relevant to this discussion of the consequences of upward comparison is Tesser's self-evaluation maintenance model (SEM; Tesser, 1988; Tesser & Campbell, 1980, 1983). The SEM proposes that two processes—reflection and comparison—affect self-evaluation with close others who perform well. If a close other performs well on a dimension that is *not* central to one's self-definition, one benefits by being able to reflect in the success of the other (affective assimilation). But if the close other performs well on a dimension important to the self, negative affect and self-esteem result (contrast). Interestingly, closeness as conceptualized in the SEM should enhance *similarity* testing—the factor in Mussweiler's (2003a) SAM model that moves one toward *assimilation*. But in the case of clear outperformance by the close other, the *difference* in self- and other attainment may be highlighted and made highly accessible. Specifically, such a comparison may activate the *distinctive* aspects of the self (i.e., as a worse performer) producing contrast (see Broemer & Diehl, 2004). However, closeness may also incline one toward assimilation to the extent that one construes the upward target as part of the self. Specifically, the general pattern of contrast predicted by SEM when one is outperformed by a close other can be reversed following manipulations designed to "expand" the notion of self (Gardner, Gabriel, & Hochschild, 2002).

Similarity has also been implicated in understanding the outcomes of *downward* comparisons. To the extent that one views a downward outcome as possible for the self—particularly if one actively simulates this outcome—*assimilation* results (Lockwood, 2002). In three studies, Lockwood (2002) found that if participants felt as though they could be like the downward target (as when college undergraduates compared to a recent graduate who was having trouble on the job market), negative self-evaluation resulted (assimilation to the downward target). But when this vulnerability was reduced—when the comparison was to a fellow undergraduate doing much worse than the self in the immediate college context—self-evaluations were favorable (contrast). More generally, comparing with a worse-off other may be threatening only if the target's outcome is perceived to be uncontrollable, in that uncontrollability implies "it could happen to me" (Major et al., 1991). In Taylor and Lobel's (1989) review of their work with cancer patients, they

suggest that comparing downwardly, for instance, to a patient who is coping more poorly, may remind one of what the future may hold. Here, uncontrollability implies (potential) similarity.

Similarity ultimately boils down to the notion of *inclusion* in Schwarz and Bless's (1992a) inclusion–exclusion model. That is, if one can claim or envision oneself as *included* in the category of a comparison target (upward or downward), assimilation rather than contrast will occur.

Recently, Locke (2003) touched on this central notion of inclusion when he described "connections" or "solidarity" with the target of comparison. Features such as shared group membership (Mussweiler & Bodenhausen, 2002) or shared birthdays (Brown et al., 1992), or vulnerability to another's negative outcomes (Lockwood, 2002), or a primed sense of the self as socially integrated or "expanded" (Gardner et al., 2002; Stapel & Koomen, 2001) increase inclusion and assimilation.

BEHAVIORAL CONSEQUENCES

Relative to affect and cognition, behavioral reactions to social comparison have been studied rarely. Nevertheless, some research on upward comparisons does point to an important positive behavioral outcome: Making upward comparisons may generate negative affect but also feelings of inspiration that can sometimes lead to behavioral improvement (Blanton, Buunk, Gibbons, & Kuyper, 1999; Gibbons, Blanton, Gerrard, Buunk, & Eggleston, 2000; Major et al., 1991; Seta, 1982; Seta, Seta, & Donaldson, 1991; cf. Markman & McMullen, 2003). In one study, smokers who preferred upward comparisons—comparisons to those who were more successful at quitting—were more likely to have succeeded in their cessation attempts 6 months later (Gerrard, Gibbons, Lane, & Reis-Bergan, 2002, cited in Gibbons et al., 2002).

Markman and McMullen's (2003) REM, described earlier in this chapter, offers predictions about the motivational–behavioral outcomes of comparative thinking. In this model, all behavioral consequences are mediated by affect, but positive and negative affect are posited to produce different levels of behavioral *persistence*, depending on the type of task. Specifically, through "mood as information" processes, positive affect should produce less persistence on *achievement tasks* but more persistence on *enjoyment tasks* (see Markman & McMullen, 2003; Martin, Ward, Achee, & Wyer, 1993; Schwarz, 1990).

This leads to the interesting prediction that in social comparison situations that prompt positive affect—say, in upward comparison to a target whose accomplishments seem attainable—behavioral perseverance may *not* result. Instead, those inspired by the other may find themselves lacking the motivation to persist on tasks relevant to achieving the high standard. Although we are not aware of research directly testing this prediction, Markman and McMullen (2003) cite recent work by Oettingen and colleagues on the role of fantasy and expectations in motivation (Oettingen, 1996; Oettingen & Mayer, 2002). In this work, positive fantasy—"mentally enjoying future outcomes" (Oettingen & Mayer, 2002, p. 1199)—has been shown to undermine motivation and success because it takes energy away from activity that would otherwise be directed toward achieving the goal. Thus, upward comparison targets may create high expectations for own future success, which are inspiring and motivating. But at the same time, those who compare upward may find themselves fantasizing rather than exerting effort toward meeting the high standard. Apparently, this can be overcome, however, if perceivers are led to explicitly contrast their positive fantasies with *reality*. This makes both present and future accessible and seems to produce enhanced motivation and success (Oettingen, Pak, & Schnetter, 2001).

Researchers have also linked comparison outcomes to basic self-regulatory strategies (Pennington & Roese, 2003). Specifically, Markman and McMullen (2003) have relied on Higgins's (1998) distinction between *promotion* and *prevention* strategies in predicting likely outcomes of social and other comparisons. "An upward comparison represents a desirable outcome and thus may activate promotion goals to obtain that outcome, whereas a downward comparison represents an undesirable outcome and thus may activate prevention goals so that the outcome does not occur" (Markman & McMullen, 2003, p. 252). Lockwood (2002) found that simulation of the self as like an attainable downward comparison target indeed enhanced *prevention* goals (see also Lockwood, Sadler, Fyman, & Tuck, 2004).

THE IMPORTANCE OF CONSTRUAL

All the findings we have reviewed on the consequences of social comparison highlight the role that construals play in this process. Downward comparison, traditionally thought to be self-enhancing, can have adverse consequences just as upward comparison, traditionally considered detrimental, can have positive consequences. And much of this depends on what one does with the comparison information—one can perceive similarity or not with the comparison other; one can reconstrue information such that the self is protected even in light of negative information (DeVellis et al., 1991). The genius effect—the tendency to perceive those who outperform us as particularly brilliant—provides a very telling example of the importance of construal (Alicke, LoSchiavo, Zerbst, & Zhang, 1997). By "aggrandizing the outperfomer" we can stave off the self-esteem threat posed by underperforming: The "genius" attribution may make the other so discrepant from the self that comparison is not even viable.

Some studies have also documented *differing* effects of the *same* social comparison. For example, married women experienced more *positive* mood (assimilation) but more *negative* evaluations of their own marriage (contrast) when they compared with an upward versus downward comparison standard (another married woman; Buunk & Ybema, 2003). These authors suggest that "the mood evoked by social comparison is primarily a primitive, direct, and automatic response," whereas evaluation of one's marriage is "a separate, cognitive, process"

(p. 614). Thus, one can feel inspired but concerned about one's own marriage when comparison to a better-off other occurs.

Taken as a whole, the literature relevant to comparison consequences leads to the conclusion, in the words of Buunk and colleagues (1990), that either direction of comparison—upward or downward—"has its ups and downs." Furthermore, although we have focused on affective responses in simplistic positive–negative terms, it is also the case that a diverse set of emotions may be experienced in response to social comparison (Smith, 2000). The impact of a comparison will ultimately be determined not by its direction but by the way in which it is perceived and interpreted by the comparer.

SUMMARY AND CONCLUSIONS

This chapter has considered the role of standards in social judgment, focusing on how the use of standards may produce either assimilation or contrast effects. We have conceptualized standards broadly, arguing that they can take the form of contextual cues (as in priming research, or simply in the accumulation of exposure to stimuli), the self, other people, social groups, or internalized guides.

We began with a consideration of how standards are selected, and we emphasized that while standards may often be directly supplied or thrust on the individual, there is also room for the perceiver to actively choose or avoid comparative standards. In this way, particular needs can be met and/or certain outcomes of comparison avoided. Thus, for example, the individual can actively choose to avoid self-awareness (as a means of escaping the negative comparison of the self to standards) or can choose an upward rather than downward comparison target (depending on the desire to self-improve). We next considered a number of models of social judgment that have emerged in recent years. Each focuses largely on *outcomes* of the comparison process and whether and when assimilation versus contrast effects emerge. Finally, we considered specific social judgment phenomena—cases in which social standards guide judgment of self and others—and attempted to relate them to the broader models.

In considering each domain of self- and other judgment, we focused on how assimilation and contrast effects are *both* evident under different conditions, and how these conditions include both cognitive and motivational factors:

1. In the case of self as a standard to judge others, assimilation effects are reflected in the positive relationships between own standing on a trait dimension and judgments of another person's standing (Lambert & Wedell, 1991), and in false consensus effects (Ross et al., 1977; Sherman et al., 1984). But judgments of others can also be contrasted from the self, as in research on egocentric definitions of trait concepts—whereby self and other judgments are negatively related (Dunning & Cohen, 1992), and in false uniqueness effects (Marks, 1984; Suls & Wan, 1987). Contrastive patterns seem to be driven by self-esteem concerns, as they are particularly likely when self-esteem is threatened (e.g., Beauregard & Dunning, 1998), but cognitive factors may be responsible as well. Judgments of unambiguous or *extreme* stimuli, and judgments along dimensions where the self's standing is extreme, are also likely to prompt contrast, as are judgments relative to *distinct* standards—such as one's specific score on the SAT (rather than one's standing on an abstract trait). These features—ambiguity of stimuli, extremity of standard, and distinctness of standard—are highlighted in a number of the judgment models considered earlier in the chapter (perhaps most notably in Stapel and colleagues' ICM).

2. In the case of stereotypes as standards to judge others, the dominant theoretical models and data suggest *assimilation*. Using language from a variety of judgment models we considered, the individual group member is "included" (Schwarz & Bless, 1992a) or "interpreted" (Stapel & Koomen, 1998) in the broad frame of the stereotype, "set" in the context of the stereotype (Martin, 1986), hypothesized to be similar to the broad category (Mussweiler, 2003a, 2003b), "simulated" as a member of the category (Markman & McMullen, 2003), and judged consistently with the group-based "expectation" (Manis & Paskewitz, 1984a, 1984b). At the same time, however, other research suggests that stereotypes can produce contrast effects, as when an African American law school applicant is judged more favorably than a comparable White applicant (Linville & Jones, 1980). These contrast effects are particularly likely when judgments are made on subjective response scales (Biernat et al., 1991), and when targets violate expectancies (Jussim et al., 1987). Again, extremity versus ambiguity of targets is implicated here—ambiguous targets, or those who can be construed in ways that "fit" the category, are assimilated; extreme, nonfitting targets are not. Motivational factors have also been identified in this literature as *diminishing* the assimilative impact of stereotypes. For example, interdependence goals and motivations to avoid stereotyping tend to reduce stereotype-consistent judgments, though they typically do not invoke contrast.

3. Internalized guides operate as standards to judge the self. In research on objective self-awareness theory (Duval & Wicklund, 1972) and self-discrepancy theory (Higgins, 1987), the focus is primarily on *contrast* effects: Comparison of the self to internalized guides or standards typically prompts affective contrast. This is because our actual behavior typically does not meet our high standards or ideals, and the salience of discrepancy prompts negative affect. At the same time, the negative affect produced by becoming aware of the self's discrepancy with a standard has a motivational quality geared toward *assimilative* change—typically, moving the self toward the standard (see Monteith et al., 2002; Silvia & Duval, 2001). Similarly, the construct of possible selves implicates behavioral assimilation (Markus & Nurius, 1986, p. 954), as do models of self-regulation (Carver & Scheier, 1998; Higgins, 1998). Thus, this body of work highlights that an initial *contrastive* outcome (negative affect based on failure to meet standards) can ultimately trigger *assimilation*—the meeting of standards or goals.

4. Finally, we considered the large literature on social comparison theory, in which other individuals are used as standards to judge the self. Historically, this perspective was dominated by a *contrast* perspective: Upward comparisons were generally presumed to prompt negative self-evaluation; downward comparisons to prompt positive self-evaluations (e.g., Morse & Gergen, 1970). And indeed, there is considerable evidence supporting such patterns of contrast. However, evidence of *assimilation* has not been rare, as when cancer patients' downward comparisons prompt negative affect (Taylor & Lobel, 1989), and when upward comparisons are inspiring (Lockwood & Kunda, 1997). Key to social comparison outcomes is the *similarity* principle, highlighted in Mussweiler's (2003a) SAM: To the extent that there is similarity between self and the comparison other—in terms of shared group membership, feature overlap, attainability, or perceived uncontrollability of outcome (such that self could be like the other)—assimilation seems to be the more likely outcome (see also Collins, 1996, 2000). The social comparison literature has also highlighted that motives may affect which comparison other is chosen or even whether or not similarity is perceived. Motives to self-enhance, to improve the self, or even to self-verify may guide all aspects of the comparison process.

In thinking about the "big picture" of standards and judgment, we find ourselves disinclined to conclude this chapter by offering a new or comprehensive model of social judgment. Instead, we suggest a broad framework and a set of factors or considerations that should be taken into account when predicting the outcome of comparative judgment.

First, we suggest that the process of standard use in judgment can be described in terms of three general stages: standard selection/activation, target-standard comparison, and judgment outcome (assimilation or contrast). These stages are similar to those outlined by Mussweiler (2003a, 2003b), and we also agree with Mussweiler that information made accessible at each of the first two stages—information that becomes the focus of attention—is the most proximal contributor to the judgment outcome. We suggest, too, that at each of the first two stages, both "hot" and "cold" factors are operative.

The first stage, standard selection/activation, can involve (1) the explicit choice of a standard—as when participants in a social comparison study select the reference group with whom they would like to compare, (2) a standard being directly thrust upon the perceiver—as when one is told to compare the target X to standard Y, or (3) a standard being subtly activated in the environment—as when a prior context makes a particular category or exemplar salient. Motivations of the perceiver matter most in the first case, when choice of standard takes us toward goals of self-enhancement, accuracy (self-evaluation), self-improvement, self-verification, or a host of other motives (including epistemic goals or curiosity; see Dunning, 2001). But even when a standard is explicitly or subtly activated in a given context, the perceivers' motives may affect their construal of the standard, as when

an upward social comparison standard is viewed as highly similar to the self (presumably reflecting some self-enhancement goals; Wheeler, 1966). In addition, of course, standard activation or selection implicates an array of cognitive features that ultimately contribute to judgment outcomes. Relevant but *extreme* or *distinct* standards set the stage for contrast; more *moderate*, *broad*, or *abstract* standards set the stage for assimilation (e.g., see Herr, Sherman, & Fazio, 1983).

At the second stage, target–standard comparison, motives of the judge may again play a role. When, for example, the self is a relevant standard for judging others, one may project the self on to others to gain validation for one's views, or may distinguish the self from others to establish uniqueness or rarity of favorable qualities (McFarland & Miller, 1990). And, of course, when an internal standard becomes salient, pointing to a self-discrepancy, the individual can choose to escape self-awareness to stop the comparison altogether (Greenberg & Musham, 1981). Again, however, this stage is affected by colder cognitive processes as well. To the extent that the standard and target share features, or can be included in the same category, assimilation is likely; to the extent that the standard is out of awareness but active—as in many priming studies—assimilation is also likely. Ambiguous targets are also more likely to prompt assimilation than extreme or unambiguous targets.

It is through these two paths then—the motives of the perceiver and the cognitive features of the standard, the target, and the standard–target overlap/distinctness—that judgment in the form of assimilation or contrast results. This judgment stage can also be affected by additional factors, such as the language in which judgment is rendered (with *objective* or *common-rule* judgments more likely to produce assimilation; Biernat et al., 1991; Krantz & Campbell, 1961), and motives to *correct* for the (unwanted) influence of a standard on judgment (which typically produces contrast [Martin, 1986], but can produce assimilation [Wegener & Petty, 1997]).

The goal of this chapter has been to make the process and outcomes of standard use a bit more transparent—to highlight basic principles by which standards influence social judgment. We have not considered *non*social judgment here, and thus we have said nothing about how federal agencies decide that drugs are or are not safe, or that cars and toys meet or do not meet safety standards. We have also avoided discussion of legal decision making or admissions policies or ethical standards, where social judgments *are* ultimately involved. In these cases, a variety of factors outside our purview surely contribute to the *development* of standards (what constitutes a safe drug? what is "reasonable doubt" in a criminal case? what are ethical standards of conduct?). And in a number of these cases, judging a stimulus relative to a standard may be a fairly *objective* process (as when initial screening for graduate school applicants requires the applicant achieve some minimum standard on the GRE). At the same time, however, we suspect that motives of the perceiver or judge can influence which standards are activated in a given context, or how information about the standard and target are processed. Features of the standard and

target—including extremity, distinctness, standard–target overlap—also determine the outcome of the comparative process. In short, we suspect that the models and phenomena reviewed in this chapter are broadly applicable to the use of standards in everyday life.

NOTES

1. From a "related attributes" perspective, it has been hypothesized that people should prefer dissimilar others when it comes to evaluating opinions, a prediction that has also received support (e.g., Goethals & Zanna, 1979; Gorenflo & Crano, 1989)

2. This derives from the "maximum–minimum" principle of Gestalt theory: a maximum simplicity of perfect articulation that produces contrast and a minimum simplicity of uniformity that produces assimilation. Stimulus differences are minimized to create homogeneity within a category and maximized to accentuate category differences (see Tajfel & Wilkes, 1963).

3. However, Krueger and Clement (1994) have demonstrated that the false consensus bias remains even after providing participants with other statistical information and educating them about the bias (see also Alicke & Largo, 1995).

4. Krueger (1998) has suggested that an egocentric *perceptual* process underlies the false consensus effect. That is, "the perception of consensus is assumed to be part of the initial encoding of the stimulus rather than the outcomes of subsequent higher level processes" (Krueger, 1998, p. 202). Such an account need not imply motivation or cognition directly.

5. Though most research in the self-discrepancy tradition has focused on failures to meet standards, recent work from a regulatory focus perspective (Higgins, 1997, 1998) discusses the assimilative consequences of meeting ideal standards. Specifically, when promotion-focused, success (the "gain" of meeting an ideal standard) is thought to evoke cheerfulness-related emotions. But when prevention-focused, success (the "nonloss" of meeting an ought standard) is thought to evoke quiescence-related emotions (Higgins, Shah, & Friedman, 1997; Idson, Liberman, & Higgins, 2000). This latter effect is a kind of affective contrast—emotion is more favorable than the undesired (and successfully avoided) outcome.

REFERENCES

Abele, A. E., & Petzold, P. (1998). Pragmatic use of categorical information in impression formation. *Journal of Personality and Social Psychology*, *75*, 347–358.

Agostinelli, G., Sherman, S. J., Presson, C. C., & Chassin, L. (1992). Self-protection and self-enhancement biases in estimates of population prevalence. *Personality and Social Psychology Bulletin*, *18*, 631–642.

Alicke, M. D., & Largo, E. (1995). The role of the self in the false consensus effect. *Journal of Experimental Social Psychology*, *31*, 28–47.

Alicke, M. D., LoSchiavo, F. M., Zerbst, J., & Zhang, S. (1997). The person who outperforms me is a genius: Maintaining perceived competence in upward social comparison. *Journal of Personality and Social Psychology*, *73*, 781–789.

Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, *7*, 275–295.

Allport, F. W. (1924). The group fallacy in relation to social science. *Journal of Abnormal and Social Psychology*, *19*, 60–73.

Anderson, N. H. (1974). Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), *Advances in ex-*
perimental social psychology (Vol. 7, pp. 1–101). New York: Academic Press.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Avant, L. L., & Helson, H. (1973). Theories of perception. In B. B. Wolman (Ed.), *Handbook of general psychology* (pp. 419–448). Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Handbook of social cognition: Vol. 1. Basic processes* (pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, *77*, 525–537.

Beauregard, K. S., & Dunning, D. (1998). Turning up the contrast: Self-enhancement motives prompt egocentric contrast effects in social judgment. *Journal of Personality and Social Psychology*, *74*, 606–621.

Beauregard, K. S., & Dunning, D. (2001). Defining self-worth: Trait self-esteem moderates the use of self-serving trait definitions in social judgment. *Motivation and Emotion*, *25*, 135–161.

Berkowitz, L. (1960). The judgmental process in personality functioning. *Psychological Review*, *67*, 130–142.

Berry, D. S., & Zebrowitz-McArthur, L. (1988). What's in a face?: Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin*, *14*, 23–33.

Bettencourt, B. A., Dill, K. E., Greathouse, S. A., Charlton, K., & Mulholland, A. (1997). Evaluations of ingroup and outgroup members: The role of category-based expectancy violation. *Journal of Experimental Social Psychology*, *33*, 244–275.

Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, *58*, 1019–1027.

Biernat, M., & Kobrynowicz, D. (1997). Gender- and race-based standards of competence: Lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, *72*, 544–557.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5–20.

Biernat, M., Manis, M., & Kobrynowicz, D. (1997). Simultaneous assimilation and contrast effects in judgments of self and other. *Journal of Personality and Social Psychology*, *73*, 254–269.

Biernat, M., Manis, M., & Nelson, T. F. (1991). Comparison and expectancy processes in human judgment. *Journal of Personality and Social Psychology*, *61*, 203–211.

Biernat, M., Vescio, T. K., & Billings, L. S. (1999). Black sheep and expectancy violation: Integrating two models of social judgment. *European Journal of Social Psychology*, *29*, 523–542.

Biernat, M., Vescio, T. K., & Manis, M. (1998). Judging and behaving toward members of stereotyped groups: A shifting standards perspective. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 151–175). Hillsdale, NJ: Erlbaum.

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, *7*, 420–430.

Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of ingroup versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, *36*, 519–530.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr. (Ed.), *Stereotype activation and in-*

hibition. Advances in social cognition (pp. 1–52). Mahwah, NJ: Erlbaum.

Bosveld, W., Koomen, W., & van der Pligt, J. (1994). Selective exposure and the false consensus effect: The availability of similar and dissimilar others. *British Journal of Social Psychology*, *33*, 457–466.

Bosveld, W., Koomen, W., & van der Pligt, J. (1996). Estimating group size: Effects of category membership, differential construal and selective exposure. *European Journal of Social Psychology*, *26*, 523–535.

Bosveld, W., Koomen, W., van der Pligt, J., & Plaisier, J. W. (1995). Differential construal as an explanation for false consensus and false uniqueness effects. *Journal of Experimental Social Psychology*, *31*, 518–532.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer, Jr. (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. B. (1996). When stereotypes lead to stereotyping: The use of stereotypes in person perception. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 254–275). New York: Guilford Press.

Brewer, M. B., & Weber, J. G. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, *66*, 268–275.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, *36*, 917–927.

Broemer, P., & Diehl, M. (2004). Evaluative contrast in social comparison: The role of distinct and shared features of the self and comparison others. *European Journal of Social Psychology*, *34*, 25–38.

Brown, D. R. (1953). Stimulus similarity and the anchoring of subjective scales. *American Journal of Psychology*, *66*, 199–214.

Brown, J. D., Novick, N. J., Lord, K. A., & Richards, J. M. (1992). When Gulliver travels: Social context, psychological closeness, and self-appraisals. *Journal of Personality and Social Psychology*, *62*, 717–727.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology*, *59*, 1238–1249.

Buunk, B. P., & Ybema, J. F. (2003). Feeling bad, but satisfied: The effects of upward and downward comparison upon mood and marital satisfaction. *British Journal of Social Psychology*, *42*, 613–628.

Buunk, B. P., Ybema, J. F., van der Zee, K., Schaufeli, W. B., & Gibbons, F. X. (2001). Affect generated by social comparisons among nurses high and low in burnout. *Journal of Applied Social Psychology*, *31*, 1500–1520.

Campbell, J. (1986). Similarity and uniqueness: The effects of attribute type, relevance, and individual differences in self-esteem and depression. *Journal of Personality and Social Psychology*, *50*, 281–294.

Carpenter, S. (1988). Self-relevance and goal-directed processing in the recall and weighting of information about others. *Journal of Experimental Social Psychology*, *24*, 310–332.

Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review*, *5*, 345–356.

Carver, C. S., Blaney, P. H., & Scheier, M. F. (1979). Reassertion and giving up: The interactive role of self-directed attention and outcome expectancy. *Journal of Personality and Social Psychology*, *37*, 1859–1870.

Carver, C. S., Lawrence, J. W., & Scheier, M. F. (1999). Self-discrepancies and affect: Incorporating the role of feared selves. *Personality and Social Psychology Bulletin*, *25*, 783–792.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behaviors*. New York: Springer-Verlag.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* New York: Cambridge University Press.

Carver, C. S., & Scheier, M. F. (2002). The hopeful optimist. *Psychological Inquiry*, *13*, 288–290.

Catrambone, R., & Markus, H. (1987). The role of self-schemas in going beyond the information given. *Social Cognition*, *5*, 349–368.

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluation. *Psychological Bulletin*, *119*, 51–69.

Collins, R. L. (2000). Among the better ones: Upward assimilation in social comparison. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 159–171). New York: Kluwer Academic.

Combs, A. W. (1959). *Individual behavior: A perceptual approach to behavior* (rev. ed.). New York: Harper & Row.

Cooley, C. H. (1964). *Human nature and the social order*. New York: Charles Scribner's Sons. (Original work published 1902)

Crocker, J., & Major, B. (1989). Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review*, *96*, 608–630.

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, *52*, 177–193.

Davis, J. A. (1966). The campus as a frog pond: An application of the theory of relative deprivation to career decisions of college men. *American Journal of Sociology*, *72*, 17–31.

Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology*, *25*, 1–17.

DeVellis, R. F., Blalock, S. J., Holt, K., Renner, B. R., Blanchard, L. W., & Klotz, M. L. (1991). Arthritis patients' reactions to unavoidable social comparisons. *Personality and Social Psychology Bulletin*, *17*, 392–399.

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817–830.

Dijksterhuis, A., Spears, R., & Lépinasse, V. (2001). Reflecting and deflecting stereotypes: Assimilation and contrast in impression formation and automatic behavior. *Journal of Experimental Social Psychology*, *37*, 286–299.

Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D., Koomen, W., van Knippenberg, A., et al. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology*, *75*, 862–871.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584.

Dornbush, S. M., Hastorf, A. H., Richardson, S. A., Muzzy, R. E., & Vreeland, R. S. (1965). The perceiver and the perceived: Their relative influence on the categories of interpersonal cognition. *Journal of Personality and Social Psychology*, *3*, 434–440.

Dunning, D. (1993). Words to live by: The self and definitions of social concepts and categories. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 99–126). Hillsdale, NJ: Erlbaum.

Dunning, D. (2001). On the motives underlying social cognition. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Volume 1: Intraindividual processes* (pp. 348–374). London: Blackwell.

Dunning, D. (2003). The relation of self to social perception. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 421–441). New York: Guilford Press.

Dunning, D., & Beauregard, K. S. (2001). Regulating impressions of others to affirm the self. *Social Cognition*, *18*, 198–222.

Dunning D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology*, *63*, 341–355.

Dunning, D., & Hayes, A. F. (1996). Evidence for egocentric comparison in social judgment. *Journal of Personality and Social Psychology*, *71*, 213–229.

Dunning, D., & McElwee, R. O. (1995). Idiosyncratic trait definitions: Implications for self-description and social judgment. *Journal of Personality and Social Psychology*, *68*, 936–946.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082–1090.

Dunning, D., Perie, M., & Story, A. L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology*, *61*, 957–968.

Duval, T. S., Duval, V. H., & Mulilis, J. P. (1992). Effects of self-focus, discrepancy between self and standard, and outcome expectancy favorability on the tendency to match self to standard or to withdraw. *Journal of Personality and Social Psychology*, *62*, 340–348.

Duval, T. S., & Lalwani, N. (1999). Objective self-awareness and causal attributions for self-standard discrepancies: Changing self or changing standards of correctness. *Personality and Social Psychology Bulletin*, *25*, 1220–1229.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. San Diego, CA: Academic Press.

Eiser, J. R. (1990). *Social judgment*. Pacific Grove, CA: Brooks/Cole.

Eiser, J. R., & Stroebe, W. (1972). *Categorization and social judgment*. New York: Academic Press.

Eiser, J. R., & van der Pligt, J. (1982). Accentuation and perspective in attitudinal judgment. *Journal of Personality and Social Psychology*, *42*, 224–238.

Fabrigar, L. R., & Krosnick, J. A. (1995). Attitude importance and the false consensus effect. *Personality and Social Psychology Bulletin*, *21*, 468–479.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf & Hartel.

Felson, R. B. (1990). Comparison processes in parents' and children's appraisals of academic performance. *Social Psychology Quarterly*, *53*, 264–273.

Felson, R. B. (1993). The (somewhat) social self: How others affect self-appraisals. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 1–26). Hillsdale, NJ: Erlbaum.

Festinger, L. (1954a). Motivation leading to social behavior. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 2, pp. 191–218). Lincoln: University of Nebraska Press.

Festinger, L. (1954b). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Foddy, M., & Smithson, M. (1989). Fuzzy sets and double standards: Modeling the process of ability inference. In J. Berger, M. Zelditch, & B. Anderson (Eds.), *Sociological theories in progress: New formulations* (pp. 73–99). Newbury Park, CA: Sage.

Foschi, M. (1992). Gender and double standards for competence. In C. L. Ridgeway (Ed.), *Gender, interaction, and inequality* (pp. 181–207). New York: Springer-Verlag.

Foschi, M. (1998). Double standards: Types, conditions, and consequences. *Advances in Group Processes*, *15*, 59–80.

Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status*. London: Oxford University Press.

Freud, S. (1956). Further remarks on the defense neuropsychoses. In *Collected papers of Sigmund Freud* (Vol. 1, pp. 155–182). London: Hogarth Press. (Original work published 1924)

Gardner, W. L., Gabriel, S., & Hochschild, L. (2002). When you and I are "we," you are not threatening: The role of self-expansion in social comparison. *Journal of Personality and Social Psychology*, *82*, 239–251.

Gastorf, J. W., & Suls, J. (1978). Performance evaluation via social comparison: Performance similarity versus related-attribute similarity. *Social Psychology*, *41*, 297–305.

Gibbons, F. X., Blanton, H., Gerrard, M., Buunk, B., & Eggleston, T. (2000). Does social comparison make a difference? Optimism as a moderator of the relation between comparison level and academic performance. *Personality and Social Psychology Bulletin*, *26*, 637–648.

Gibbons, F. X., Lane, D. J., Gerrard, M., Reis-Bergan, M., Lautrup, C. L., Pexa, N. A., et al. (2002). Comparison-level preferences after performance: Is downward comparison theory still useful? *Journal of Personality and Social Psychology*, *83*, 865–880.

Giladi, E. E., & Klar, Y. (2002). When standards are wide of the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. *Journal of Experimental Psychology: General*, *131*, 538–551.

Gilbert, D. T., Giesler, R., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, *69*, 227–236.

Gilbert, D. T., Pelham, B., & Krull, D. S. (1988). On cognitive busyness. When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*, 733–740.

Gilovich, T. (1990). Differential construal and the false consensus effect. *Journal of Personality and Social Psychology*, *59*, 623–634.

Goethals, G. R. (1986). Fabricating and ignoring social reality: Self-serving estimates of consensus. In J. Olson, C. P. Herman, & M. P. Zanna (Eds.), *Relative deprivation and social comparison: The Ontario symposium on social cognition* (Vol. 4, pp. 135–157). Hillsdale, NJ: Erlbaum.

Goethals, G. R., Allison, S. J., & Frost, M. (1979). Perceptions of the magnitude and diversity of social support. *Journal of Experimental Social Psychology*, *15*, 570–581.

Goethals, G. R., & Darley, J. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Socoial comparison processes: Theoretical and empirical perspectives* (pp. 259–278). Washington, DC: Hemisphere.

Goethals, G. R., Messick, D., & Allison, S. (1991). The uniqueness bias: Studies of constructive social comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 149–176). Hillsdale, NJ: Erlbaum.

Goethals, G. R., & Zanna, M. P. (1979). The role of social comparison in choice shifts. *Journal of Personality and Social Psychology*, *37*, 1469–1476.

Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 287–312). New York: Guilford Press.

Gorenflo, D. W., & Crano, W. D. (1989). Judgmental subjectivity/objectivity and locus of choice in social comparison. *Journal of Personality and Social Psychology*, *57*, 605–614.

Green, J. D., & Sedikides, C. (2001). When do self-schemas shape social perception? The role of descriptive ambiguity. *Motivation and Emotion*, *25*, 67–83.

Greenberg, J., & Musham, C. (1981). Avoiding and seeking self-focused attention. *Journal of Research in Personality*, *15*, 191–200.

Gross, S. R., & Miller, N. (1997). The "golden section" and bias in perceptions of social consensus. *Personality and Social Psychology Review*, *1*, 241–271.

Hakmiller, K. L. (1966). Threat as a determinant of downward comparison. *Journal of Experimental Social Psychology*, *2*(Suppl. 1), 32–39.

Heider, F. (1958). *The psychology of interpersonal relations*. Oxford, UK: Wiley.

Helgeson, V. S., & Mickelson, K. D. (1995). Motives for social comparison. *Personality and Social Psychology Bulletin*, *21*, 1200–1209.

Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *American Journal of Psychology*, *60*, 1–29.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology*, *51*, 1106–1115.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the conse-

quences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology*, *19*, 323–340.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1990). Personality, social psychology, and person-situated relations: Standards and knowledge activation as a common language. In L. A. Pervin (Ed.), *Handbook of personality* (pp. 301–338). New York: Guilford Press.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press.

Higgins, E. T. (1999). When do self-discrepancies have specific relations to emotions? The second generation question of Tangney, Niedenthal, Covert, & Barlow (1998). *Journal of Personality and Social Psychology*, *77*, 1313–1317.

Higgins, E. T., & Lurie, L. (1983). Context, categorization and recall: The "change-of-standard" effect. *Cognitive Psychology*, *15*, 525–547.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Higgins, E. T., Shah, J., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as a moderator. *Journal of Personality and Social Psychology*, *72*, 515–525.

Higgins, E. T., & Stangor, C. (1988). A "change-of-standard" perspective on the relations among context, judgment and memory. *Journal of Personality and Social Psychology*, *54*, 181–192.

Higgins, E. T., Strauman, T., & Klein, R. (1986). Standards and the process of self-evaluation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 23–63). New York: Guilford Press.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.

Holmes, D. S. (1968). Dimensions of projection. *Psychological Bulletin*, *69*, 248–268.

Holmes, D. S. (1978). Projection as a defense mechanism. *Psychological Bulletin*, *85*, 677–688.

Holyoak, K. J., & Gordon, P. C. (1983). Social reference points. *Journal of Personality and Social Psychology*, *44*, 881–887.

Houston, D. A., & Sherman, S. J., (1995). Cancellation and focus: The role of shared and unique features in the choice process. *Journal of Experimental Social Psychology*, *31*, 357–378.

Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *Journal of Abnormal and Social Psychology*, *55*, 244–252.

Hovland, C. I., & Sherif, M. (1952). Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology*, *47*, 822–832.

Hyman, H. (1942). The psychology of subjective status. *Psychological Bulletin*, *39*, 473–474.

Idson, L. C., Liberman, N., & Higgins, E. T. (2000). Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective on hedonic intensity. *Journal of Experimental Social Psychology*, *36*, 252–274.

Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology*, *65*, 69–84.

James, W. (1948). *Principles of psychology*. New York: Holt. (Original work published 1890)

Judd, C. M., & Harackiewicz, J. M. (1980). Contrast effects in attitude judgment: An examination of the accentuation hypothesis. *Journal of Personality and Social Psychology*, *38*, 390–398.

Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology*, *52*, 536–546.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Karniol, R. (2003). Egocentrism versus protocentrism: The status of self in social prediction. *Psychological Review*, *110*, 564–580.

Kawada, C. L. K., Oettingen, G., Gollwitzer, P. M., & Bargh, J. A. (2004). The projection of implicit and explicit goals. *Journal of Personality and Social Psychology*, *86*, 545–559.

Kelley, H. H. (1952). Two functions of reference groups. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (2nd ed., pp. 410–430). New York: Holt, Rinehart & Winston.

Kihlstrom, J. F., Cantor, N., Albright, J. S., Chew, B. R., Klein, S. B., & Niedenthal, P. M. (1988). Information processing and the study of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp.145 178). San Diego, CA: Academic Press.

Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology*, *73*, 885–901.

Klein, W. M. (1992). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, *72*, 763–774.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110*, 499–519.

Krantz, D. L., & Campbell, D. T. (1961). Separating perceptual and linguistic effects of Context shifts upon absolute judgments. *Journal of Experimental Psychology*, *62*, 35–42.

Krech, D., & Crutchfield, R. S. (1948). *Theory and problems of social psychology*. New York: McGraw-Hill.

Krueger, J. (1998). Enhancement bias in descriptions of self and others. *Personality and Social Psychology Bulletin*, *24*, 505–516.

Krueger, J. (2000). The projective perception of the social world: A building block of social comparison processes. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 323–351). New York: Plenum/Kluwer.

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*, 596–610.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint satisfaction theory. *Psychological Review*, *103*, 284–308.

Lambert, A. J., & Wedell, D. H. (1991). The self and social judgment: Effects of affective reaction and "own position" on judgments of unambiguous and ambiguous information about others. *Journal of Personality and Social Psychology*, *61*, 884–897.

Lemaine, G. (1974). Social differentiation and social originality. *European Journal of Social Psychology*, *4*, 17–52.

Lemon, N., & Warren, N. (1976). Salience, centrality, and self-relevance of traits in construing others. *British Journal of Social and Clinical Psychology*, *13*, 119–124.

Lewicki, P. (1983). Self-image bias in person perception. *Journal of Personality and Social Psychology*, *45*, 384–393.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers*. Oxford, UK: Harpers.

Leyens, J. P., Yzerbyt, V., & Schadron, G. (1994). *Stereotypes and social cognition*. London: Sage.

Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of outgroup members. *Journal of Personality and Social Psychology*, *38*, 689–703.

Locke, K. D. (2003). Status and solidarity in social comparison: Agentic and communal values and vertical and horizontal directions. *Journal of Personality and Social Psychology*, *84*, 619–631.

Lockwood, P. (2002). Could it happen to you? Predicting the impact of downward comparisons on the self. *Journal of Personality and Social Psychology*, *82*, 343–358.

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, *73*, 91–103.

Lockwood, P., & Kunda, Z. (1999). Increasing the salience of one's best selves can undermine inspiration by outstanding role models. *Journal of Personality and Social Psychology*, *76*, 214–228.

Lockwood, P., Sadler, P., Fyman, K., & Tuck, S. (2004). To do or not to do: Using positive and negative role models to harness motivation. *Social Cognition*, *22*, 422–450.

Lombardi, W. J., Higgins, E. T., & Bargh, J. A. (1987). The role of consciousness in priming effects on categorization. *Personality and Social Psychology Bulletin*, *13*, 411–429.

Major, B., Testa, M., & Bylsma, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 237–260). Hillsdale, NJ: Erlbaum.

Manis, M., Biernat, M., & Nelson, T. F. (1991). Comparison and expectancy processes in human judgment. *Journal of Personality and Social Psychology*, *61*, 203–211.

Manis, M., & Blake, J. B. (1963). Interpretation of persuasive messages as a function of prior immunization. *Journal of Abnormal and Social Psychology*, *66*, 225–230.

Manis, M., & Moore, J. C. (1978). Summarizing controversial messages: Retroactive effects due to subsequent information. *Social Psychology*, *41*, 62–68.

Manis, M., & Paskewitz, J. R. (1984a). Judging psychopathology: Expectation and contrast. *Journal of Experimental Social Psychology*, *20*, 363–381.

Manis, M., & Paskewitz, J. R. (1984b). Specificity and contrast effects: Judgments of psychopathology. *Journal of Experimental Social Psychology*, *20*, 217–230.

Manis, M., Paskewitz, J. R., & Cotler, S. (1986). Stereotypes and social judgment. *Journal of Personality and Social Psychology*, *50*, 461–473.

Manstead, A. S. R. (1982). Perceived social support for opinions: A test of the magnitude and diversity hypothesis. *British Journal of Social Psychology*, *21*, 35–41.

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality and Social Psychology Review*, *7*, 244–267.

Marks, G. (1984). Thinking one's abilities are unique and one's opinions are common. *Personality and Social Psychology Bulletin*, *10*, 203–208.

Marks, G., & Miller, N. (1987). Ten years of research on the false consensus effect: An empirical and theoretical review. *Psychological Bulletin*, *102*, 72–90.

Markus, H., Crane, M., Bernstein, S., & Siladi, M. (1982). Self-schemas and gender. *Journal of Personality and Social Psychology*, *42*, 38–50.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Markus, H., & Smith, J. (1981). The influence of self-schemata on the perception of others. In N. Cantor & J. Kihlstrom (Eds.), *Personality, cognition, and social interaction* (pp. 233–262). Hillsdale, NJ: Erlbaum.

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in perception of others. *Journal of Personality and Social Psychology*, *49*, 1494–1512.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 195–216). Hillsdale, NJ: Erlbaum.

Martin, L. L., & Seta, J. J. (1983). Perceptions of unity and distinctiveness as determinants of attraction. *Journal of Personality and Social Psychology*, *44*, 755–764.

Martin, L. L., Seta, J. J., & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. *Journal of Personality and Social Psychology*, *59*, 27–37.

Martin, L. L., Ward, D. W., Achee, J. W., & Wyer, R. S., Jr. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, *64*, 317–326.

McFarland, C., Buehler, R., & MacKay, L. (2001) Affective responses to social comparison with extremely close others. *Social Cognition*, *19*, 547–586.

McFarland, C., & Miller, D. T. (1990). Judgments of self-other similarity—Just like other people, only more so. *Personality and Social Psychology Bulletin*, *16*, 475–484.

McMullen, M. N. (1997). Affective contrast and assimilation in counterfactual thinking. *Journal of Experimental Social Psychology*, *33*, 77–100.

Merton, R. K., & Kitt, A. S. (1950). Contributions to the theory of reference group behavior. In R. Merton & P. F. Lazarsfeld (Eds.), *Continuities in social research: Studies in the scope and method of "The American Soldier"* (pp. 40–105). Glencoe, IL: Free Press.

Miller, C. T. (1982). The role of performance-related similarity in social comparison of abilities: A test of the related attributes hypothesis. *Journal of Experimental Social Psychology*, *18*, 513–523.

Miller, D. T., & Prentice, D. A. (1996). The construction of social norms and standards. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 799–829). New York: Guilford Press.

Monin, B., & Norton, M. I. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus, false polarization, and pluralistic ignorance in a water conservation crisis. *Personality and Social Psychology Bulletin*, *29*, 559–567.

Monteith, M. J. (1996). Contemporary forms of prejudice-related conflict: In search of a nutshell. *Personality and Social Psychology Bulletin*, *22*, 461–473.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology*, *65*, 469–485.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operations of cues for control. *Journal of Personality and Social Psychology*, *83*(5), 1029–1050.

Monteith, M. J., Devine, P. G., & Zuwerink, J. R. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology*, *64*, 198–210.

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, *16*, 148–156.

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., et al. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, *21*, 262–283.

Mullen, B., Dovidio, J. F., Johnson, C., & Copper, C. (1992). In-group out-group differences in social projection. *Journal of Experimental Social Psychology*, *28*, 422–440.

Mullen, B., Driskell, J. E., & Smith, C. (1989). Availability and social projection: The effects of sequence of measurement and wording of question on estimates of consensus. *Personality and Social Psychology Bulletin*, *15*, 84–90.

Mullen, B., & Goethals, G. R. (1990). Social projection, actual consensus and valence. *British Journal of Social Psychology*, *29*, 279–282.

Mullen, B., & Hu, L. (1988). Social projection as a function of cog-

nitive mechanisms: Two meta-analytic integrations. *British Journal of Social Psychology*, 27, 333–356.

Murstein, B. I., & Pryer, R. S. (1959). The concept of projection: A review. *Psychological Bulletin*, 56, 353–374.

Mussweiler, T. (2003a). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110, 472–489.

Mussweiler, T. (2003b). When egocentrism breeds distinctiveness—Comparison processes in social prediction: Comment on Karniol (2003). *Psychological Review*, 110, 581–584.

Mussweiler, T., & Bodenhausen, G. V. (2002). I know you are, but what am I?: Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology*, 82, 19–32.

Mussweiler, T., & Rüter, K. (2003). What friends are for! The use of routine standards in social comparison. *Journal of Personality and Social Psychology*, 85, 467–481.

Mussweiler, T., Rüter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology*, 40, 689–696.

Newcomb, T. M. (1942). Community roles in attitude formation. *American Sociological Review*, 7, 621–630.

Newcomb, T. M. (1943). *Personality and social change*. Fort Worth, TX: Dryden Press.

Newman, L. S., Higgins, E. T., & Vookles, J. (1992). Self-guide strength and emotional vulnerability: Birth order as a moderator of self-affect relations. *Personality and Social Psychology Bulletin*, 18, 402–411.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.

Oettingen, G. (1996). Positive fantasy and motivation. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 236–259). New York: Guilford Press.

Oettingen, G., & Mayer, D. (2002). The motivating function of thinking about the future: Expectations versus fantasies. *Journal of Personality and Social Psychology*, 83, 1198–1212.

Oettingen, G., Pak, H., & Schnetter, K. (2001). Self-regulation of goal-setting: Turning free fantasies about the future into binding goals. *Journal of Personality and Social Psychology*, 80, 736–753.

Ostrom, T. M., & Upshaw, H. S. (1968). Psychological perspective and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 217–242). San Diego, CA: Academic Press.

Oyserman, D. (2001). Self concept and identity. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology* (pp. 449–517). Malden, MA: Blackwell Press.

Oyserman, D., Bybee, D., Terry, K., & Hart-Johnson, T. (2004). Possible selves as roadmaps. *Journal of Research in Personality*, 38, 130–149.

Parducci, A. (1956). Direction of shift in the judgment of single stimuli. *Journal of Experimental Psychology*, 51, 169–178.

Parducci, A. (1963). Range-frequency compromise in judgment. *Psychological Monographs*, 77, 1–29.

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, 72, 407–418.

Parducci, A. (1983). Category ratings and the relational character of judgment. In H. G. Geissler, H. F. J. M. Buffort, E. L. J. Leeuwenberg, & V. Sarris (Eds.), *Modern issues in perception* (pp. 89–105). Berlin: VEB Deutscher Verlag der Vissenschaffen.

Parducci, A., Knobel, S., & Thomas, C. (1976). Independent contexts for category ratings: A range-frequency analysis. *Perception and Psychophysics*, 20, 360–366.

Pennington, G. L., & Roese, N. J. (2003). Regulatory focus and temporal distance. *Journal of Experimental Social Psychology*, 39, 563–576.

Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology*, 29, 137–165.

Plant, E. A., & Devine, P. G. (2001). Response to other-imposed pro-Black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology*, 37, 486–501.

Postman L., & Miller, G. A. (1945). Anchoring of temporal judgments. *American Journal of Psychology*, 58, 43–53.

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279–301.

Ross, L. D., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, 35, 817–829.

Ruvolo, A. P., & Markus, H. R. (1992). Possible selves and performance: The power of self-relevant imagery. *Social Cognition*, 10, 95–124.

Sanitioso, R., & Kunda, Z. (1991). Ducking the collection of costly evidence: Motivated use of statistical heuristics. *Journal of Behavioral Decision Making*, 4, 161–176.

Schneider, D. J. (2004). *The psychology of stereotyping*. New York: Guilford Press.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Volume 2. Foundations of social behavior* (pp. 527–561). New York: Guilford Press.

Schwarz, N., & Bless, H. (1992a). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Schwarz, N., & Bless, H. (1992b). Scandals and the public's trust in politicians: Assimilation and contrast effects. *Personality and Social Psychology Bulletin*, 18, 574–579.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Schwarz, N., Münkel, T., & Hippler, H. J. (1990). What determines a "perspective"? Contrast effects as a function of the dimension tapped by preceding questions. *European Journal of Social Psychology*, 20, 357–361.

Seta, J. J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology*, 42, 281–291.

Seta, J. J., Martin, L., & Capehart, G. (1979). Effects of contrast and generalization on the attitude similarity-attraction relationship. *Journal of Personality and Social Psychology*, 37, 462–467.

Seta, J. J., Seta, C. E., & Donaldson, S. (1991). The impact of comparison processes on coactor's frustration and willingness to expend effort. *Personality and Social Psychology Bulletin*, 17, 560–568.

Sherif, M. (1936). *The psychology of social norms*. New York: Harper & Brothers.

Sherif, M. (1948). The necessity of considering current issues as part and parcel of persistent major problems, illustrated by the problem of prejudice. *International Journal of Opinion and Attitude Research*, 2, 63–68.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Sherman, S. J., Presson, C. C., & Chassin, L. (1984). Mechanisms underlying the false consensus effect: The special role of threats to the self. *Personality and Social Psychology Bulletin*, 10, 127–138.

Sherman, S. J., Presson, C. C., Chassin, L., Corty, E., & Olshavsky, R. (1983). The false consensus effect in estimates of smoking prevalence: Underlying mechanisms. *Personality and Social Psychology Bulletin*, 9, 197–207.

Shrauger, J. S., & Patterson, M. B. (1976). Self-evaluation and the selection of dimensions for evaluating others. *Journal of Personality*, 42, 569–585.

Sigelman, C. (1991). Social distance from stigmatized groups: False

consensus and false uniqueness effects on responding. *Rehabilitation Psychology*, *36*, 139–151.

Silvia, P. J., & Duval, T. S. (2001). Objective self-awareness theory: Recent progress and enduring problems. *Personality and Social Psychology Review*, *5*, 230–241.

Smith, R. H. (2000). Assimilative and contrastive emotional reactions to upward and downward social comparisons. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 173–200). New York: Kluwer Academic.

Spears, R., & Manstead, A. S. R. (1990). Consensus estimation in social context. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 1, pp. 81–109). Chichester, UK: Wiley.

Stapel, D. A., & Blanton, H. (2004). From seeing to being: Subliminal social comparisons affect implicit and explicit self-evaluations. *Journal of Personality and Social Psychology*, *87*, 468–481.

Stapel, D. A., & Koomen, W. (1996). Differential consequences of trait inferences: A direct test of the trait-referent hypothesis. *European Journal of Social Psychology*, *26*, 827–837.

Stapel, D. A., & Koomen, W. (1997). Using primed exemplars during impression formation: Interpretation or comparison? *European Journal of Social Psychology*, *27*, 357–367.

Stapel, D. A., & Koomen, W. (1998). When stereotype activation results in (counter) stereotypical judgments: Priming stereotype-relevant traits and exemplars. *Journal of Experimental Social Psychology*, *34*, 136–163.

Stapel, D. A., & Koomen, W. (2000). How far do we go beyond the information given? The impact of knowledge activation on interpretation and inference. *Journal of Personality and Social Psychology*, *78*, 19–37.

Stapel, D. A., & Koomen, W. (2001). I, we, and the effects of others on me: How self-construal moderates social comparison effects. *Journal of Personality and Social Psychology*, *80*, 766–781.

Stapel, D. A., Koomen, W., & van der Plight, J. (1996). The referents of trait inferences: The impact of trait concepts versus actor-trait links on subsequent judgments. *Journal of Personality and Social Psychology*, *70*, 437–450.

Stapel, D. A., Koomen, W., & van der Pligt, J. (1997). Categories of category accessibility: The impact of trait concept versus exemplar priming on person judgments. *Journal of Experimental Social Psychology*, *33*, 47–76.

Stapel, D. A., & Schwarz, N. (1998). The Republican who did not want to become president: Colin Powell's impact on evaluations of the Republican party and Bob Dole. *Personality and Social Psychology Bulletin*, *24*, 690–698.

Stapel, D. A., & Winkielman, P. (1998). Assimilation and contrast as a function of context-target similarity, distinctness, and dimensional relevance. *Personality and Social Psychology Bulletin*, *24*, 634–646.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Suls, J. M., & Wan, C. K. (1987). Social comparison choices for evaluating a sex- and age related ability. *Personality and Social Psychology Bulletin*, *4*, 102–105.

Suls, J., Wan, C. K., Barlow, D. H., & Heimberg, R. G. (1990). The fallacy of uniqueness. Social consensus perceptions of anxiety disorder patients and community residents. *Journal of Research in Personality*, *24*, 415–432.

Suls, J., & Wills, T. A. (Eds.). (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum

Tajfel, H. (1957). Value and the perceptual judgment of magnitude. *Psychological Review*, *64*, 192–204.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, *25*, 79–97.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Social and Clinical Psychology*, *54*, 101–114.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*, 569–575.

Taylor, S. E., Wood, J. V., & Lichtman, R. R. (1983). It could be worse: Selective evaluation as a response to victimization. *Journal of Social Issues*, *39*, 19–40.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–226). New York: Academic Press.

Tesser, A., & Campbell, J. (1980). Self-definition: The impact of the relative performance and similarity of others. *Social Psychological Quarterly*, *43*, 341–347.

Tesser, A., & Campbell, J. (1983). Self-definition and self-evaluation maintenance. In J. Suls & A. Greenwald (Eds.), *Social psychological perspectives on the self* (pp. 1–31). Hillsdale, NJ: Erlbaum.

Thornton, D., & Arrowood, A. J. (1966). Self-evaluation, self-enhancement, and the locus of social comparison. *Journal of Experimental Social Psychology, 2*(Suppl. 1), 40–48.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, *33*, 529–554.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454–463.

Upshaw, H. S. (1962). Own attitudes as an anchor in equal appearing intervals. *Journal of Abnormal and Social Psychology*, *64*, 85–96.

Upshaw, H. S. (1969). Stimulus range and the judgmental unit. *Journal of Experimental Social Psychology*, *5*, 1–11.

Upshaw, H. S., & Ostrom, T. M. (1984). Psychological perspective in attitude research. In J. R. Eiser (Ed.), *Attitudinal judgment* (pp. 23–41). New York: Springer-Verlag.

Van Knippenberg, A. (1978). Status differences, comparative relevance, and intergroup differentiation. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 171–199). London: Academic Press.

Van Knippenberg, A., Wilke, H., & de Vries, N. K. (1981). Social comparison on two dimensions. *European Journal of Social Psychology*, *11*, 267–283.

Volkmann, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 273–294). New York: Harper.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 177–254). New York: Academic Press.

Wallen, R. (1943). Individuals' estimates of group opinions. *Journal of Social Psychology*, *17*, 269–274.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141–208). New York: Academic Press.

Wever, E. G., & Zener, K. E. (1928). The method of absolute judgment in psychophysics. *Psychological Review*, *35*, 466–493.

Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, *2*(Suppl. 1), 27–31.

Wheeler, L., Koestner, R., & Driver, R. E. (1982). Related attributes in the choice of comparison others: It's there, but it isn't all there is. *Journal of Experimental Social Psychology*, *18*, 489–500.

Wheeler, L., Martin, R., & Suls, J. (1997). The proxy model of social comparison for self-assessment of ability. *Personality and Social Psychology Review*, *1*, 54–61.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*, 245–271.

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*, 231–248.

Wood, J. V., Taylor, S. E., & Lichtman, R. R. (1985). Social compari-

son in adjustment to breast cancer. *Journal of Personality and Social Psychology*, *49*, 1169–1183.

Wood, J. V., & van der Zee, K. (1997). Social comparisons among cancer patients: Under what conditions are comparisons upward and downward? In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 299–328). Mahwah, NJ: Erlbaum.

Yzerbyt, V. Y., Schadron, G., Leyens, J. P., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, *66*, 48–55.

Zanna, M. P., Goethals, G. P., & Hill, J. F. (1975). Evaluating a sex-related ability: Social comparison with similar others and standard setters. *Journal of Experimental Social Psychology*, *11*, 86–93.

Zebrowitz, L. A., Kendall-Tackett, K., & Fafel, J. (1991). The influence of children's facial maturity on parental expectations and punishment. *Journal of Experimental Child Psychology*, *52*, 221–238.

Zebrowitz, L., & Lee, S. Y. (1999). Appearance, stereotype-incongruent behavior, and social relationships. *Personality and Social Psychology Bulletin*, *25*, 569–584.

Decisions Constructed Locally

Some Fundamental Principles of the Psychology of Decision Making

ELDAR SHAFIR

Behavioral research into the making of decisions has been influenced by two popular and rather compelling views of the human agent. The first, "folk psychology" view, consists of our intuitive understanding of the decisions that we make and of the factors that motivate and underlie them. The second, normative, "rational agent model," presents a more analytic, a priori, analysis of what it means to make rational choices and has come to dominate much of economics and the rest of the social sciences, as well as the formulation and conduct of policy. Interestingly, part of what has made the normative treatment so appealing has been its general affinity with intuition. Putting aside certain technical requirements that even the most ardent proponents of the rational model recognize as highly idealized, preferences, according to the normative theory, generally are assumed to respect simple rules of stability and well ordering that most naive respondents, upon a moment's reflection, readily endorse. Similarly, people's intuitions regarding value maximization, self-control, planning, situational influences, and the stability and reliability of preferences are, to a first approximation, aligned with normative expectations, even if folk psychology recognizes certain normative assumptions as extreme.

The empirical findings, on the other hand, are often nonnormative and counterintuitive. Not only are preferences often inconsistent with the normative requirements of decision making, they violate simple intuitive expectations as well. And, as in other areas of empirical investigation, observations that are unexpected and counterintuitive tend to trigger further thinking and

research and can lead to new insight and understanding.

In this chapter, we explore some fundamental insights and general principles that characterize our current understanding of decision making. Some of these principles, while integral to our understanding of decision-making processes, are quite general, forming part of a psychological portrayal of human information processing that is not unique to the making of decisions; other insights pertain more specifically to the decision-making context. As with all reviews of this kind, the perspective provided here is somewhat idiosyncratic, heavily influenced by the author's views and predilections; it by no means provides a complete and definitive summary of this rich and diverse field of investigation. (Some good collections of relevant readings are Goldstein & Hogarth, 1997; Holyoak & Morrison, 2005; Koehler & Harvey, 2004; among others.)

In what follows, we consider the tension between normative and descriptive insights that underlies the modern study of decision making. We begin by discussing two prominent features that render people's decision behavior profoundly different from that envisioned by normative treatments. One is the overwhelming tendency to make "local" as opposed to "global" decisions; the other is our proclivity to "construct" our preferences once a decision needs to be made. Following an initial discussion of these two essential characteristics, we proceed to consider various empirical studies that have explored the logic and consequence of these attitudes, as well as some of the psychological traits or processes that

underlie them. In the third section, we consider a number of features that are integral to the decision-making context, including attitudes toward risk and loss and the repeated and temporal aspects of decision making. In the fourth section, we explore further some of the psychological processes that appear to play a prominent role in people's construction of their preferences. These include response compatibility, decisional conflict, reasons, emotions, and personal identity, as well as unconscious influences and automatic processes. The fifth section provides some concluding remarks.

A PERSONAL ANECDOTE

Consider the following brief personal anecdote. As a graduate student, I would walk a few times a week from home to work, and back, along the streets of Cambridge, Massachusetts. One day, while listening to a lecture on "local optima," I realized that I had been doing something remarkable: I had been taking one route every morning, on the way to work, but a different route in the evening, on the way back home. As it turned out, there was no good reason for this inconsistency: There were no inclines, no views to enjoy, no dark or crowded spaces. Rather, in an attempt to walk around an irregularly shaped parking garage situated en route, I took, in each direction, what felt like the most efficient detour. As portrayed in Figure 14.1, because of how the garage was configured, that detour led to one route on the way to school and to an alternate route on the way back home. Not surprisingly, my old Cambridge map shows one route to be somewhat longer than the other. To the extent that my sole purpose was to cover the distance efficiently, I had taken the "wrong" route half the time.

This anecdote captures the essence of the account of decision making that unfolds in this chapter. It is characterized by two distinguishing features that I shall argue are at the core of decision-making behavior. The first is that decisions are made "locally." In my commuting everyday, I had not bothered to plan my entire route ahead of time. Clearly, had I determined the shortest route from home to work, I could then have opted for the same route back from work to home. Instead, I left home with a clear destination and a general direction and proceeded to make little decisions along the way, intended eventually to lead me to where I wanted to go. My behavior was the outcome not of a "global" plan, intended to maximize efficiency, but rather of a series of local decisions, each looking intuitively right but jointly yielding a suboptimal pattern.

The second characteristic of my commuting experience, closely related to the first, was that at each juncture of the route my preferences were not simply revealed (for I did not yet have any), but rather, these preferences were "constructed" at the moment of having to decide. When I arrived at the parking garage (at least on the first days) I had no memory or other available record to consult about whether I ought to take a right or a left.[1] Instead, only once there, with the strangely shaped garage in full display, did I undertake to decide where to turn,
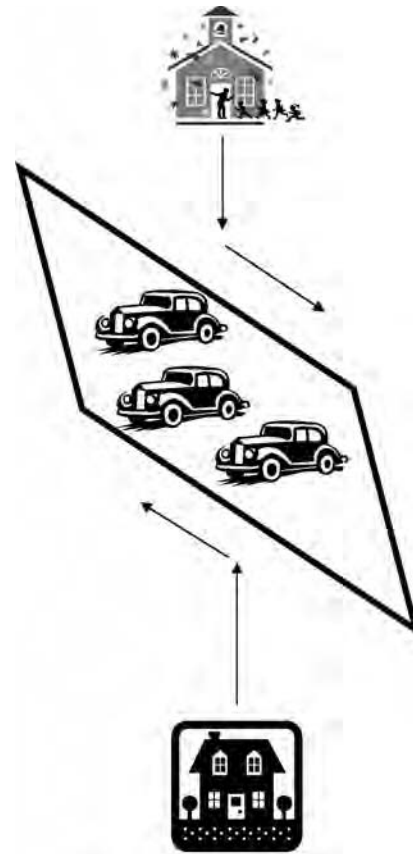
FIGURE 14.1. Local optimization.

and that decision was heavily influenced by the features of the context that were locally salient.

Normative as well as intuitive analyses require, implicitly if not explicitly, that decisions be considered at a level relevant to the outcomes of interest. When making a choice, the available alternatives ought to be evaluated in terms of their ability to deliver the outcomes of interest, with relative disregard for immaterial detail encountered along the way. If my primary aim is to take the shortest route from home to work, it would be a mistake to end up on a longer route just because I took a quick turn on some random corner. Likewise, when shopping for a car, it would be a mistake to opt for a mediocre alternative when I could do better just because the salesman had a warm smile. Similarly, if I want to save as much as I can for retirement, it would be a mistake to choose an option that is likely to save me less in the long run just because it is expected to do better next week.[2] People, however, tend to make their decisions locally, paying great attention to minor deviations from the status quo rather than to their overall, long-term impact. People tend to consider decisions one at a time, rather than as part of some relevant accumulation over many similar instances. And they let various accidental features of the context—whether an option is in their possession or not, the default option, the presence or absence of comparable alternatives—influence their decision, often with long-term consequences.

SOME BASIC FEATURES OF THE DECISION-MAKING CONTEXT

We begin our survey with some basic terminology and a few observations that are considered "classical" by now. They have been extensively reviewed and amply discussed, and they form the core of a successful program of research commonly referred to as *behavioral decision making*. At the core of this program are fundamental psychological principles that underlie the making of decisions, and that account for the systematic discrepancies that arise between the normative treatment and the empirical findings. We begin with a quick review of these (for further exposition and discussion, see Kahneman & Tversky, 2000; LeBouef & Shafir, 2005b; Payne, Bettman, & Johnson, 1992; Shafir & Tversky, 1995).

Risk and Value

Decisions are often risky, in the sense that their outcomes are not known with certainty. In making decisions under risk, people have to consider two factors, the desirability of the outcomes and their probability of occurrence. Classical decision theory is largely concerned with the question of how these factors are, or should be, combined.

Consider a choice between a risky prospect that offers a 50% chance to win $200 (and 50% chance to win nothing) and the alternative of receiving $100 for sure. As it turns out, most people prefer the sure gain over the gamble, although the two prospects have the same expected value. (The expected value of a gamble is a weighted average where each possible outcome is weighted by its probability of occurrence; the expected value of the gamble above is .50 * $200 + .50 * 0 = $100.) Preference for a sure thing over a gamble of equal expected value is called *risk averse*.

Now suppose that we are forced to choose between a prospect that offers a 50% chance to lose $200 (and 50% chance to lose nothing) versus losing $100 for sure. In this problem, most people reject the sure loss and prefer to take an even chance at losing more or nothing. As in the previous choice involving gains, these prospects have the same expected value. This preference for a risky prospect over a sure outcome that has the same expected value is an instance of *risk seeking*. In fact, with the exception of prospects that involve very small probabilities, risk aversion is generally observed in choices involving gains, whereas risk seeking tends to hold in choices involving losses.

The aforementioned risk tendencies can be explained by the notion of diminishing sensitivity. Just as the impact of a candle is greater when it is brought into a dark room than into a room that is well lit, so the impact of an additional $100 is greater when it is added to a gain of $100 than when it is added to a gain of $800. This principle was first formalized by early 18th-century mathematicians, who proposed that subjective value, or utility, is a concave function of money, and that preferences ought to be described using expected utility instead of expected value. According to expected utility, the worth of a gamble offering a 50% chance to win $200 (and 50% chance to win nothing) is .50 * $u($200), where u is the person's concave utility function (where $u(0) = 0$). It follows from such a function that the subjective value attached to a gain of $100 is more than 50% of the value attached to a gain of $200, which entails preference for the sure $100 gain and, hence, risk aversion.

Risk seeking for losses can be explained by assuming similarly diminishing sensitivity to negative outcomes, yielding a subjective value function for losses that is convex. According to such a function, the worth of a gamble that offers a 50% chance to lose $200 is greater (i.e., less negative) than that of a sure loss of $100. That is, .50 * $u(-$200) > u(-$100)$. This implies a risk-seeking preference for the gamble over the sure loss.

Conjoining the concave and convex functions for the contexts of gains and losses, respectively, yields the S-shaped value function illustrated in Figure 14.2. This function forms part of a descriptive theory of choice, known as prospect theory (Kahneman & Tversky, 1979, 1984; Tversky & Kahneman, 1992). The value function of prospect theory has three important properties: (1) it is defined on gains and losses rather than total wealth; (2) it is steeper for losses than for gains; and (3) it is concave for gains and convex for losses.

The first property is closely related to my daily commute behavior portrayed in Figure 14.1. It captures the fact that people tend to treat outcomes as gains and losses defined relative to a neutral reference point, rather than in terms of total wealth. (Just like I turned left or right based on what felt beneficial at that particular point, not on what would prove most efficient overall.) The second property states that losses generally loom larger than corresponding gains, which is captured in Figure 14.1 by a function that is steeper for losses than for gains (i.e., $u($X) < -u(-$X)$). The third property of the value function captures the aforementioned risk attitudes: risk aversion in the domain of gains and risk seeking in the domain of losses.
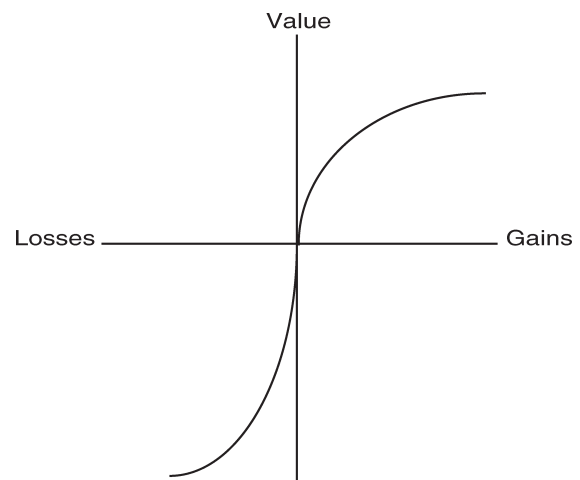
FIGURE 14.2. Prospect theory's value function.

Risk Attitudes and Framing

Although there is a presumption that people are entitled to their own preferences and that those described above are as legitimate as any others, people's attitudes toward risk can soon be seen to generate problematic decision patterns, reminiscent of the ones I showed on my way to school. One such pattern is illustrated by the following pair of decisions (where the bracketed numbers indicate the percentage of respondents who chose each option.)

Problem 1 :
Assume yourself richer by $300 than you are today.
You have to choose between
- a sure gain of $100   [72%]
- 50% chance to gain $200 and 50% chance
  to gain nothing   [28%]

Problem 2:
Assume yourself richer by $500 than you are today.
You have to choose between
- a sure loss of $100   [36%]
- 50% chance to lose nothing and 50% chance
  to lose $200   [64%]

Most subjects presented with the choice involving gains in problem 1 are risk averse, whereas most subjects presented with the choice between losses in problem 2 are risk seeking. The two problems, however, are essentially identical: When the initial sums are added to the expected outcomes, both problems amount to a choice between $400 for sure versus an even chance at $300 or $500.

The predictable responses to the two problems capture a fundamental feature of "local" decision making. They show that people do not combine the initial payment with the anticipated choice outcomes as required by normative analysis. (For, if these were combined into a global representation, the two problems would appear indistinguishable.) And, as a consequence of this local decision evaluation, the same choice problem framed in alternative ways leads to systematically different choices (much as different encounters of the same parking garage led me to take alternate routes). When the same problem framed in slightly different but logically equivalent ways leads to different decisions, this is known as a *framing effect*. Framing effects constitute a violation of an essential requirement of the rational theory of choice, known as *description invariance*, according to which equivalent representations of a choice problem should yield the same preferences (Tversky & Kahneman, 1986).

Framing effects are pervasive and can be observed even when the same respondents answer both versions of a problem (LeBouef & Shafir, 2003). Furthermore, they are found in the choices of both naive and sophisticated respondents. For example, experienced physicians were found to be just as susceptible to the effects of framing—concerning alternative treatments for lung cancer described in terms of mortality rates or in terms of survival rates—as were graduate students or clinic patients (McNeil, Pauker, Sox, & Tversky, 1982).

In fact, the tendency to adopt the locally provided frame is quite general and does not depend on the triggering of alternate risk attitudes. It can lead, among others, to "attribute-framing" effects (Levin, Schneider, & Gaeth, 1998). For example, people are more favorably inclined toward medical procedures when their chances of success, rather than failure, are highlighted. Similarly, ground beef, which can be described as 75% lean or 25% fat, tends to be evaluated more favorably under the former than the latter description (Levin, 1987; see also Levin, Schnittjer, & Thee, 1988). And a community with a 3.7% crime rate tends to be allocated greater police resources than one described as 96.3% "crime free" (Quattrone & Tversky, 1988).

Attribute-framing manipulations affect the perceived quality of items through a minor alteration of their description. Part of the impact of such semantic factors may be due to spreading activation (Collins & Loftus, 1975) wherein positive concepts (e.g., "crime free") activate associated positive concepts and negative concepts activate other negative concepts. In addition, because of the psychophysics of number perception, a 96.3% "crime free" rate suggests that virtually all are law-abiding, whereas a 3.7% as opposed to 0% crime rate appears substantial and suggests the need for intervention (Quattrone & Tversky, 1988). In line with the foregoing discussion, such framing effects are the result not of an all-encompassing, context-independent process of evaluation but, rather, of a "situated cognition," built of low-level perception and immediate association, which leaves decisions largely at the mercy of local, and often accidental, features of specific exposure.

Local versus Global Decisions

To further capture the problematic nature of decisions made locally, recall that what typically matters is the person's overall state, not the specific gains and losses encountered along the way. This, however, is often violated when people make their decisions locally, paying great attention to local deviations from the status quo rather than to their overall, long-term impact.

Whereas normative thinking typically preoccupies itself with the most efficient routes, people often optimize "locally," as they navigate around specific obstacles. However, decisions that are compelling locally often add up to inferior overall patterns. Consider, for example, the following two choices, one involving gains and the other involving losses (taken from Tversky & Kahneman, 1981, p. 454):

Imagine that you face the following pair of concurrent decisions. First examine both decisions, then indicate the options you prefer.

Decision (i). Choose between:
  A: a sure gain of $240   [84%]
  B: 25% chance to gain $1000 and 75% chance
      to gain nothing   [16%]

Decision (ii). Choose between:
  C: a sure loss of $750   [13%]
  D: 75% chance to lose $1000 and 25% chance
       to lose nothing   [87%]

As expected from the foregoing discussion, the majority choice in decision (i), involving gains, is risk averse, and the majority choice in decision (ii), involving losses, is risk seeking. Simple calculation, however, shows that the combination of A and D, which is chosen by a majority of respondents, yields a combined outcome (25% chance to win $240 and 75% chance to lose $760) that is strictly inferior to the unchosen alternative (combination B and C; 25% chance to win $250 and 75% chance to lose $750). Two individually compelling local decisions thus lead to a pattern that is globally inferior.

Many patterns of decision inconsistency documented in the literature would not arise were decisions considered from a more global perspective. Various framing manipulations would be of little consequence were people to go beyond the provided frame and represent the decision outcomes in a canonical manner that is description invariant. Interestingly, unlike some more purely cognitive domains (e.g., vision or language) wherein canonical representations are often formed automatically, the representation of decision outcomes does not typically undergo canonical transformation (cf. Kahneman, 2003). As described next, a key ingredient in the local representation of decision outcomes is what are seen as potential losses.

Loss Aversion

Many decisions present trade-offs between gains and losses, and losses generally loom larger than corresponding gains, as captured in Figure 14.2 by the fact that the value function is steeper in the negative than in the positive domain. This asymmetry in the evaluation of positive and negative outcomes, called loss aversion, occurs because the pain associated with giving up a good is greater than the pleasure associated with obtaining it. An immediate implication of loss aversion is that people will not accept an even chance to win or lose $X, because the loss of $X is more aversive than the gain of $X is attractive. (In fact, people will typically accept an even-chance prospect only when the gain is more than twice as large as the loss; Tversky & Kahneman, 1991.) This yields "endowment effects," wherein the mere possession of a good (such that parting with it is rendered a loss) can lead to higher valuation of the good than if it were not in one's possession. A classic experiment illustrates this point (Kahneman, Knetsch, & Thaler, 1990). Participants were arbitrarily assigned to be *sellers* or *choosers*. Sellers were each given an attractive mug, which they could keep, and asked to indicate the lowest amount for which they would sell the mug. Choosers were not given a mug but were instead asked to indicate the amount of money that the mug was worth to them. (Further procedural details were designed to promote truthful estimates: an official market price, $X, was to be revealed; all those who valued the mug at more than $X received a mug, whereas those

who valued the mug below $X received $X.) Note that all participants, sellers and choosers, faced the same decision: determining a price at which they would prefer money over the mug. Because participants were randomly assigned to be sellers or choosers, standard expectations are that the two groups would value the mugs similarly. Loss aversion, however, suggests that the sellers would set a higher price (for what they were about to "lose") than the choosers. Indeed, sellers' median asking price was twice that of choosers.

Another manifestation of loss aversion is a reluctance to trade, illustrated, for example, in a study in which half of the subjects were given a decorated mug and the rest were given a bar of chocolate (Knetsch, 1989). Later, each subject was shown the alternative gift option and offered the opportunity to trade his or her gift for the other. Because the initial gift allocation was arbitrary and transaction costs minimal, economic theory predicts that about half the participants would exchange their gifts. Loss aversion, on the other hand, predicts that most participants would be reluctant to give up a gift in their possession (a loss) to obtain the other (a gain). Indeed, only 10% of participants chose to trade. This contrasts sharply with standard analysis, in which the value of a good does not change when it becomes part of one's endowment.

Loss aversion emerges as an important empirical generalization that has implications for a wide range of decisions. It promotes stability rather than change, because the disadvantages of departing from the status quo loom larger than the advantages presented by the alternative (Samuelson & Zeckhauser, 1988). It has thus been suggested that loss aversion can hinder the negotiated resolution of disputes. If each disputant sees the opponent's concessions as gains but its own concessions as losses, agreement will be hard to reach because each will perceive itself as relinquishing more than it stands to gain (Kahneman & Tversky, 1995.) A skillful mediator may facilitate agreement by framing concessions as bargaining chips rather than as losses, or by setting all sides' reference points low, thus requiring compromises over outcomes that are mostly perceived as gains (cf. Kahneman, 1992).

Note that loss aversion is an attitude inherently confined to a local setting; it is based on departure from a current reference point, not on the conditions reached after the loss. Thus, a person A, who has just lost $1 million in the stock market may be more upset than person B, who has lost only $10,000, despite the fact that, at the end of this bad day for both, person A may be worth twice as much as person B. In my decision to surround the garage in Figure 14.1, a perceived "loss" corresponded to the reluctance to walk backward rather than forward at that particular juncture; concern with losing ground at that point did not allow for the possibility that the avoidance of such loss may ultimately lead to a longer route overall.

Isolated versus Joint Evaluation

An interesting and systematic pattern of local influence is observed when people evaluate options in isolation as

opposed to when they directly compare them. In general, systematic discrepancies in the weighting of option attributes can emerge when options are evaluated in isolation as opposed to when they are compared. Certain attributes, for example, are difficult to evaluate in isolation; others tend to overwhelm in direct comparisons. In one example, capturing what Hsee (1996) has called an evaluability effect, subjects are invited to consider two secondhand music dictionaries, one with 20,000 entries and a cover that's slightly torn, the other with 10,000 entries and a cover that is like new. When evaluating only one dictionary at a time, respondents, who have little notion of how many entries to expect, are willing to pay more for the dictionary with the new rather than the torn cover. In contrast, when these dictionaries are evaluated concurrently most people prefer the dictionary with more entries, despite its inferior cover (Hsee, 1996; Hsee, Loewenstein, Blount, & Bazerman, 1999).

For another example, consider a job that pays $80,000 a year at a firm where one's peers receive $100,000, compared to a job that pays $70,000 while coworkers are paid $50,000. Consistent with the fact that most people prefer higher incomes, a majority of second-year MBA students who compared the two options preferred the job with the higher absolute income, despite the lower pay relative to one's peers. When the jobs are considered separately, however, the precise attractiveness of one's own salary is hard to gauge, whereas earning less than one's peers renders the former job relatively less attractive than the latter. Indeed, the majority of MBA students who evaluated the two jobs separately anticipated higher satisfaction in the job with the lower salary but the higher relative position, obviously putting greater weight on the latter attribute in the context of separate evaluation (Bazerman, Schroth, Shah, Diekmann, & Tenbrunsel, 1994).

In the same vein, decision principles that are hard to apply in isolated settings may prove decisive in direct comparisons, yielding systematic shifts in attribute weights. Kahneman and Ritov (1994), for example, asked participants for the maximum they were willing to contribute to several environmental programs. One program was dedicated to saving dolphins in the Mediterranean Sea; another funded free medical checkups for farmworkers at risk for skin cancer. When asked which program they would rather support, the vast majority chose the medical checkups for farmworkers, presumably based on the principle that human lives matter more than those of animals. However, when asked separately for the largest amount they would be willing to pay for each intervention, respondents, moved by the animals' touching plight, were willing to pay more for the dolphins than for workers' checkups. In a similar application, potential jurors awarded comparable dollar amounts to plaintiffs who had suffered either physical or financial harm when the cases were evaluated separately. However, in concurrent evaluation, award amounts were substantially larger for physical as opposed to financial harms, affirming the notion that personal harm is perceived as the graver offense (Sunstein, Kahneman, Schkade, & Ritov, 2001).

Attribute weights, which are normatively assumed to remain stable, are thus seen to systematically shift and give rise to patterns of inconsistent preferences. Such shifts are largely the outcome of local decision contexts, which can trigger different comparison and evaluation processes, yielding differential relative attribute weights. As with other localized decisions, this process can lead to suboptimal decisions about outcomes that will be experienced in a different future context. A compelling illustration is provided by Hsee and colleagues (1999) in the context of a hypothetical purchasing decision at an audio store. Suppose that a pair of speakers has superior sound but is aesthetically unattractive relative to a second pair, which looks much better but has a slightly less impressive sound. At the store, with the two options in direct comparison, one is likely to opt for the better sounding speaker that, after all, one might reason, matters more than how a speaker looks. But then, at home with only the chosen speakers and nothing to compare them with, the slight superiority in sound is likely to prove inconsequential, whereas the unattractiveness of the speakers may persist and prove quite noticeable.

Discrepancies between separate versus concurrent evaluation may have profound implications for intuition and for policy. Outcomes in everyday life are typically experienced one at a time: A person lives through one scenario or another. Normative intuitions, on the other hand, often arise from comparative introspection: we entertain a scenario along with its alternatives. When an event triggers reactions that partly stem from its being experienced in isolation, important aspects of that experience may be misconstrued by intuitions that arise from comparative evaluation (for further discussion, see Shafir, 2002).

Repeated Decisions

As illustrated in Figure 14.1 and other previous examples, a local approach to decision making can give rise to decisions that are inconsistent with preferences that emerge from a more "global" evaluation. In essence, a global perspective requires one to ignore features of the decision that are momentarily salient in favor of other, often less salient considerations that have long-run repercussions. This tension can be clearly observed in the context of repeated decisions, where local features of a decision can be in tension with considerations that prove to matter more in the long run.

In fact, decisions that occur on a regular basis are often more meaningful when evaluated "in the long run." The decision to diet, to exercise, or to practice a musical instrument makes very little difference on any single occasion and can only be carried out under a long-term perspective that trumps the person's short-term preference for cake over vegetables, for sleeping late rather than jogging early, or for reading a good book rather than rehearsing a boring musical piece. In many instances that are interpreted as involving myopia, impatience, or the lack of self-control, people fail to take a long-term perspective when evaluating instances of recurring decisions, where the repeated perspective might be more ap-

propriate than regarding each choice as an isolated event.

When offered a 50% chance to win $2,000 and a 50% chance to lose $500, most participants refused to play this gamble once, but the majority were happy to play the gamble five times. Furthermore, when given the choice, most preferred to play the gamble six times rather than five. Apparently, fear of possibly losing the single gamble is compensated for by the high likelihood of ending up ahead in the repeated play. Other participants were asked to imagine already having played the gamble five times (outcome as yet unknown) and were given the option to play one additional such gamble: The majority of participants rejected this additional opportunity. Although participants preferred to play the gamble six times rather than five overall, once they had finished playing five and had readjusted their point of reference, the additional opportunity to play one more gamble was immediately "segregated" and treated as a single instance, which—as is known from the single-gamble version—participants prefer to avoid (Redelmeier & Tversky, 1992).

Along similar lines, physicians often can think of their patients "individually" (i.e., a patient at a time) or "globally" (as groups of patients presenting similar problems). In several studies, Redelmeier and Tversky (1990) found that physicians were more likely to take "extra measures," such as ordering an expensive medical test or recommending an in-person consultation, when they considered the treatment of an individual patient than when they considered a larger group of similarly afflicted patients. Personal concerns loom larger when patients are considered individually than when they are considered "in general," with the latter group more likely to highlight efficiency concerns. Because physicians tend to see patients one at a time, this discrepancy predicts a pattern of individual decisions that is often inconsistent with what these physicians would endorse collectively. For a more mundane example, people report greater willingness to wear a seatbelt—and to support pro-seatbelt legislation—when they are shown statistics concerning lifetime risk of being in a fatal accident, instead of the dramatically lower risk associated with any single auto trip (Slovic, Fischoff, & Lichtenstein, 1988).

Related patterns prompted Kahneman and Lovallo (1993) to argue that decision makers often err by treating each decision as unique, rather than as part of a series of similar decisions made over a lifetime (or, in the case of organizations, made by many workers). They distinguish an "inside view" of situations and plans, characterized by a focus on the peculiarities of the case at hand, from an "outside view," guided by analysis of a large number of relevant cases. Whereas the outside view, based, for example, on base rates, tends to lead to more accurate evaluation, people routinely adopt an inside view, which typically overweighs the particulars of the given case at the expense of more general and reliable (e.g., base rate) considerations. The inside view can thus generate overconfidence (Kahneman & Lovallo, 1993), as well as undue optimism, for example, regarding the chances of completing projects by early deadlines (Buehler, Griffin,

& Ross, 1994). In particular, the myopia that emerges from treating repeated decisions as unique leads to overly bold predictions and to the neglect of considerations that ought to matter in the long run. For example, although they know that past product launches have routinely run over budget and behind schedule, managers may convince themselves that this time will be different because the team is excellent and the product exceptional.

Temporal Considerations

A nontrivial task in the making of decisions, and one that is intimately related to considerations of repeated transactions, is the need to decide how much weight to give to outcomes extended into the future relative to if they were to occur right now. Various forms of uncertainty (regarding nature, one's own tastes, etc.) justify some degree of discounting in the calculation of the present value of future goods. The sum of $1,000 received a year from now is typically worth less than $1,000 received today. As it turns out, observed discount rates tend to be unstable and often influenced by factors, such as the size of the good and its temporal distance, that are not subsumed under standard normative analyses (for reviews, see Ainslie, 2001; Frederick, Loewenstein, & Donoghue, 2002; Loewenstein & Thaler, 1989). For example, although many prefer an apple today over two apples tomorrow, virtually nobody prefers one apple in 30 days over two apples in 31 days (Thaler, 1981). Because discount functions are nonexponential, a one-day delay has greater impact when that day is near than when it is far (see also Loewenstein & Prelec, 1992). Similarly, when asked what amount of money in the future would be comparable to receiving a specified amount today, people typically require about $60 in one year to match $15 now, but they are satisfied with $4,000 in a year instead of $3,000 today. This implies discount rates of 300% in the first case and 33% in the second.

Such discrepancy in the discount rates that are applied to small versus large amounts is, again, reminiscent of my excessive concern with the small detour in front of the parking garage while trying to cover a larger distance. To the extent that we engage in numerous transactions through time, imposing great discount rates on small amounts right now ignores the fact that repeated small amounts will add up to be large in the long run. This yields systematic inconsistency in attitudes toward small decisions as opposed to the full trajectory, which in many instances—my walk to school, for example—ought to count for more.

Excessive discounting turns into myopia, which can often be observed in people's attitudes toward future outcomes (see, e.g., Elster, 2000; Elster & Loewenstein, 1992). Loewenstein and Thaler (1989), for example, discuss an intervention in West Virginia in which the high school dropout rate was reduced by one-third when potential dropouts were threatened with the loss of their driving privileges. This immediate threat had a significantly greater impact than the far more serious but more distant socioeconomic implications of failing to graduate

from high school. In a similar vein, physicians apparently lament the fact that warning about the risk of skin cancer from excessive sun exposure has less impact than warning about sun exposure's tendency to cause acne. In fact, "quit smoking" campaigns have begun to stress the immediate benefits of quitting (quick reduction in chances of heart attack, improved ability to taste foods within 2 days, and so on) more prominently than the substantial long-term benefits (American Lung Association, 2003). Similar reasoning applies in the context of medical self-examinations and the promotion of safe-sex practices, where immediate discomfort or gratification can overwhelm much greater, but temporally distant, considerations (see also Schelling 1980, 1984). Myopia presents an extreme case of decisions made locally, with insufficient regard for more distant and global considerations. It stems from strong impulses triggered by local context that tend to overwhelm the construction of future preferences.

Trope and Liberman (2000, 2003) present compelling evidence suggesting that temporal distance systematically alters people's mental representations of future events and that this systematically changes the valuation of those events. They propose that changes in preference occur because temporal distance from actual engagement in an activity changes the way the activity is represented. Specifically, the more temporally distant events are, the more likely they are to be represented in terms of a few abstract features that convey their perceived essence ("high-level construals") rather than in terms of more concrete and secondary details ("low-level construals"), which gain in weight as the events grow closer. Indeed, these authors suggest that construal levels underlie a broad range of evaluative and behavioral consequences of psychological, not merely temporal, distance from events. We next turn to some other, nontemporal features that impact preference construction.

SOME ELEMENTS OF DECISION CONSTRUCTION

In line with construal processes in other areas of psychology, preferences appear to be constructed, not merely revealed, in the making of decisions (Slovic, 1995). The study of choice, as Payne, Bettman, and Schkade (1999) have suggested, often looks more like architecture, with preferences constructed in context, rather than as archaeology, with preferences being merely uncovered. And while this has major implications for the study of normative–descriptive tensions, it also sheds light on the processes that underlie the evaluation of alternatives.

Compatibility

Because people often do not have reliable views about the relative importance of various dimensions, the weights assigned to those dimensions can be influenced by relatively immaterial changes in the task, the description, and the nature of the options under consideration. Even simple monetary gambles, for example, differ on

payoffs and the chance to win. Respondents' preferences among such gambles can be assessed in different, but logically equivalent, ways (see Schkade & Johnson, 1989, for a review). Thus, participants may be asked to choose among the gambles, or, alternatively, they may estimate their maximum willingness to pay for each gamble. Although logically equivalent, these procedures often result in differential weightings of attributes and, consequently, in inconsistent preferences.

In line with the principle of stimulus–response compatibility, long recognized by students of human perception and motor performance (Fitts & Seeger, 1953; Wickens, 1984; see Shafir, 1995, for discussion), the weight given an attribute can be enhanced by its compatibility with a required response. Thus, a gamble's potential payoff can be weighted more heavily in a pricing task (in which both the price and the payoff are expressed in the same—monetary—units) than in choice (Slovic, Griffin, & Tversky, 1990; Tversky, Sattath, & Slovic, 1988; although see Frederick & Shafir, 2005). For example, consider two gambles: one offers an 8/9 chance to win $4 and the other a 1/9 chance to win $40. People typically choose the high-probability gamble but assign a higher price to the high-payoff gamble, thus expressing conflicting preferences (Grether & Plott, 1979; Lichtenstein & Slovic, 1971, 1973; Tversky, Slovic, & Kahneman, 1990). This pattern is known as the *preference reversal phenomenon* (Slovic & Lichtenstein, 1983). It has been replicated among professional gamblers in a Las Vegas casino (Lichtenstein & Slovic, 1973), and it has also been documented in numerous studies involving nonmonetary options, including choice between highway safety programs, job candidates, and interventions intended to address environmental problems (Kahneman & Ritov, 1994; Slovic et al., 1990; Tversky et al., 1988).

For another type of response compatibility, imagine having to choose or, alternatively, having to reject, one of two options. Logically speaking, the two tasks are interchangeable: If people prefer one option they will reject the second, and vice versa. However, people tend to focus on the relative strengths of options (more compatible with choosing) when they choose, and on weaknesses (compatible with rejecting) when they reject. As a result, options' positive features (the pros) loom larger in choice, whereas their negative features (the cons) are weighted relatively more during rejection. In one study, respondents were presented with pairs of options—an enriched option, with various positive and negative features, and an impoverished option, with no real positive or negative features (Shafir, 1993). For example, consider two vacation destinations: one with a variety of positive and negative attributes, such as gorgeous beaches and great sunshine but cold water and strong winds, and another that is neutral in all respects. Some respondents were asked which destination they preferred; others decided which to forgo. Because positive features are weighed more heavily in choice and negative features matter relatively more during rejection, the enriched destination was most frequently chosen and rejected. Overall, its choice and rejection rates summed to 115%, significantly more than the impoverished destination's

85%, and more than the 100% expected if choice and rejection were complementary (see also Downs & Shafir, 1999; but see Wedell, 1997, for an alternative interpretation).

Thinking back to our opening anecdote, it is worth noting that the cognitive principle of compatibility is itself highly "localized" in character. The compatibility between an option's features and the response is a function of the specific features that are locally prominent and the response that is required at the moment of decision. The prominent features may change with time or with perspective, and they may become more or less important with compounding or with frame. But such potential changes are unlikely to have any impact on the compatibility that emerges during the processing of the information that is salient at the moment of decision.

Decisional Conflict

Interestingly, according to the classical normative view, options are each assigned an attractiveness, or "utility" rating. In this view, there is not much room for decisional conflict. Regardless of how utilities are computed, once they have been assigned, independently for each alternative, the person is predicted to choose the option assigned the highest subjective utility. In contrast, according to the picture that emerges from the present discussion, choices can be hard to make. People need to construct their preferences in the context of making a decision, and they often do so by looking for a good reason, compelling rationale, for choosing one option over another. At times, compelling rationales are easy to come by and to articulate, whereas other times no easy rationale presents itself, rendering the conflict between options hard to resolve. Such conflict can prove aversive and can lead people to postpone the decision or to select a "default" alternative. The proclivity to resort to compelling rationales in an attempt to minimize conflict appears benign; nonetheless, it can generate preference patterns that are fundamentally different from those predicted by normative accounts based on value maximization.

Tversky and Shafir (1992a) presented subjects with pairs of options, such as bets varying in probability and payoff (after they were first invited to familiarize themselves with all options, to avoid the triggering of different expectations). Subjects could choose one of the two options or, instead, they could pay to add an additional option from a set of alternatives observed earlier. One group received the following two options:

Conflict:
x) 65% chance to win $15
y) 30% chance to win $35

Others received a similar choice except that option y was replaced by option x':

Dominance:
x) 65% chance to win $15
x') 65% chance to win $14

Note that gamble x is common to both conditions, whereas x' is strictly dominated by x, whereas y is a viable alternative. A moment's study should reassure the reader that the *conflict* condition presents an objectively superior set of alternatives to those offered in the *dominance* condition.

Subjects were asked to indicate whether they wanted to add another gamble or select between the available alternatives. Their preferred gamble from the resulting set (with or without the added option) was played out and their payoff was proportional to the amount of money they earned minus the fee they paid for the added gambles.

On average, subjects requested an additional alternative 64% of the time in the conflict condition, and only 40% of the time in the dominance condition ($p < .05$). The tendency to search for additional options, in other words, was greater when the choice among alternatives was harder to rationalize, although the quality of available options was superior, than when there was a compelling reason and the decision was easy, despite the fact that the available options were now less good.

According to standard assumptions of value maximization, a person should search for additional alternatives if and only if the expected (subjective) value of searching exceeds that of the best alternative currently available. Because the best alternative offered in the dominance condition was also available in the conflict condition, value maximization implies that the percentage of subjects who seek an additional alternative cannot be greater in the conflict than in the dominance condition.

It appears that the search for additional alternatives depends not only on the value of the best available option but also on the difficulty of choosing among the available options. In situations of dominance, for example, there are clear and indisputable reasons for choosing one option over another (e.g., "This bet offers an equal chance to win more!"). Having a compelling argument for choosing one of the options reduces the temptation to look for additional alternatives. When the choice involves conflict, on the other hand, reasons for choosing any one of the options are less immediately available and the decision is more difficult to justify (e.g., "Should I go for a greater chance to win less, or for a smaller chance to win more?"). In the absence of compelling reasons for choice, there is a greater tendency to search for other alternatives.

As the foregoing pattern clearly illustrates, decision conflict is a sentiment that is largely a feature of the local context; it is determined by the ease or difficulty of the present decision and can be quite independent of the actual worth of the available alternatives. Thus, a good alternative, when placed next to another good alternative, can make for a difficult decision, whereas a weaker alternative, when placed next to one that is substantially worse, can make the decision appear easy. The option the person ends up with, therefore, can be largely an accident of the particular set of alternatives encountered at the moment of decision. Reminiscent of my walk to school (Figure 14.1), where I had optimized not the en-

tire route but, rather, the specific detour around the garage standing in the way, so the subjects above appear to be choosing options as a function not of an overall preference ordering but of the specific decision conflict produced by the locally available alternatives.

Conflict and the Status Quo

One way to avoid conflict in choice is to opt for what appears like no choice at all, namely, the status quo. In one example (Tversky & Shafir, 1992a), participants who were purportedly looking to buy a CD player were presented with a Sony player that was on a 1-day sale for $99, well below the list price. Two-thirds of the participants said they would buy such a CD player. Another group was presented with the same Sony player, and also with a top-of-the-line Aiwa player for $159. In the latter case, only 54% expressed interest in buying either option, and 46% preferred to wait until they learned more about the various models. The addition of an attractive option increased conflict and diminished the number of people who ended up with either player, despite the fact that most preferred the initial alternative to the status quo. This violates what is known as the regularity condition, which states that the "market share" of an existing option—in this case, the status quo—cannot be increased by enlarging the offered set (see also Tversky & Simonson, 1993).

A related pattern was documented in an upscale grocery store, where shoppers encountered tasting booths that offered the opportunity to taste any of 6 jams in one condition, or any of 24 jams in the second (Iyenger & Lepper, 2000). In the 6-jams condition 40% of shoppers stopped to have a taste and, of those, 30% proceeded to purchase a jam. In the 24-jam condition, a full 60% stopped to taste, but a mere 3% purchased. Presumably, the conflict generated by so many options proved hard to resolve. In a similar vein, Huberman, Iyengar, and Jiang (2004) found that employees' participation in 401(k) retirement savings plans drops as the number of fund options proposed by their employer increases.

Decisional conflict tends to favor default alternatives, much as it advantages the status quo. In one study, students agreed to fill out a questionnaire in return for $1.50. Upon completion of the questionnaire, half of the respondents were offered the opportunity to exchange the $1.50 (the default) for one of two prizes: a metal *Zebra* pen, or a pair of plastic *Pilot* pens. The remaining subjects were only offered the opportunity to exchange the $1.50 for the *Zebra*. The pens were shown to subjects, who were informed that each prize regularly costs about $2. The results were as follows. Twenty-five percent opted for the payment over the *Zebra* when the *Zebra* was the only alternative offered, but a reliably greater 53% chose the payment over the *Zebra or* the *Pilot* pens when both options were available (Tversky & Shafir, 1992a). Whereas the majority of subjects took advantage of the opportunity to obtain a valuable alternative when only one was offered, the availability of competing valuable alternatives increased the tendency to retain the default option.

Related effects have been documented in decisions made by expert physicians and legislators (Redelmeier & Shafir, 1995). In one scenario, due to caseload in the operating room, neurologists and neurosurgeons were asked to decide which of several patients awaiting surgery ought to be operated on first. Half the respondents were presented with information about two patients, a woman in her early 50s and a man in his 70s. Others saw the same two patients along with a third, a woman in her early 50s who was highly comparable to the first, so that it was difficult to think of a rationale for choosing either woman over the other. As predicted, more physicians (58%) chose to operate on the older man in the latter version, where the two highly comparable women generated decisional conflict, than in the former version (38%), in which the choice was between only one younger woman and the man. (Note that chance would have acted in the opposite direction.)

Adherence to the default or status quo has also been observed in a couple of naturally occurring "experiments." One was in the context of insurance decisions, when New Jersey and Pennsylvania both introduced the option of a limited right to sue, entitling automobile drivers to lower insurance rates. The two states differed in what was offered as the default option: New Jersey motorists needed to acquire the full right to sue (transaction costs were minimal: a signature), whereas in Pennsylvania, the full right to sue was the default, which could then be forfeited in favor of the limited alternative. Whereas only about 20% of New Jersey drivers chose to acquire the full right to sue, approximately 75% of Pennsylvania drivers chose to retain it. The difference in adoption rates had financial repercussions estimated at nearly $200 million (Johnson, Hershey, Meszaros, & Kunreuther, 1993). A second naturally occurring "experiment" was recently observed in Europeans' decisions regarding being potential organ donors (Johnson & Goldstein, 2003). In some European nations drivers are by default organ donors unless they elect not to be, whereas in other, comparable European nations they are, by default, not donors unless they choose to be. Observed rates of organ donors are almost 98% in the former nations and about 15% in the latter, a remarkable difference given the low transaction costs and the significance of the decision.

The addition of options can generate conflict and increase the tendency to refrain from choosing. Other options, on the other hand, can lower conflict and increase the likelihood of making a choice. *Asymmetric dominance* refers to the fact that in a choice between options A and B, a third option, A′, can be added that is clearly inferior to A (but not to B), thereby increasing the choice likelihood of A (Huber, Payne, & Puto, 1982). For example, a choice between $6 and an elegant pen presents some conflict for participants. But when a less attractive pen is added to the choice set, the superior pen clearly dominates the inferior pen, thus providing a rationale for choosing the elegant alternative and increasing the percentage of those choosing the elegant pen over the cash. Along related lines, a *compromise effect* has been observed

wherein the addition of a third, extreme option makes a previously available option appear as a reasonable compromise, thus increasing its popularity (Simonson, 1989; Simonson & Tversky, 1992).

Standard normative accounts do not deny conflict, nor, however, do they assume any direct influence of conflict on choice. (Under utility maximization, there does not appear to be much room for conflict: Either utility differences are large and the decision is easy, or they are imperceptibly small and the decision is of little import.) In actuality, people are concerned with making the "right" choice, which gives decisional conflict an influence considerably greater than suggested by considerations of value maximization. Conflict is an integral aspect of decision making, and the phenomenology of conflict, which can be altered via the addition or removal of alternatives, yields systematic choice patterns that often violate standard normative predictions.

What is most remarkable, and lies at the heart of the tension with the normative account, is the "local" character, the heavy context dependency, of the observed choices. The normative account assumes a well-behaved ordering of preferences, with any observed choice revealing a glimpse into that order. In contrast, people's actual choices are often the result of the local resolution of conflict between alternatives, with the option chosen often not the option that would have been chosen had context differed by just a little, and often in rather trivial ways. A major contributor to the local construction of preference is people's reliance on what feel like good reasons or compelling rationales for making a decision. We turn to a brief consideration of such reasons next.

Reasons

Reminiscent of the notion of the "need for closure" (Kruglanski & Webster, 1996), people confronted with nontrivial decisions like to feel like they have made the "right choice." In an attempt to make the right choice, people often look for good reasons, for compelling arguments, for choosing one option over another (Shafir, Simonson, & Tversky, 1993). Relying on good reasons seems like sound practice; the converse, making a choice without good reason, seems unwise. At the same time, abiding by this practice can be problematic because the reasons that come to mind often are heavily context dependent, limited to what is introspectively accessible, and not necessarily those that guide, or ought to guide, the decision at hand. For example, participants who were asked to analyze *why* they felt the way that they did about some jams they had a chance to taste showed less agreement with "expert" ratings of the jams than did those who merely stated their preferences without explanation (Wilson & Schooler, 1991). Preference can be altered in line with reasons that come readily to mind, but those reasons may be heavily influenced by salience, availability, or local context. (Further inquiry into the logic of reason salience is likely to bring forth a panoply of other, nuanced considerations, such as the distinction between the availability of reasons—for example, a comparison to other options, which may not be available when the op-

tion is considered in isolation—as opposed to the salience or accessibility of reasons, which may be influenced by description or method of elicitation; see Higgins, 1996.) Focusing on a set of temporarily available or salient reasons can cause one to lose sight of other, perhaps more valid considerations (Wilson, Dunn, Kraft, & Lisle, 1989).

In fact, people are not always aware of their reasons for acting and deciding (see, e.g., Bargh, 1997; Nisbett & Wilson, 1977). In one example, participants, presented with four identical pairs of stockings and asked to select one, showed a marked preference for the option on the right end of the display. However, despite the evidence that choice was governed by position, no participant mentioned position as the reason for his or her choice. Respondents easily generated "reasons" (in which they cited attributes such as texture), but the reasons they provided bore little resemblance to those that actually guided choice (Nisbett & Wilson, 1977). Related work by Schwarz and colleagues on the experience of reason accessibility (e.g., Sanna, Schwarz, & Small, 2002; Schwarz, 1998; Schwarz et al., 1991) has documented paradoxical patterns wherein the availability of a greater number of reasons that felt more difficult to generate yields more reluctant decisions than a small subset of those reasons produced easily.

Because nuanced factors that are salient in the local context can determine which reasons emerge as most apparent, a reliance on reasons can induce great preference malleability. In one study (Tversky & Shafir, 1992b), college students were asked to imagine that having just taken and passed a difficult exam, they now had an opportunity to buy an attractive vacation package over the Christmas holidays at a low price. Respondents faced the following options: They could choose to buy the vacation package, they could forego the vacation package, or they could pay a $5 fee to defer the decision by a day. The majority elected to buy the vacation package, and less than a third elected to delay the decision. A second group was asked to imagine that they had taken the exam and failed and would need to retake it after the Christmas holidays. They were then presented with the same choice, and as before, the majority elected to buy the vacation package, and less than a third preferred to defer. However, when a third group of participants was to suppose they did not know whether they had passed or failed the exam, the majority preferred to pay to defer the decision until the next day, when the exam result would be known, and only a minority were willing to commit to the trip without knowing. Apparently, participants were comfortable booking the trip when they had clear reasons for the decision—celebrating when they passed the exam or recuperating when they had failed—but were reluctant to commit when their reasons for the trip were uncertain. This pattern, which violates the *sure thing principle* (Savage, 1954), has been documented in a variety of contexts, including gambling and strategic interactions as manifested in prisoner's dilemma games (Shafir, 1994; Shafir & Tversky, 1992).

The tendency to focus on information that is unavailable at the moment of decision, as suggested by the fore-

going studies, can lead to delays that have a significant impact. Consider the following scenario (Bastardi & Shafir, 1998):

> For some time, you have considered adding a compact disc (CD) player to your stereo system. You now see an ad for a week-long sale offering a very good CD player for only $120, 50% off the retail price. Recently, however, your amplifier broke. You learn that your warranty has expired and that you have to pay $90 for repairs.

One group (the "simple" condition) was asked whether they would buy the CD player during the sale, and the vast majority (91%) said they would. Another ("uncertain") group was presented with the same scenario but were told that they would not know until the next day whether their warranty covered the $90 repairs. These respondents could opt to wait until the following day (when they knew about the warranty) to decide whether to buy the CD player. Sixty-nine percent elected to wait. Those who chose to wait then learned that the warranty had expired and would not cover repairs. Upon receiving the news, the majority decided not to buy the CD player. Note that this contrasts sharply with the unequivocal choice to buy the CD player when the $90 repair costs were a given. Although they faced the same decision, only 55% (over all those who waited and those who did not) chose to buy the CD player in the uncertain condition, when they did not know but could pursue information about the repair costs, compared with 91% in the certain condition, when repair costs were known from the start. The decision to pursue information that is locally unavailable can focus attention on the information obtained and thereby trigger reasons for making the choice that would have had little impact otherwise, ultimately distorting preference (Bastardi & Shafir, 1998). Similar patterns have been replicated in a variety of contexts, including one involving professional nurses in a renal-failure ward, more of whom expressed willingness to donate a kidney (to a hypothetical relative) when they had purportedly been tested and learned that they were eligible than when they had known they were eligible from the start (Redelmeier, Shafir, & Aujla, 2001). A reliance on reasons in choice renders decision makers susceptible to a variety of contextual, procedural, and circumstantial nuances that make alternative potential reasons salient and thus help shape preferences, often in discrepant and possibly suboptimal ways.

Transient Frames of Mind

Emotions

Emotional reactions are another major facilitator in allowing local contexts to play a prominent role in determining preference. Much of literature and the arts are devoted to the eternal tension between passion and reason and the influence that heightened states of arousal can have on actions that conflict with long-term interests. But at a more mundane level as well, emotional reactions, often undetected, can contribute to a variety of local influences on decision making. Indeed, transient moods influence choice and judgment in ways that neither rationality assumptions nor intuition predicts (Zajonc, 1980; Zajonc & Markus, 1982). Negative moods, for example, can increase the perceived likelihood of bad outcomes and the frequency of undesirable events (such as homicides) and to decrease judged life satisfaction, while positive moods act in the opposite direction (Johnson & Tversky, 1983, Schwarz & Clore, 1983). Furthermore, those in positive moods often engage in attempts at "mood maintenance," among other things through greater risk aversion (Isen & Geva, 1987; Isen, Nygren, & Ashby, 1988). In a related vein, Trope and his colleagues (e.g., Trope, Ferguson, & Raghunathan, 2000) have explored the influence of mood as a resource that serves in coping with emotionally challenging messages. Raghunathan and Trope (2002), for example, document the differential impact of a good mood when processing counterattitudinal messages that are or are not relevant to the decision maker.

Interestingly, moods with the same valence can have differential effects on judgment; for example, anger and fear, both negative emotions, seem to have different effects on individuals' appraisal of control and, consequently, yield optimism and pessimism about future risks, respectively (Lerner & Keltner 2001). Similarly, Raghunathan and Pham (1999) have suggested that sad individuals tend to be more risk prone, whereas anxious individuals are more risk averse. They attribute these tendencies to the notion that anxiety and sadness convey different information to the decision maker and prime different goals, with anxiety promoting an implicit goal of uncertainty reduction and sadness an implicit goal of reward achievement. In fact, the salience of alternate goals, often induced by mood, is another contributor to the effects of local contexts, with different goals often primed by contextual nuances and interim subgoals possibly drawing attention away from goals that are higher in the hierarchical goal structure (Kruglanski et al., 2002; Shah & Kruglanski, 2002; Trope, Igou, & Burke, 2006).

An "affect heuristic" has been proposed, according to which spontaneous and effortless judgments are often made through quick consultation of images marked by positive and negative affective feelings (Finucane, Alhakami, Slovic, & Johnson, 2000; Slovic, Finucane, Peters, & MacGregor, 2002). The role of emotional reactions can be witnessed, for example, in the inverse relationship commonly observed between perceived risks and benefits, such that activities that are thought by people to have great benefits are seen by those same people as presenting few risks, and vice versa. Typical of a heuristic outcome, this unlikely inverse relationship, purportedly mediated by affect, is strengthened under time pressure (Finucane et al., 2000; Fischhoff, Slovic, Lichtenstein, Reid, & Combs, 1978). Because, in line with an affect heuristic, emotional reactions can generate powerful associations, respondents show a relative insensitivity, once emotionally powerful events are considered, to these events' actual probability of occurrence (Rottenstreich & Hsee, 2001). In a similar vein, both the perceived frequency of common events and the perceived likelihood of risks such as nuclear power are re-

lated to the amount of dread that they arouse (Fischhoff et al., 1978; Lichtenstein, Slovic, Fischoff, Layman, & Combs, 1978).

Emotionally evaluative responses can have a non-negligible effect on decision. For one example, mere cognitive effort can generate negative emotions and influence the probability that the alternative producing such effort will be chosen (Garbarino & Edell, 1997). For a different example, participants are apparently willing to pay more to insure, and are more likely to seek compensation for, an item that is emotionally meaningful than for an emotionally neutral but equally valuable item (Hsee & Kunreuther, 2000). More generally, Loewenstein, Weber, Hsee, and Welch (2001) suggest that "anticipatory emotions" (e.g., emotional reactions to potential outcomes in a risky situation) can influence the cognitive appraisal of decision situations and can impact choice. Shiv and Fedorikhin (1999) consider situations in which consumers are influenced by automatically evoked affect as well as by more controlled cognitions, and they present findings suggesting that when processing resources are limited, spontaneous affective reactions have a greater impact relative to cognitions compared to when the availability of processing resources is high.

Transient emotions, often triggered by local contextual factors, can thus influence the construction of preference. Grocery shopping while very hungry, for example, is likely to lead to purchases that would not have been made under normal circumstances (cf. Loewenstein, 1996). But, as with other contextual influences, even when people are aware of being in the grip of a transient emotion, they typically fail to "correct" adequately. In one study, for example, respondents were asked to predict whether they would be more bothered by thirst or by hunger if trapped without food and water. Some were asked before exercising (when they were not especially thirsty) whereas others were approached immediately after exercising (and, thus, were thirsty). Postexercise, 92% indicated that they would be more troubled by thirst than by hunger, whereas preexercise only 61% did (Van Boven & Loewenstein, 2003). In general, people tend to underestimate the degree to which various contextual changes will impact their sentiments and preferences (e.g., Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Van Boven, Dunning, & Loewenstein, 2000). This contributes to what look like myopic decisions, as people honor present inclinations, not fully appreciating the extent to which these may be attributable to local factors that may soon change.

Automaticity and Priming

A variety of priming effects and automatic processes further contribute to decisions that are responsive to highly local influences, which often have little relevance to eventual consumption. At one extreme are phenomena such as mere priming, where mere repeated exposure to objects, even subliminally, can increase their liking (Bornstein, 1989; Zajonc, 1968). In a classic priming study, participants received a "word perception test" in which either creativity, reliability, or a neutral topic was primed. Participants then completed an ostensibly unrelated "product impression" survey that gauged their opinions of various cameras. Cameras advertised for their creative potential were rated more attractive by those primed for creativity than by those exposed to words related to reliability or a neutral topic (Bettman & Sujan, 1987). Momentary priming can influence ensuing preferences, rendering salient criteria that would otherwise have been considered less important. Due to the highly local nature of such salience effects, product consumption is often likely to occur long after such criterion salience has dissipated (see Mandel & Johnson, 2002; Verplanken & Holland, 2002).

Then, there are a variety of automatic and imperceptible influences, where, for example, a momentary, context-induced physical reaction can influence an ensuing decision. Thus, people who happened to have a waitress touch them lightly on the shoulder are found to tip higher than those who were not touched (Crusco & Wetzel, 1984; Schwarz, 1990, Schwarz & Clore, 1983.) Or, for another example, male participants who encounter a woman on a fear-arousing suspension footbridge are likely, due to misattributed arousal produced by the wobbly bridge, to consider the woman more attractive and are consequently more likely to follow up with a phone call than others who meet the woman before the bridge (Dutton & Aron, 1974). The judged funniness of cartoons can be influenced by the perceived state of one's zygomaticus muscle (involved in smiling) in contexts in which smiling is either facilitated or inhibited, and evaluative processes can be influenced by unrelated tasks such as the Isometric flexion ("approach behavior") or extension ("avoidance behavior") of the upper arm (Neumann & Strack, 2000).

Unavoidable "local" contextual influences occur also during simple interpersonal interaction, and they have been studied, and used, by marketing researchers. Consider, for example, the effects of hypothetical questions on respondents' subsequent decisions. One line of investigation has shown that people's prediction of their future behavior, although inaccurate, can affect their actual behavior. In one experiment (Sherman, 1980), college students were asked to write counterattitudinal essays. In a prior, seemingly unrelated survey, half the students were asked to predict whether they would comply with such a request, and many predicted they would not. The eventual rate of compliance among those who made this prediction was much lower than among those who had not made an earlier prediction. Subjects had thus mispredicted their own behavior (because many would have written the essay had they not been asked to predict in advance). Nonetheless, the actual rate of compliance was close to that predicted. In effect, people went on to decide in a manner consistent with their own mispredictions. Related research has shown the potential role of such self-erasing errors in increasing voter turnout simply by asking people to predict whether they will vote (Greenwald, Carnot, Beach, & Young, 1987).

Even when faced with questions that are purely hypothetical, respondents are unable to prevent a substantial biasing effect on their behavior, particularly when the

questions appear relevant (Fitzsimons & Shiv, 2001). Thus, gauging attitudes toward consumer products can increase attitude accessibility and impact consumer behavior (Chapman, 2001; Fazio, Powell, & Williams, 1989). For example, Morwitz, Johnson, and Schmittlein (1993) found that merely asking consumers whether they intended to purchase an automobile or a personal computer increased their subsequent purchase rate. Follow-up interviews suggest that the effects of hypothetical questions on choice occur beyond awareness and, as a result, are quite difficult to counteract.

A rich and fascinating literature documents the many ways that mere exposure, simple priming, subliminal perception, and unconscious inferences alter judgment and choice. A common theme running through many of these phenomena is the fact that they are heavily context sensitive, induced by nuances that can shift at any moment. The extent to which they impact the local construction of preferences may have little to do with the contexts surrounding consumption. The person touched on the shoulder leaves a larger tip without receiving better service, and the man who calls the person he met on the bridge may be disappointed when they meet again on firm ground. Reminiscent of my detour around the parking garage (Figure 14.1), these are actions that often may prove tangential to one's greater aims and occasionally may cause distortion and error.

Identities

Constructed preferences can further fluctuate as a function of personal identities that happen to be locally salient. A working mother, for example, might think of herself primarily as a mother when in the company of her children but may see herself primarily as a professional while at work. The list of potential identities is extensive (Turner, 1985), with some of a person's identities (e.g., "mother") conjuring up strikingly different values and ideals from others (e.g., "CEO").

Identity salience has been shown to affect various behaviors, including resistance to persuasion (Kelley, 1955), reactions to advertisements (Forehand, Deshpandé, & Reed, 2002), and the rating of consumer products (Reed, 2004). In one study, Asian American women (who hold identities—Asian vs. woman—that entail conflicting expectations about mathematical ability) scored higher on a math test after completing a brief survey that evoked their ethnicity than did those who first completed a survey that evoked gender (Shih, Pittinsky, & Ambady, 1999).

Along similar lines, decisions can be seen to depend on local and happenstance fluctuations in identity salience. In one study, college students whose "academic" identities had been triggered were more likely to opt for more academic periodicals (e.g., *The Economist*) than were those whose "socialite" identities had been made salient. Similarly, Chinese Americans whose American identities were evoked adopted more stereotypically American preferences (e.g., for individuality and competition over collectivism and cooperation) compared to when their Chinese identities had

been triggered (LeBoeuf, 2002; LeBoeuf & Shafir, 2005a).

Whereas these findings are consistent with research on the fluidity of the self-concept (cf. Higgins, 1987; Markus & Kunda, 1986), individual personality traits can also have an impact that depends on local context. As documented by Higgins and his colleagues, the fit between the manner of making a decision and the person's regulatory orientation can influence people's level of engagement with and their evaluation of outcomes. For example, participants assigned a higher price to a coffee mug if they had chosen it using a strategy that fit their orientation (eager strategy/promotion; vigilant strategy/prevention) than a strategy that did not fit (Avnet & Higgins, 2006; Higgins, 2000; Higgins, Idson, Freitas, Spiegel, & Molden, 2003).

Although choice and evaluation are typically expected to reveal stable and coherent patterns that correspond to the person's overall preference rankings, evoked identities tend to activate concepts and priorities that are associated with particular tastes and values, which, in turn assimilate preferences to the evoked identity (cf. Bargh, Chen, & Burrows, 1996; Higgins, Rholes, & Jones, 1977). Preference thus tends to align with currently salient identities, yielding predictable tension anytime there is a mismatch between the identity that does the choosing and the one likely to do the consuming, as when a parent, while at work, commits to an evening meeting only to regret missing her child's soccer game once back at home.

Utility Prediction

Some recent research has explored the question whether people actually maximize the expected experienced utility of a decision (i.e., the hedonic experience that the decision will bring) as opposed to merely the decision utility (i.e., the utility perceived at the moment of decision) (Kahneman, 1994). In general, utility mispredictions are common. When predicting the impact on their lives of various events, people tend to focus too heavily on those specific events, consequently overestimating the impact these will have on their lives and on their life satisfaction (Schkade & Kahneman, 1998; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000). In a similar vein, thinking about their own mood appears to increase decision makers' concern with the hedonic consequences of decisions, thereby promoting mood-regulatory activities and altering preference, possibly in favor of otherwise less preferred alternatives (Caruso & Shafir, 2005).

In making such forecasts, people tend, among other things, to neglect the eventual dissipation of satiation (Simonson, 1990); they often misremember options previously encountered due to inferences made at present or to the overweighting of final moments (Kahneman, Frederickson, Schreiber, & Redelmeier, 1993; Mather, Shafir, & Johnson, 2000), and they often fail to anticipate increases in liking due to mere exposure or endowment (Kahneman & Snell, 1992; Van Boven et al., 2000) and tend to underappreciate the extent to which they will be able to maintain a level of satisfaction in the face of ad-

versity (Gilbert et al., 1998, Kahneman, Diener, & Schwarz, 1999).

One common cause for the misprediction of future utility is the tendency to anchor on the present and to focus heavily on the change between how things are now and how they will become, as opposed to how they might be, in steady state, once they have changed. Like preferences, utility predictions are often formed during the decision making process, and present-biased utility expectations are likely to result in courses of action that often fail to maximize future well-being.

CONCLUDING REMARKS

This chapter has considered people's proclivity to make "local" decisions. Such decisions are shaped by considerations and impulses that are triggered by the local context of decision, rather than by a more global evaluation, which—as dictated by a normative analysis—would be guided by foresight and planning and by considerations of compounding and consistency.

At the core of this account lies a classic case of divergence in disciplinary sensitivities. Whereas the notion of "construal" is now second nature to social psychologists, it is far less obvious to common intuition, and even less so in economics. Economic thinking concerns itself with states of the world. The axioms that underlie the normative treatment are stated in terms of extensional outcomes, where each option, representing a state of the world, is clearly and uniquely designated, independently of how precisely it may be described or perceived. The problem with this perspective is a painfully obvious but profoundly consequential fact about human nature: People's decisions are not directly about objective (extensional) states of the world but about the mental (intensional) representations of those states. And the mental representation of states of the world is subject to a rich psychology, involving embellishment, interpretation, and distortion, much of it arising at the time of having to make a decision. As a consequence, trivial contextual features can have an unwarranted impact, and nuanced differences in otherwise logically equivalent representations of options can generate widely different perceptions, trigger different reactions, and yield different preferences.

The implications of this tension are profound. Although people tend to endorse the normative desiderata upon reflection, their decision behaviors, driven as they are by intuitive evaluation, automatic processes, and locally triggered impulses, often violate the normative requirements in systematic and predictable ways (Shafir & LeBoeuf, 2002). We have recently conducted a field experiment in South Africa to assess the relative importance of subtle psychological features compared to price in the decision to take a loan from a local lender (Bertrand, Karlan, Mullainathan, Shafir, & Zinman, 2005). Some 57,000 incumbent clients of the lender were sent letters offering large, short-term loans at randomly chosen interest rates. Consistent with standard economics, those offered higher rates were less likely to take a

loan than those with access to lower rates. In addition, various "psychological" features on the offer letter, which did not affect offer terms or economic content, were also independently randomized. One of the features consisted of the number of sample possible loans displayed: The offer letter displayed either one example of a loan size and term, along with the respective monthly repayment information, or multiple (four) such examples of loan sizes and terms. In contrast with standard economic prediction, and in line with the foregoing discussion regarding decisional conflict and the proliferation of alternatives, we found higher takeup of the loan under the one-example description than under the multiple-example version. The magnitude of this effect was large: The more limited description of the offer had the same positive effect on takeup as dropping the monthly interest on these loans by 2 percentage points. (In a similar vein, Huberman et al., 2004, find that employees' participation in 401(k) plans drops as the number of fund options proposed by their employer increases.) We also randomized the presence or absence of a smiling woman's picture in the bottom corner of the offer letters. For the men in the sample, the presence of that picture had the same positive effect on takeup as dropping the monthly interest on the loans by 4.5 percentage points! On average, any one psychological manipulation we tried had the same effect as a one-half-percentage-point change in the monthly interest rate on loans that averaged one-third of the borrowers' monthly income.

People's decisions are influenced by a variety of factors. Some factors, such as a concern with fairness, the tendency to defer difficult decisions, and even attitudes toward gains and losses, may be, at least in part, conscious and intentional. Others, such as those based on automatic processing, happen entirely out of awareness. Most people who are not well versed in current behavioral research, for example, would most likely be unwilling to entertain the possibility that a touch on the shoulder or a picture in the corner of a letter predispose them to make substantially different economic decisions. Choice behavior is the outcome of a variety of cognitive and affective processes. Some of these—as well as others that are not true about us—we consciously and proudly endorse; others we have no introspective access to or control over. What is common to many of these is that they are creatures of the moment. They arise from a local context and are employed in the construction of preferences we need to form. But local contexts change, and with them the cues and features that impinge upon our valuation of alternatives. As a result of this malleability, our decisions often violate the most basic consistency requirements, including the requirements of independence of irrelevant alternatives, and of independence from irrelevant descriptive or procedural nuances. Also, because these often diverge, the context-sensitive character of our decision making tends to call into question the presumed continuity between the contexts of decision and those surrounding consumption. Normative treatments tend to focus less on specific decisions and more on the aggregate outcomes they produce in combination. Much

of our mental apparatus, on the other hand, both conscious and not, focuses heavily on each decision in turn, with limited attention devoted to aggregate outcomes. Our tendency to construct decisions in context, quirky, often sophisticated, but highly localized, produces some excellent constructions as well as others that are not so good. Typically, however, these occur in the absence of an overall plan, leading to patterns we would have wished to design otherwise had we been able to think about it differently.

NOTES

1. Of course, as time progressed, my commute became habitual and was performed by rote. This, in fact, captures another key feature of many decisions. They are first made under the influence of local context and later become thoughtless and "automatic."
2. Of course, the horizon that ought optimally to be taken in making such decisions can be a matter of some debate and may depend on the options under consideration. Classic economic theories of saving often assume maximization over one's life cycle. For many other decisions—such as how best to plan our Christmas vacation—shorter horizons are sufficient. This issue does not have an impact on the present discussion.

REFERENCES

Ainslie, G. (2001). *Breakdown of will.* New York: Cambridge University Press.

American Lung Association. (2003). What are the benefits of quitting smoking? Available: *www.lungusa.org/tobacco/quit_ben.html*

Avnet, T., & Higgins, E. T. (2006). How regulatory fit impacts value in consumer choices and opinions. *Journal of Marketing Research*, *43*(1), 1–10.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and Stereotype activation on action. *Journal of Personality and Social Psychology*, *71*(2), 230–244.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology 75*, 19–32.

Bazerman, M. H., Schroth, H. A., Shah, P. P., Diekmann, K. A., & Tenbrunsel, A. E. (1994). The inconsistent role of comparison others and procedural justice in reactions to hypothetical job descriptions: Implications for job acceptance decisions. *Organizational Behavior and Human Decision Processes*, *60*, 326–352.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., & Zinman, J. (2005). *What's psychology worth?: A field experiment in the consumer credit market.* Unpublished manuscript, University of Chicago.

Bettman, J. R., & Sujan, M. (1987). Effects of framing on evaluation of comparable and non-comparable alternatives by expert and novice consumers. *Journal of Consumer Research*, *14*, 141–154.

Bornstein, R. F. (1989). Exposure and effect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*, 366–381.

Caruso, E.M., & Shafir, E. (2005). Now that I think about it, I'm in the mood for laughs: Decisions focused on mood. *Journal of Behavioral Decision Making*, *18*, 1–15.

Chapman, K. J. (2001). Measuring intent: There's nothing "mere" about mere measurement effects. *Psychology and Marketing*, *18*, 811–841.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Crusco, A. H., & Wetzel, C. G. (1984). The Midas touch: The effects of interpersonal touch on restaurant tipping. *Personality and Social Psychology Bulletin*, *10*, 512–517.

Downs, J. S., & Shafir, E. (1999). Why some are perceived as more confident and more insecure, more reckless and more cautious, more trusting and more suspicious, than others: Enriched and impoverished options in social judgment. *Psychonomic Bulletin and Review*, *6*, 598–610.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, *30*(4), 510–517.

Elster, J. (2000). *Ulysses unbound: Studies in rationality, precommitment, and constraints.* New York: Cambridge University Press.

Elster, J., & Loewenstein, G. (Eds.). (1992). *Choice over time.* New York: Russell Sage.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research*, *16*, 280–288.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*, 1–17.

Fischhoff, B., Slovic, P., Lichtenstein, S., Reid, S., & Coombs, B. (1978). How safe is safe enough?: A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, *9*, 127–152.

Fitts, P. M., & Seeger, C. M. (1953). S–R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*, 199–210.

Fitzsimons, G. J., & Shiv, B. (2001). Nonconscious and contaminative effects of hypothetical questions on subsequent decision making. *Journal of Consumer Research*, *28*, 224–238.

Forehand, M. R., Deshpandé, R., & Reed, A., II. (2002). Identity salience and the influence of differential activation of the social self-schema on advertising response. *Journal of Applied Psychology*, *87*(6), 1086–1099.

Frederick, S., Loewenstein, G., & Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*, 351–401.

Frederick, S., & Shafir, E. (2005). *Principled choices and mindless matches: Tales about tradeoffs.* Unpublished manuscript, Massachusetts Institute of Technology.

Garbarino, E. C., & Edell, J. A. (1997). Cognitive effort, affect, and choice. *Journal of Consumer Research*, *24*(2), 147–158.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*, 617–638.

Goldstein, W. M., & Hogarth, R. M. (Eds.). (1997). *Research on judgment and decision making: Currents, connections, and controversies.* Cambridge, UK: Cambridge University Press.

Greenwald, A. G., Carnot, C. G., Beach, R., & Young, B. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology*, *72*, 315–318.

Grether, D., & Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, *69*, 623–638.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319–340.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*, 1140–1153.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*(2), 141–154.

Holyoak, K., & Morrison, R. (Eds.). (2005). *Cambridge handbook of thinking and reasoning*. New York: Cambridge University Press.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation of preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, *67*, 247–257.

Hsee, C. K., & Kunreuther, H. C. (2000). The affection effect in insurance decisions. *Journal of Risk and Uncertainty*, *20*, 141–159.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *5*, 576–590.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*, 90–98.

Huberman, G., Iyengar, S. S., & Jiang, W. (2004). *Defined contribution pension plans: Determinants of participation and contribution rates*. Manuscript under review.

Isen, A. M., & Geva, N. (1987). The influence of positive affect on acceptable level of risk: The person with a large canoe has a large worry. *Organizational Behavior and Human Decision Processes*, *39*, 145–154.

Isen, A. M., Nygren, T. E., & Ashby, F. G. (1988). Influence of positive affect on the subjective utility of gains and losses: It is just not worth the risk. *Journal of Personality and Social Psychology*, *55*, 710–717.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*, 995–1006.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*, 1338–1339.

Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7, 35–51.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*, 20–31.

Kahneman, D. (1992). Reference points, anchors, and mixed feelings. *Organizational Behavior and Human Decision Processes*, *51*, 296–312.

Kahneman, D. (1994). New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, *150*, 18–36.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: The foundations of hedonic psychology*. New York: Russell Sage.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*, 401–405.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economics*, *98*, 1325–1348.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*, 17–31.

Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty*, *9*, 5–38.

Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, *5*, 187–200.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341–350.

Kahneman, D., & Tversky, A. (1995). Conflict resolution: A cognitive perspective. In K. J. Arrow, R. H. Mnookin, L. Ross, A. Tversky, & R. B. Wilson (Eds.), *Barriers to conflict resolution* (pp. 45–60). New York: Norton.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge, UK: Cambridge University Press.

Kelley, H. H. (1955). Salience of membership and resistance to change of group-centered attitudes. *Human Relations*, *8*, 275–289.

Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, *79*, 1277–1284.

Koehler, D., & Harvey, N. (Eds.). (2004). *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell.

Kruglanski, A. W., Shah, J., Fischbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 331–376). New York: Academic Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "seizing" and "freezing." *Psychological Review*, *103*(2), 263–283.

LeBoeuf, R. A. (2002). *Alternating selves and conflicting choices: Identity salience and preference inconsistency*. Unpublished doctoral dissertation, Princeton University.

LeBoeuf, R., & Shafir, E. (2003). Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making*, *16*, 77–92.

LeBoeuf, R. A., & Shafir, E. (2005a). *Alternating selves and conflicting choices: Identity salience and preference inconsistency*. Manuscript under review.

LeBoeuf, R., & Shafir, E. (2005b). Decision making. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 243–265). New York: Cambridge University Press.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*, 146–159.

Levin, I. P. (1987). Associative effects of information framing. *Bulletin of the Psychonomic Society*, *25*, 85–86.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*, 149–188.

Levin, I. P., Schnittjer, S. K., & Thee, S. L. (1988). Information framing effects in social and personal decisions. *Journal of Experimental Social Psychology*, *24*, 520–529.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 551–578.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46–55.

Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preferences in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, *101*, 16–20.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*, 272–292.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics*, *107*, 573–597.

Loewenstein, G., & Thaler, R. H. (1989). Intertemporal choice. *Journal of Economic Perspectives*, *3*, 181–193.

Loewenstein, G., Weber, E., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Mandel, N., & Johnson, E. J. (2002). When Web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, *29*, 235–245.

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, *51*(4), 858–866.

Mather, M., Shafir, E., & Johnson, M. K. (2000). Misremembrance of options past: Source monitoring and choice. *Psychological Science, 11,* 132–138.

McNeil, B. J., Pauker, S. G., Sox, H. C., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine, 306,* 1259–1262.

Morwitz, V. G., Johnson, E. J., & Schmittlein, D. (1993). Does measuring intent change behavior? *Journal of Consumer Research, 20,* 46–61.

Neumann, R., & Strack, F. (2000). "Approach and avoidance": The influence of proprioceptive and exteroceptive encoding of affective information. *Journal of Personality Social Psychology, 79,* 39–48.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43,* 87–131.

Payne, J.W., Bettman, J. R., & Schkade, D. A. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty, 19,* 243–270.

Quattrone, G. A., & Tversky, A. (1988). Contrasting rational and psychological analyses of political choice. *American Political Science Review, 82,* 719–736.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes, 79*(1), 56–77.

Raghunathan, R., & Trope, Y. (2002). Walking the tightrope between feeling good and being accurate: Mood as a resource in processing persuasive messages. *Journal of Personality and Social Psychology, 83,* 510–525.

Redelmeier, D. A., & Shafir, E. (1995). Medical decision making in situations that offer multiple alternatives. *Journal of the American Medical Association, 273,* 302–305.

Redelmeier, D., Shafir, E., & Aujla, P. (2001). The beguiling pursuit of more information. *Medical Decision Making, 21,* 376–381.

Redelmeier, D. A., & Tversky, A. (1990). Discrepancy between medical decisions for individual patients and for groups. *New England Journal of Medicine, 322,* 1162–1164.

Redelmeier, D. A., & Tversky, A. (1992). On the framing of multiple prospects. *Psychological Science, 3,* 191–193.

Reed, A., II. (2004). Activating the self-importance of consumer selves: Exploring identity salience effects on judgments. *Journal of Consumer Research, 31*(2), 286–295.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science, 12,* 185–190.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1,* 7–59.

Sanna, L., Schwarz, N., & Small, E. M. (2002). Accessibility experiences and the hindsight bias: I knew it all along versus it could never have happened. *Memory and Cognition, 30,* 1288–1296.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Schelling, T. (1980). The intimate contest for self-command. *Public Interest, 60,* 94–118.

Schelling, T. (1984). Self-command in practice, in policy, and in theory of rational choice. *American Economic Review, 74,* 1–11.

Schkade, D. A., & Johnson, E. J. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes, 44,* 203–231.

Schkade, D. A., & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science, 9,* 340–346.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 527–561). New York: Guilford Press.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review, 2*(2), 87–99.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195–202.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513–523.

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How goal pursuit is affected by accessible alternatives. *Journal of Experimental Social Psychology, 38,* 368–382.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory and Cognition, 21,* 546–556.

Shafir, E. (1994). Uncertainty and the difficulty of thinking through disjunctions. *Cognition, 50,* 403–430.

Shafir, E. (1995). Compatibility in cognition and decision. In J. R. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Decision making from the perspective of cognitive psychology* (Vol. 32, pp. 247–274). New York: Academic Press.

Shafir, E. (2002). Cognition, intuition, and policy guidelines. In R. Gowda & J. C. Fox (Eds.), *Judgments, decisions, and public policy* (pp. 71–88). New York: Cambridge University Press.

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53,* 491–517.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49,* 11–36.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology, 24*(4), 449–474.

Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), *An invitation to cognitive science: Vol. 3. Thinking* (2nd ed., pp. 77–100). Cambridge, MA: MIT Press.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology, 39,* 211–221.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80–83.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research, 26*(3), 278–292.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158–174.

Simonson, I. (1990). The effect of purchase quantity and timing on variety seeking behavior. *Journal of Marketing Research, 27,* 150–162.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 289–295.

Slovic, P. (1995). The construction of preference. *American Psychologist, 50,* 364–371.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397–420). New York: Cambridge University Press.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1988). Response mode, framing, and information-processing effects in risk assessment. In D. E. Bell, H. Raiffa, & A. Tversky (Eds.), *Decision making: Descriptive, normative, and prescriptive interactions* (pp. 152–166). Cambridge, UK: Cambridge University Press.

Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. In R. Hogarth (Ed.), *Insights in decision making: Theory and applications* (pp. 5–27). Chicago: University of Chicago Press.

Slovic, P,, & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review, 73,* 596–605.

Sunstein, C. R., Kahneman, D., Schkade, D., & Ritov, I. (2001). *Pre-*

dictably incoherent judgments. Chicago: University of Chicago Press.

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters*, *8*, 201–207.

Trope, Y., Ferguson, M., & Raghunathan, R. (2000). Mood-as-a-resource in processing self-evident information. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 256–272). Mahwah, NJ: Erlbaum.

Trope, Y., Igou, E. R., & Burke, C. (2006). *Mood as resource in structuring goal pursuit.* Unpublished manuscript, New York University.

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology*, *79*(6), 876–889.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*(3), 403–421.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 77–121). Greenwich, CT: JAI Press.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and psychology of choice. *Science*, *211*, 453–458.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, *59*, 251–278.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, *106*, 1039–1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384.

Tversky, A., & Shafir, E. (1992a). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, *3*, 358–361.

Tversky, A., & Shafir, E. (1992b). The disjunction effect in choice under uncertainty. *Psychological Science*, *3*, 305–309.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*, 1178–1189.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *American Economic Review*, *80*, 204–217.

Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology*, *79*, 66–76.

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin*, *29*, 1159–1168.

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: Effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology*, *82*, 434–447.

Wedell, D. H. (1997). Another look at reasons for choosing and rejecting. *Memory and Cognition*, *25*, 873–887.

Wickens, C. D. (1984). *Engineering psychology and human performance.* Columbus, OH: Merrill.

Wilson, T. D., Dunn, D. S., Krafit, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental and social psychology* (pp. 123–205). San Diego, CA: Academic Press.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*, 181–192.

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*, 821–836.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs*, *9*(2), 1–27.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151–175.

Zajonc, R. B., & Markus, H. (1982). Affective and cognitive factors in preferences. *Journal of Consumer Research*, *9*(2), 123–131.

Psychological Distance

Nira Liberman
Yaacov Trope
Elena Stephan

WHAT IS PSYCHOLOGICAL DISTANCE?

Psychologically distant things (objects, events) are those that are not present in the direct experience of reality. We would like to set aside the philosophical questions regarding the objective existence of reality and focus, instead, on people's subjective experiences of reality, divergence from such experiences, and the consequences of those. People believe that they directly experience themselves and their immediate surroundings at the present moment. Anything that is not present is distal. It may be thought of, constructed, or reconstructed, but it cannot be experienced directly. There are different reasons for things not to be present in the immediate reality experienced by me. Things may belong to the past or to the future (e.g., my first year of marriage, my first year of school), to spatially remote locations (e.g., my parents' house, the North Pole), to other people (the way my best friend or a person from another culture perceives the present situation), and to hypothetical alternatives to reality, what could or might have been but never materialized (e.g., had I married another person or had I had wings). These alternatives to the directly experienced reality define, respectively, four dimensions of psychological distances—temporal distance, spatial distance, social distance, and hypotheticality. In each pair of examples of distal things, the first example is more proximal than the second. We would like to propose that in relation to psychological distance, these various distance dimensions are anchored on a single starting point (zero distance

point), which is my direct experience of the here and now. Anything else—other times, other places, experiences of other people, and hypothetical alternatives to reality—is a mental construct.

This analysis suggests a basic relationship between psychological distance and construal; that is, any distancing (i.e., moving beyond direct experience) involves construal. Based on construal-level theory (CLT; Liberman & Trope, 1998; Trope & Liberman, 2003), we distinguish between extents (levels) of construal and propose that more distal entities, which are more remote from direct experience, are construed on a higher level (i.e., involve more construal). The reason for this change in construal could be that as we move away from direct experience of things, we have less information about those things. Thus, we know less about the more distant future and the more distant past, we know less about more remote places, we know less about acquaintances than about close friends, and, finally, we know less how the world would have been with more remote (i.e., less probable and more difficult to imagine) alternatives to reality. In all those cases, lack of knowledge about the more remote entities—people, events, places, alternatives—requires representing them more abstractly than proximal entities.

Sometimes, in the absence of knowledge about distal entities, one may extrapolate from proximal ones. For example, if I do not know how my friend felt about a movie, I may reason that she is similar to me and assume that her feelings were similar to mine. Ordinarily, simi-

larity decreases with distance (e.g., the distant future becomes increasingly less similar to the present and more distal people are increasingly less similar to ourselves) and with it the ability to project (i.e., extrapolate) direct experience. Schematic construals are employed instead. For example, if I assume that the experiences of a person from another culture are dissimilar to mine, then I need to resort to general schemas in order to construe those experiences. More generally, in the absence of any specific knowledge about either a distal or a proximal entity, the former would be less likely construed in concrete terms than the latter.

We thus propose that because there is less information about distal entities than proximal entities, people typically construe the former more abstractly than the latter. As a result, an association is formed between psychological distance and level of construal, and the use of more abstract construals for more distal entities may be generalized beyond the conditions that initially gave rise to the association. That is, people may construe distal entities at a high level and proximal entities at a low level even when information on these entities is equally available or obtainable. Differential knowledge explains how the association between distance and level of construal was formed, but it does not necessarily explain every specific case of construing distal entities at a high level.

It is important to note that estimations of distance are subjective. What determines level of construal is not the physical distance as measured in meters or temporal distance as measured in seconds but, rather, the subjective experience of distances as large or small. The same objective distance may look differently due to psychological factors, some of which are "cognitive" (e.g., division into more stages may enhance perceived distance), and some are "motivational" (e.g., in the service of self-enhancement, one's negative behaviors may seem more distant than one's positive behaviors). This chapter does not address the question of how objective units of distance translate into subjective experiences. Rather, we address the effects of perceived distance on other outcomes, such as level of construal, prediction, affect, and choice.

PLAN OF THE CHAPTER

The next section discusses in more detail the concept of level of construal and the association between level of construal and psychological distance. That section addresses two implications of this association, namely, that psychological distance would produce higher levels of construal and that, conversely, high levels of construal would enhance perceived distance. The third section examines the effects of psychological distance on confidence in prediction, intensity of affective reactions, and evaluation and choice. We present evidence suggesting that the effects of various distance dimensions are similar to each other and are mediated by level of construal. The fourth section further proposes that the different psychological distances are interrelated and to some extent interchangeable. That is, distancing an object on one dimension may be exchanged for distancing the object on another dimension.

PSYCHOLOGICAL DISTANCE AND LEVEL OF CONSTRUAL

What Is Level of Construal?

Before examining how psychological distance and level of construal affect each other, we discuss in more detail the concept of level of construal. Our basic premise is that high-level construals change the meaning of events by forming more abstract representations of the events. A defining property of concrete representations is that they lend themselves to multiple abstract representations. For example, the concrete action "waving a hand" could be identified more abstractly as threatening or, alternatively, as being friendly (Vallacher & Wegner, 1987), and "a poodle" could be classified as a pet or, alternatively, as a mammal (Rosch & Lloyd, 1978). An abstract representation is selected among different possible abstractions according to context-defined relevance, which, in turn, may be affected by one's goals. Thus, if one's goal is to buy a poodle for a child, then "a pet" is a relevant category and friendliness is relevant, but if one's goal is scientific zoological classification, then "a mammal" is a relevant category and friendliness is of less relevance.

Usually, some aspects of the focal stimulus are more closely related to one interpretation or categorization (the hand movement was relatively fast; poodles are friendly), while other aspects are more closely related to other interpretations or categorizations (the person who waved her hand seemed to smile; poodles are warm blooded). Moving from a concrete representation to an abstract representation involves deciding on one of the alternative abstract representations. This, in turn, means omitting the features that are perceived to be less important while retaining those considered more central or important to the abstract construct in question. For example, in replacing "waving a hand" with the more abstract construal "showing friendliness," the fact that one used one's hand is omitted (Semin & Fiedler, 1988). Similarly, in representing a poodle as a pet, warm-bloodedness is omitted (Rosch & Lloyd, 1978). Like irrelevant details, details that are inconsistent with the chosen abstract representation are omitted from the representation or assimilated into it. For example, the detail that the hand waving was slightly faster than usual would be omitted or modified once the "being friendly" interpretation is chosen.

Because abstract representations necessarily impose one of possibly many alternative interpretations, and because irrelevant or inconsistent details are omitted from the abstract representation or assimilated to it, abstract representations may be expected to be simpler, less ambiguous, and more prototypical than concrete representations (Fiske & Taylor, 1991; Smith, 1998). Abstraction thus involves moving to a more schematic, simple, and coherent representation.

There are multiple levels of abstractness, as one could construct less inclusive or more inclusive categories of objects (e.g., poodle, dog, or mammal). In feature-based theories of categorization, more inclusive categories have fewer features and therefore are simpler than concrete categories (Rosch & Lloyd, 1978). Actions also form hierarchies of abstractness (e.g., giving money, helping, and being a good person) with each level of abstractness containing less concrete details about the specific type of action performed and the objects it involved (Semin & Fiedler, 1988; Trope, 1989). In the same way, traits form hierarchies (e.g., an excellent guitarist, musical, and talented) such that more abstract traits are less detailed about the behaviors, objects, circumstances, and people they involve (Hampson, John, & Goldberg, 1986). Goal-directed actions form hierarchies too, as goals could be translated into more abstract, superordinate goals (Carver & Scheier, 2000; Miller, Galanter, & Pribram, 1960; Vallacher & Wegner, 1987). In such hierarchies, each action (e.g., study for an exam) has a superordinate, abstract level, which answers the question of why the action is performed (e.g., do well) and a subordinate, concrete level, which supplies the details of how the action is to be performed (e.g., read a textbook).

It is important to note that abstract representations are not simply more impoverished than concrete representations. Rather, many times, they contain additional information (i.e., high-level meaning) about the value of the stimulus and its relations to other stimuli. For example, the construal "a mammal" entails many characteristics not directly observed in the original "poodle" stimulus (e.g., ways of reproduction), and it places it within a larger context of living things, thus specifying its relations to other animal species (e.g., it informs us that a poodle is closer to mice than to fish). Thus, the process of abstraction involves not only a loss of specific, idiosyncratic, and incidental information but also ascription of new meaning via the top-down process of induction from stored knowledge. Abstract representations contain less information about the unique instance but might, nevertheless, contain more general information.

In sum, we suggest, in line with other theories of mental construal, that abstract mental models represent the detailed and possibly ambiguous information contained in a real event in a relatively simple and coherent mental model that connects it to stored knowledge. Moreover, it is possible to think about levels of abstractness as a gradual reduction in incidental details and in complexity of representations. We further suggest that higher-level, more abstract and coherent representations are formed of more distal entities.

The rest of this section presents evidence in support of the construal-level hypothesis, which is that distancing produces more abstract, higher-level construals. We also examine the reverse direction of influence, namely, higher-level construals leading to perception of increased psychological distance. To that end, we examine new and existing social psychological literature on future and past temporal distance, physical distance, social distance, and hypotheticality.

The Effect of Psychological Distance on Level of Construal

The Effect of Temporal Distance on Level of Construal

CONSTRUAL OF ACTIONS

Research conducted in the framework of temporal construal theory has examined the effects of future temporal distance on level of construal of actions, objects, other people, and oneself. For example, in one study participants imagined themselves engaging in various activities (e.g., reading a science fiction book) either tomorrow or next year and described these activities (Liberman & Trope, 1998, study 1, part 1). The study found that people used more superordinate, high-level descriptions of distant activities (e.g., "getting entertained") and low-level descriptions for near future activities (e.g., "flipping pages"). A related, forced-choice study used an adapted version of Vallacher and Wegner's (1989) Level of Personal Agency questionnaire that was originally designed to assess stable individual differences in action identification (Liberman & Trope, 1998, study 1, part 2). The questionnaire presents 19 activities (e.g., "locking a door"), each followed by two restatements, one corresponding to the how (low-level) aspects of the activity (e.g., "putting a key in the lock") and the other corresponding to the why (high-level) aspects of the activity (e.g., "securing the house"). As predicted, participants chose significantly more high level, why restatements when the activities were described as occurring in the distant future than when the same activities were described as occurring in the near future.

CATEGORY BREADTH

If the distant future is represented more abstractly, then individuals should use broader categories to classify objects for distant-future situations than for near-future situations. To test this prediction, Liberman, Sagristano, and Trope (2002, study 1) asked participants to imagine an event (e.g., camping trip) in either the upcoming weekend or a weekend a few months later and to classify a given set of 38 objects related to the event (e.g., tent and toothbrush) into as many mutually exclusive and exhaustive groups as they deemed appropriate. Participants were instructed to assume that their classification was final and could not be altered later. The authors simply counted the number of groups into which participants classified the objects from each scenario. The results showed that participants used fewer categories when they imagined the objects in a distant-future scenario than a near-future scenario. This finding is consistent with the CLT assumption that distant-future events are represented in terms of relatively high-level, abstract categories, whereas near-future events are represented in terms of relatively low-level, concrete categories.

PRINCIPLES AND VALUES

Eyal, Liberman, Sagristano, and Trope (2006) found that the distant future is construed in terms of abstract

principles more than the near future. For example, in one study, participants read about distant-future and near-future situations that involved an abstract principle or a dilemma (e.g., "In a few days/in a few years, the university will set to increase the number of minority students") and asked them to choose a description of this situation either in terms of a global principles or in terms of a lower-level action, devoid of moral implications (e.g., "endorsing affirmative action" vs. "making changes to admission lists"). It was found that distant-future situations were perceived in terms of high-level principles more than near-future situations. Presumably, principles more easily apply to the distant future, but as the situation gets closer in time, morals and ideologies seem to lose their relevance.

CAUSAL ATTRIBUTION

A series of studies by Nussbaum, Trope, and Liberman (2003) examined level of construal of social targets as a function of future temporal distance. A considerable amount of person perception research has demonstrated the correspondence bias, namely, the tendency to attribute situationally constrained behavior to the corresponding personal disposition (see Gilbert & Mallone, 1995; Jones, 1979). In terms of CLT, this bias reflects a high-level construal of behavior in terms of abstract, decontextualized dispositions (see Fiedler, Semin, Finkenauer, & Berkel, 1995; Semin & Fiedler, 1988; Semin & Smith, 1999). According to CLT, the correspondence bias would be more likely with more distal social targets.

Nussbaum and colleagues (2003, study 1) used the Jones and Harris (1967) attitude attribution paradigm to test this hypothesis with respect to temporal distance. Student participants from Tel-Aviv University read an essay arguing in favor of Israel's withdrawal from Lebanon. (The study was conducted a few months before Israel's withdrawal from Lebanon in June 2000.) They were told that the essay was written by a student who had been instructed either to express her own opinion (unconstrained condition) or to argue in favor of withdrawal (situationally constrained condition). Participants were asked to estimate the likelihood that the writer would express pro-withdrawal attitudes in a variety of near-future (next day) or distant-future (a year later) situations (e.g., express pro-withdrawal attitudes in a conversation with friends and attend a pro-withdrawal rally). The results showed that the judged likelihoods of essay-consistent (pro-withdrawal) behavior in the near future were moderated in view of the situational constraints, whereas the judged likelihoods for the more distant future were high regardless of situational constraints. Thus, whereas near-future predictions showed substantial situational discounting, distant-future predictions showed little or no situational discounting. These findings demonstrate that the correspondence bias, the tendency to underweight low-level situational constraints on observed behavior, is more pronounced when this behavior is used for predicting the distant future than the near future.

BEHAVIOR CONSISTENCY

CLT predicts that people would expect others to behave more consistently across different situations in the distant future than in the near future. Nussbaum and colleagues (2003, study 2) tested this hypothesis by asking participants to predict an acquaintance's behavior in four different situations (e.g., a birthday party and waiting in line in the supermarket) in either the near future or the distant future. Participants predicted the extent to which their acquaintances would display 15 traits (e.g., behave in a friendly vs. unfriendly manner) representative of the Big Five personality dimensions (extraversion, agreeableness, conscientiousness, emotional stability, intellect). Cross-situational consistency was assessed by computing, for each of the 15 traits, the variance in predicted behavior across the four situations and the correlations among the predicted behaviors in the four situations. As hypothesized, the results showed that participants expected others to behave more consistently across distant-future situations than across near-future situations. This was manifested in both lower cross-situational variance and higher cross-situational correlations for distant-future behavior predictions than for near-future behavior predictions.

CONSTRUAL OF THE SELF

The latter study was replicated with the self as the target. Nussbaum, Liberman, and Trope (2004) made people imagine themselves in different situations, either in the near future or in the distant future, and indicate the extent to which their behavior in those situations would reflect each of the Big Five traits. As with describing other people, it was found that in the distant future people expected themselves to exhibit traits more consistently across situations. This study suggests, in line with CLT, that people are more likely to use abstract, decontextualized trait concepts in predicting distant-future than near-future behaviors of both other people and themselves.

Nussbaum and colleagues (2004) also examined the effect of future temporal distance on the level of construal of self using Linville's (1985) self-complexity paradigm. Participants described themselves to another person, who was expected to read the description in either the near future or the distant future. For that end, they chose cards with self-descriptive adjectives and grouped them into categories. The results indicated that participants used fewer categories and that the categories were more distinct from each other in the distant-future condition than in the near-future condition. It seems, then, that a temporally distal perspective on the self fosters a less complex representation thereof than a proximal temporal perspective.

In sum, extensive research conducted within the framework of temporal construal theory demonstrates that future temporal distance enhances level of construal of actions, situations, objects, and people. Are there similar effects for other dimensions of psycho-

logical distance? We turn now to examine this question.

The Effect of Past Temporal Distance on Level of Construal

Closely related to CLT is the idea that past temporal distance is associated with higher construal levels. Semin and Smith (1999) asked participants to recall significant events from either the distant past (at least a year ago) or the recent past (the last 2 weeks) and analyzed the abstractness level of the responses. They found more abstract language in descriptions of more distant events. The authors suggested that past events are gradually moved into a long-term storage system, which uses an abstract, schematic form of information representation. Consistent with this view is also the research on temporal effects on attribution, which has shown that over time, observers' attributions about an actor's behaviors become more dispositional and less situational, presumably because the specific details of the situations fade away more rapidly than the more general inferences about the actor's personality (Frank & Gilovich, 1989; see also Funder & Van Ness, 1983; Moore, Sherrod, Liu, & Underwood, 1979; Nigro & Neisser, 1983; Peterson, 1980; but see Burger, 1986).

These findings are consistent with memory consolidation, namely, the tendency for concrete details to fade away from memory more rapidly than general abstractions, which makes memories of the distant past more abstract than memories of the recent past (Bartlett, 1932; Wyer & Srull, 1986). However, assuming that memory is an active, reconstructive process, CLT proposes that increasing past temporal distance is associated not only with a loss of specific details but also with an increased use of high-level, schematic construals (Ross, 1989). For example, Ross (1989) found that young adults tend to think that their personal adjustment (i.e., their sense of personal worth and self-reliance) and social skills improve with age. In an actual assessment, however, these scores appeared relatively stable, with the same average self-ratings between the ages of 25 and 45 (Woodruff & Birren, 1972). Interestingly, at age 45 people tended to rate their adjustment at the age of 25 as much lower than it actually was, presumably because they subscribed to a general theory that adjustment and social skills increase over the years (Ross, 1989). Thus, it is not only the case that information about their actual standing at 25 was lost, but it was also replaced with schema-derived inferences. Consistent with this idea, McDonald and Hirt (1997) showed that over time, student participants not only forgot a fellow student's grades but also were more likely to infer those from overall expectancies and general attitudes toward that person, thus producing a more coherent picture of the target person over time. Similarly, Mitchell, Thompson, Peterson, and Cronk (1997) found that people's recollections of a bicycle trip or a trip to Europe became more positive over time, presumably as the details of these events faded from memory and were replaced with a more general, schematic (and, in

this case, positive) representation. It seems, then, that the passage of time produces not only a passive loss of detail but also an active reconstruction characteristic of higher-level construals.

The Effect of Spatial Distance on Level of Construal

Spatial distance from an object ordinarily determines the amount and kind of information that is available about the object. As one gets closer to an object, the information becomes more accurate and detailed. From the distance we see a forest; as we get closer we begin seeing the trees. A more pertinent question, however, is whether, as would be predicted by CLT, the same information about an object would be represented differently depending on whether the object is thought to be in a close or a distant location. A study by Fujita, Henderson, Eng, Trope, and Liberman (2006) examined how perceivers construe a social interaction believed to take place in either a spatially close or a distant location. New York University student participants watched and then described a video of two other NYU students casually conversing about their schooling, hobbies, and families. They were told that the video was filmed either at NYU Study Abroad in Florence, Italy (distant location), or at NYU in New York City (proximal location). Content analysis based on Semin and Fiedler's (1988) linguistic categorization model (LCM) served as to assess the level of abstractness of the descriptions. As expected, the descriptions of the taped interaction were more abstract when the interaction was said to take place in Florence than in New York City. Participants thus formed more abstract representations of the same observed social interaction when its location was said to be distant rather than near. It should be pointed out that the effect of the location was unrelated to participants' sense of familiarity with or similarity to the interactants, which was also measured in this study. These findings demonstrate the idea that spatial distance, like temporal distance, leads to higher-level construal of the information at hand.

The Effect of Social Distance on Level of Construal

The distinctions between self and other, similar and dissimilar others, familiar and unfamiliar others, ingroup members and outgroup members, and status differences, all may be considered as instances of social distance. We now review social psychological research documenting higher-level construals of more distant social targets. We then discuss the case of politeness as an indicator of social distance and examine its effects on level of construal. Finally, we examine the possibility of relating cultural differences in construal to CLT.

SELF VERSUS OTHER

Perhaps the best documented difference in construal between self and other is the actor–observer effect in attribution (Jones & Nisbett, 1972; for a review, see Gilbert, 1998). This research shows that a person's view of his or

her behavior emphasizes the role of concrete situational factors that operate at the moment of action, whereas his or her view of other people emphasizes the causal role of stable, general dispositional properties of the actor. For example, Robins, Spranca, and Mendelson (1996) found that participants in a "getting acquainted" conversation task tended to see their own behavior as caused by their partner but their partner's behavior as caused by his or her personality (see also Nisbett, Caputo, Legant, & Marecek, 1973; Saulnier & Perlman, 1981; for a review, see Watson, 1982).

Similar to what CLT suggests, Semin and Fiedler (1989) argued that the actor–observer effect reflects different levels of abstraction of action representation. In support of this proposal, Semin and Fiedler (1989; see also Fiedler et al., 1995) made participants describe either their own or another person's behaviors in a number of situations (e.g., a successful party or a failure at school) and coded their responses for abstractness. The analysis employed the LCM (Semin & Fiedler, 1988) described earlier. The results showed that observers' descriptions of behaviors had a higher proportion of abstract verbs than actors' descriptions. The reverse pattern was found for the rate of specific context supplements (e.g., space and time). From the perspective of CLT, these findings demonstrate a higher construal level of distal social targets (another person) than a proximal social target (the self).

Self–other differences might be explained as being due to differences in knowledge (people know more about themselves and the variability of their behavior over situations than about others) and differences in the salience of behaviors versus situations (the latter is more salient from one's own perspective, the former from the observer's perspective). The differential knowledge explanation is consistent with the CLT notion of how high level of construal came to be associated with distant entities, but it does not necessarily exemplify the notion of construal level. Would differences in construal emerge when knowledge is identical for near and distal social targets? Research relating abstractness of memories to the perspective in which they are recalled seems to offer a positive answer. It has been shown, for example, that personal memories of behaviors that were recalled from a third-person perspective (e.g., "try to remember your first day at school, as if you are now watching the kid you were") rather than from a first-person perspective ("try to remember your first day at school, as if you are a kid again") tended to employ dispositional (as opposed to situational) terms (Frank & Gilovich, 1989; Nigro & Neisser, 1983). In a similar vein, Libby and Eibach (2002, study 4) found that imagining performing an activity (e.g., rock climbing and playing drums) from a third-person perspective produced less vivid and rich reports of the activity than imagining the same activity in a first-person perspective. In terms of CLT, this means that a third-person perspective, which imposes more distance than a first-person perspective, induced a higher level of construal.

A third-person perspective on oneself may be induced by a situational manipulation of self-awareness, such as the presence of a mirror or a camera (Duval & Wicklund, 1972; Wicklund, 1975) but may also express a dispositional tendency for public self-consciousness (Fenigstein, Scheir, & Buss, 1975). Both sources of self-awareness seem to increase people's tendency to behave in a way that corresponds to their own personal values and socially accepted ideals (e.g., Macrae, Bodenhausen, & Milne, 1998; for a review, see Gibbons, 1990). From our perspective, values and ideals are high-level constructs (Eyal, 2005). The finding that self-awareness increases behavior–value correspondence is therefore consistent with the idea that increased distance from oneself increases the impact of high-level constructs on one's behavior. Self-awareness also decreases the effect of situational factors on behavior. For example, self-aware individuals are better able to resist situational inducements to commit moral transgressions under the cover of anonymity (Beaman, Klentz, Diener, & Svanum, 1979). Moreover, both situationally manipulated self-awareness and dispositional self-consciousness seem to enhance dispositional attributions and diminish situational attributions for one's own actions, presumably because these variables engender a view on oneself that is akin to that of an external observer (Gibbons, 1990). As noted before, general dispositions are typically more abstract than situational attributions and, therefore, constitute a higher level of construal.

In sum, people tend to construe others in higher-level terms than themselves and, in addition, to construe themselves in higher-level terms when taking the perspective of another person. Both of these effects are consistent with CLT. The latter effect is particularly informative because it is hard to interpret in terms of differential knowledge about the target.

INGROUPS VERSUS OUTGROUPS

A considerable amount of research on group perception suggests that people form higher-level construals of outgroups than their ingroups. Compared with ingroups, outgroups are described in more abstract terms (Fiedler et al., 1995; Werkman, Wigboldus, & Semin, 1999) and are perceived as more homogenous (Jones, Wood, & Quattrone, 1981; Park & Judd, 1990; Park & Rothbart, 1982), less differentiated into subgroups (Brewer & Lui, 1984; Linville, 1982; Park, Ryan, & Judd, 1992), and as possessing more structured, predictable sets of properties (Linville, Fischer, & Yoon, 1996). From the perspective of CLT, we construe outgroups more abstractly than ingroups because we typically have less direct experience with outgroups and thus perceive them as more distant. We might also have less information about the outgroup. We propose, however, that a generalized sense of greater psychological distance from the outgroup than the ingroup may independently affect construal. Support for this proposal comes from the Jones and colleagues (1981) study on ingroup heterogeneity/outgroup homogeneity. These researchers asked students, who were members of four selective clubs, to characterize members of different clubs (their own club and each of the three other groups) on a number of per-

sonal characteristics, such as introverted–extraverted, refined–crude, and athletic–nonathletic. On each trait scale, participants marked the position of the average member of each group, two values between which they believed 50% of the group members would fall, and two scale values between which they believed all group members would fall. As predicted, participants indicated wider ranges for their own club than the three other clubs, indicating more perceived variability within the ingroup as opposed to the outgroup. Importantly, this tendency was unrelated to the number of ingroup or outgroup members known, suggesting that the more abstract construals of the outgroup than the ingroup may be independent of amount of knowledge about those groups.

SOCIAL POWER

Almost any relationship can be characterized in terms of the amount of control an individual has over others' outcomes relative to the amount of control others' have over the individual's outcomes (Thibaut & Kelley, 1959). A considerable amount of social psychological research has shown that social power affects a broad range of social cognitive and self-regulatory phenomena (Fiske, 1993; Keltner, Gruenfeld, & Anderson, 2003). Does social power also affect social distance among individuals? Based on the definition of power in terms of outcome dependence, Smith and Trope (2006) recently proposed that the more powerful individuals feel more independent of others and, therefore, more distinct and separate from them. Indeed, there is evidence that individuals with more power see themselves as more different from, and thus distant from, other people than people with less power (e.g., Lee & Tiedens, 2001; Snyder & Fromkin, 1980). For example, group leaders tend to be distinctive relative to their fellow group members, and over time they become even more psychologically separate from the rest of the group (see Hogg, 2001; Hogg & Reid, 2001).

If power entails more social distance, then CLT would predict that it would predispose the more powerful individuals to form high-level construals of relevant situations. Powerful individuals might, as a result, focus on the central, most important aspects of the situation and disregard secondary, more peripheral aspects, thus forming an unequivocal orientation toward the situation. In a study related to this prediction, Overbeck and Park (2001) asked high- and low-power participants to interact via email with targets who held the opposite power role. Some of the information participants received was relevant to the task at hand and some was irrelevant. At the end of these interactions, participants were asked to list as much information from the emails as they could remember. Not only did participants in the high-power role recall overall more information from these interactions, but they were especially superior at recalling information that was relevant. Thus, high-power participants were better at distinguishing between primary and secondary information, a hallmark of abstract processing.

Powerful people also seem to use more abstract language. Guinote (2001) found that Portuguese participants used more abstract language to describe both their ethnic group and an outgroup when they were part of the majority (i.e., they were living in Portugal: higher-power group) than when they were part of a minority (i.e., they were immigrants living in Germany: lower-power group). Similarly, participants who played the role of judges used more abstract, trait-like language in referring to themselves than did participants who were workers (Guinote, Judd, & Brauer, 2002).

Smith and Trope (2006) examined how priming power affects participants' ability to abstract visual stimuli in perceptual tasks. For example, one of the studies examined participants' performance on an embedded-figures task in which a specified pattern has to be found and traced within a complex geometrical figure (Witkin, Oltman, Raskin, & Karp, 1971). As predicted, participants who were primed with high power performed better than participants who were primed with low power. Subsequent studies by Smith and Trope found that power-primed participants were (1) better at detecting patterns and thus more accurately detected the amount of covariation in a series of data; (2) focused more on primary features, making more superordinate categorizations; and (3) more accurate at detecting structure, making responses that better fit the data at hand, in a Gestalt completion task. These effects of power on level of construal were not mediated by participants' reported mood, efficacy, or effort.

It has been argued that powerful individuals are susceptible to heuristic thinking and uninhibited reliance on superficial stereotype-consistent cues (Fiske, 1993; Keltner et al., 2003). The present analysis suggests an interesting alternative to this view. Power does not seem to reduce overall attention to and utilization of the available information. Instead, it seems to focus individuals on its central, high-level aspects. It is possible, then, that the distal perspective activated by the possession of social power promotes going beyond the information given, detecting the underlying structure, and abstracting from it superordinate, central features. This power-driven construal may support the kind of long-term planning and goal pursuit that are often required of individuals in positions of power.

POLITENESS

The way individuals communicate with each other might be indicative of the social distance between them (Argyle, 1970). One general aspect of interpersonal communication is the extent to which the speaker addresses a recipient politely. Research conducted across different cultures has found that politeness is closely related to social distance. That is, speakers tend to address more politely socially distant listeners than socially close listeners (see Brown & Levinson, 1987; Holtgraves, 2002). Politeness may thus serve as an indicator of social distance and, like social distance, act to produce higher construal levels of the situation.

A series of studies by Stephan (2005) tested this hypothesis. For example, one study examined how the intention to be very polite as opposed to less polite affects the abstractness of one's language. Participants were asked to indicate how a student would address a classmate (e.g., ask for his or her notes) in a very polite way, a moderately polite way, or a less polite way. The open-ended responses were analyzed for abstractness according to the Semin and Fiedler's (1988) LCM. As predicted, abstract verbs were used more frequently and concrete verbs were used less frequently with increasing levels of politeness. Another study examined how the intention to be very polite versus less polite affected level of construal of actions. Participants imagined a person performing an action and then described it in either a polite way or a less polite way by choosing between a high-level, "why" description and a low-level, "how" description of the action (Vallacher & Wegner, 1989). For example, participants chose between describing "Shelly is reading a book" as "Shelly is gaining knowledge" or as "Shelly is following lines of print." As predicted, high-level restatements were chosen more frequently for polite descriptions than for less polite descriptions. Both studies demonstrate, then, that an increase in level of politeness resulted in higher-level construals of actions.

CULTURE

Research on cross-cultural psychology distinguishes between independent and interdependent cultures. This research suggests that Asian and South American cultures, compared to West European and North American cultures, emphasize more the interdependence and interconnectedness of the individual with the collective, rather than his or her independence from others (Berry, 1976; Markus & Kitayama, 1994). Thus, participants from China, India, Africa, and Mexico often explicitly acknowledge the importance of other people, of relations, and of the participatory, responsive, interpersonal nature of behavior. In terms of social distance, interdependent cultures may be characterized as maintaining more proximity between a person and his or her social surrounding (Markus & Kitayama, 1994). If others are more proximal in interdependent cultures, then CLT would predict that in these cultures, people would be predisposed to using low-level construals of persons, objects, and events.

Indeed, a large amount of research has documented a reduced tendency in interdependent cultures to construe others in high-level, dispositional terms, and an enhanced tendency to construe them in lower-level, situation-specific and relational terms (Miller, 1984; Morris & Peng, 1994; for a review, see Choi, Nisbett, & Norenzayan, 1999). For example, Morris and Peng (1994) found that Chinese participants explain a murder crime in situational terms (e.g., the murdered was just fired from work) whereas European Americans explain a similar crime in dispositional terms (e.g., the murderer could not control himself). Interestingly, it seems that the differences in construal between independent and interdependent cultures apply not only to people, but also to a wide range of nonsocial stimuli. Nisbett, Peng, Choi, and Norenzayan (2001) suggest, more broadly, that individuals from interdependent cultures process stimuli in a contextual manner, attend to the relationships between the focal object and the field, explain and predict events on the basis of such relationships, and rely on experience-based knowledge rather than on abstract logic. In contrast, individuals from independent cultures presumably process the same stimuli as isolated from the immediate context, focus on attributes of the object and assign it to categories, use rules about the categories to explain and predict the object's behavior, and use formal logic. For example, these researchers report a study in which Japanese and American participants were shown pictures of fish and other underwater objects and reported what they had seen. This study found that Japanese participants made more statements about background elements than did American participants. Interestingly, Nisbett and colleagues suggested that these differences in perception and cognitive styles may be traced to structural differences in the socioeconomic systems that have evolved in different cultures. Social structure presumably shapes cognition because perception of social objects is generalized to the inanimate world, and because reasoning develops in the service of solving interpersonal conflicts. Based on this assumption, Nisbett and colleagues argue more specifically that social interdependence and close social ties give rise to attention to relations between objects and to experience-based knowledge, whereas social independence gives rise to isolating objects, attending to their attributes, and assigning them to abstract categories. This idea is consistent with the CLT proposal that social distance affects level of construal.

The Effect of Level of Construal on Psychological Distance

As indicated earlier, psychological distance may not only affect construal but may be affected by construal. We argued that the association of distance to high construal levels becomes generalized beyond situations in which distance entails less knowledge. We further suggest that through this association process, the connection also became bidirectional, so that high levels of construal induce perceptions of greater distance from stimuli. We therefore predict that highly construed stimuli would be perceived as more distant in time and space, as more distant socially, and as less real.

The Effects of Level of Construal on Future Temporal Distance

Liberman, Trope, McCrae, and Sherman (in press) examined the effect of construal level on the temporal distance of activity enactment. In one of their studies, participants were first asked to indicate either "why" (i.e., high-level construal) or "how" (i.e., low-level construal) a person would perform an activity (e.g., "Ron is considering opening a bank account. Why (How) would Ron do that?") and were then asked to estimate how much time

from now the person would do the activity. As predicted, participants indicated more distant enactment times after a high-level "why" construal than after a low-level "how" construal. The authors found similar effects with other manipulations of level of construal, and with participants' estimates of the enactment time of their own activities.

The Effects of Level of Construal on Past Temporal Distance

Semin and Smith (1999, studies 2 and 3) studied the effect of linguistic abstractness on event age. They provided participants with retrieval cues of varying abstractness and examined how distant were the events that they recalled. For example, participants recalled either an occasion on which they helped somebody (i.e., concrete retrieval cue) or an occasion on which they displayed a trait of helpfulness (i.e., abstract retrieval cue). As predicted, an abstract retrieval cue prompted memories that were more than 8 months older than memories that were prompted by a concrete retrieval cue. Thus, abstractness of representation affected past temporal distance.

Semin and Smith (1999) suggest that abstract terms mark distant events and concrete terms mark recent events. They relate their findings to the existence of two separate learning and memory systems: a fast learning system and a slow learning system (McClelland, McNaughton, & O'Reilly, 1995). According to this model, the slow learning (schematic) system records repeatedly encountered regularities in the environment and uses them to fill in unobserved details and interpret new information. It uses, therefore, more abstract representations. In contrast, the fast learning (episodic) memory system uses more concrete representations and records specific events together with information about their context. From our perspective these findings demonstrate a more general principle, which is that high-level construals foster a perception of more distant past events.

The Effects of Level of Construal on Social Distance

Stephan (2005) conducted a series of studies to examine the effect of level of construal on politeness and familiarity, which were conceptualized as indicators of social distance. It was predicted that an increase in level of construal would produce a corresponding increase in perceived politeness and a decrease in perceived familiarity of a social target. For example, one study asked participants to provide either dispositional, high-level explanations or situational, low-level explanations for an actor's behavior (e.g., "Danny is explaining the class materials to another student. What aspects of Danny's personality, character, or dispositions [in the situation or the setting] could explain Danny's behavior?"). The researchers then measured the perceived familiarity of the target person. Consistent with the prediction, politeness and familiarity were higher after participants generated low-level, situational attributions than high-level, dispositional attributions. These results suggest that an increase

in level of construal produces an increase in perceived social distance, as indicated by reduction in perceived familiarity as well as in increased politeness.

The Effects of Level of Construal on Hypotheticality and Probability

A low-level construal of hypothetical events, more than a high-level construal, makes them seem more likely to become real, or, in other words, makes them seem more probable (for reviews, see Koehler, 1991; Nisbett, 1993). For example, a study by Sherman, Chialdini, Schwartzman, and Reynolds (1985) presented participants with information about a disease that supposedly was becoming prevalent on campus and asked them to imagine actually contracting the disease. For some subjects, symptoms were concrete (low energy level, muscle aches, severe headaches), whereas for others the symptoms were described more abstractly (disorientation, malfunctioning nervous system). A control group read the descriptions of the disease but did not engage in imagining contracting the disease. The results indicated that subjects who imagined contracting the disease with the concrete symptoms estimated that the likelihood of actually contracting it was greater than the control subjects imagined, who, in turn, gave higher likelihood estimates than the group of subjects who imagined abstract symptoms. Thus, construing something on a lower level makes it seem more likely. Koehler (1991) argued that low-level, detailed construals allow a more confirmatory information search. It is possible, however, that the mere existence of low-level, concrete details creates a feeling of greater reality, veridicality, and likelihood.

The literature on source monitoring in memory (Johnson, Hashtroudi, & Lindsay, 1993) is also closely related to the connection between level of construal and veridicality. One of the questions that this literature has addressed is how people know if a mental representation corresponds to reality rather than to fantasy. In other words, how do people decide that something they remember actually happened as opposed to being imagined, dreamt, or considered (e.g., did I really visit my aunt's house or did I only consider going there?). According to source monitoring theory, certain memory characteristics enhance people's confidence in an event's reality. These include perceptual detail, such as sound, smell, touch, and taste, as well as contextual information, such as the hour, day, year and season the event took place, and the relative spatial arrangement of people and of objects (Johnson, Foley, Suengas, & Raye, 1988). For example, I might think I remember the color of the floor my aunt stood on and the smell of the cooking meal in her kitchen and thus conclude that it must be the case that I really visited her house. In terms of CLT, these details pertain to a low-level construal and the process of inferring the reality of a mental representation involves using the existence of low-level details as an indication of higher likelihood and veridicality.

It is noteworthy that representations of imagined actions do actually appear to have less perceptual and contextual features than observed events (Suengas & John-

son, 1988). Stern and Rotello (2000) made participants perform some actions and imagine others (e.g., eating crackers and tying a ribbon around a pencil) and examined the memory characteristics of these actions both immediately and 1 week later. They found that performed events were clearer and richer in sensory and contextual detail than imagined events. Thus, it appears that the relation between hypotheticality and level of construal is bidirectional: not only that people infer realism (i.e., probability) from level of construal but also that realism affects level of construal. Interestingly, in the study by Stern and Rotello, the level of perceptual and contextual features of both imagined and performed events deteriorated over a period of 1 week (see Johnson et al., 1988, for a comparable effect of time on autobiographical memories), making the level of these features similar between immediate imagined actions and performed actions 1 week later. It seems, therefore, that temporal distance has an effect on level of construal that is similar to the effect of hypotheticality.

Implicit Associations between Psychological Distance and Level of Construal

Recently, Bar-Anan, Liberman, and Trope (in press) examined associations between level of construal and psychological distance using an Implicit Associations Test. Similar to other assessments of implicit associations (e.g., between stereotyped group members and stereotypical attributes; Greenwald, McGhee, & Schwartz, 1998), participants in these studies were presented with stimuli from four categories: stimuli pertaining to high-level construal (e.g., category names such as "drinks"), stimuli pertaining to low-level construal (e.g., exemplar names such as "coke"), stimuli pertaining to low psychological distance (e.g., the word "ours" or the word "friend" for the social distance), stimuli pertaining to high psychological distance (e.g., the word "theirs" or the word "stranger"). In the critical trials participants mapped stimuli from these four categories on two responses, pressing either a left key or a right key on the computer keyboard. On CLT-congruent trials, high-level stimuli were paired with distant stimuli and low-level stimuli were paired with proximal stimuli, whereas on CLT-incongruent trials high-level stimuli were paired with proximal stimuli and low-level stimuli were paired with distal stimuli. Reaction times were compared between congruent and incongruent trials to test the prediction of CLT that reaction would be faster on congruent than on incongruent trials. Each study examined one of four dimensions of psychological distance—temporal distance, spatial distance, social distance, and hypotheticality. With all four dimensions it was found that participants are faster with congruent than with incongruent pairings, suggesting that participants implicitly associate psychological distance with high-level construal and psychological proximity with low-level construal.

This set of studies extends previous lines of research on the association between level of construal and psychological distance in several important respects. First, it shows that this association cannot be explained by differential knowledge about proximal versus distal targets. Second, the research demonstrates similar effects across the four dimensions of psychological distance (temporal, spatial, social, and hypotheticality), thus suggesting that they all share a common meaning as instances of psychological distance. Third, it suggests that the association between psychological distance and construal level can be activated automatically without conscious deliberation.

COGNITIVE, AFFECTIVE, AND BEHAVIORAL EFFECTS OF PSYCHOLOGICAL DISTANCE

How does psychological distance from an event affect people's affect, cognitions, and behaviors toward those events? Research conducted in the framework of temporal construal theory addressed this question with respect to future time perspective. This section reviews the research, integrates it with literatures suggestive of similar effects of other distance dimensions, and offers new predictions for these dimensions. As in the previous section, we examine the effects of future temporal distance, past temporal distance, hypotheticality, and social distance. We also examine the effects of level of construal. As demonstrated in the previous section, level of construal is closely related to psychological distance, and therefore CLT postulates that its effects on prediction, evaluation, and choice would resemble those of psychological distancing.

Prediction

The French anthropologist Claude Levi Strauss once noted that in the Western sciences, the most distal topics (e.g., astronomy and geography) developed first whereas those that look at more proximal entities (e.g., psychology) were last to develop (Levi Strauss, 1978). We would like to propose, consistent with this observation, that abstract theoretical reasoning is easier to apply to distal targets than to proximal targets. More specifically, CLT proposes that increasing psychological distance from a future situation would make it more likely that predictions about this situation would be based on the implications of high-level rather than low-level construals. In theory-based predictions, the theory constitutes a high-level construct that promotes confident predictions, whereas low-level features, associated with incidental deviations and noise, undermine confidence. In these cases, psychological distancing, because it enhances the effects of high-level constructs and reduces the effects of low-level constructs, would enhance confidence. For example, economic theory posits that increasing interest rates causes the stock market to decline. The theory acknowledges that other factors might also affect the stock market but treats them as noise. According to CLT, psychological distancing would promote confidence in predicting that if interest rates are raised, then the stock market would fall. Thus, economists would be more confident in their prediction when considering the more distant future, when making predictions about more geographically distant markets, when making forecasts

about the investments of other people, and when the predictions concern a hypothetical or unlikely scenario.

Normatively, predictions about more distant entities should be made with less confidence because less is known about them. For example, one knows less about another person than about oneself, and hence one should be less confident when predicting the other's than one's own behavior. However, if high-level construals promote greater confidence, then people may feel no less and even more confident in predicting distal outcomes. For example, because stable personality traits and dispositions are more readily inferred about another person than about oneself (Jones & Nisbett, 1972), one may be more confident in predicting another person's correspondent behaviors than one's own (Pronin, Kruger, Savtisky, & Ross, 2001). This logic of CLT is consistent with the claim that overconfident predictions stem from relying on oversimplified representations of situations.

A great deal of social cognitive research has identified mental construal as underlying a wide range of prediction errors, including overconfidence (Dunning, Griffin, Milojkovic, & Ross, 1990; Griffin, Dunning, & Ross, 1990), the planning fallacy (Buehler, Griffin, & Ross, 1994; Kahneman & Lovallo, 1991; Kahneman & Tversky, 1979), affective forecasting errors (Dunn, Wilson, & Gilbert, 2003; Gilbert, Morewedge, Risen, & Wilson, 2004; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Gilbert & Wilson, 2000; Kahneman & Snell, 1990, 1992; Wilson, Meyers, & Gilbert, 2001). This research has shown that people often have unwarranted confidence in their prediction because their mental models of future situations are schematic and oversimplified. For example, people typically underestimate task completion times because they base their estimates on schemas of how things unfold and fail to take into consideration nonschematic events (Buehler et al., 1994). Thus, the schema of writing a paper includes reading relevant literature, analyzing data, and communicating with colleagues, but it does not include a visit from one's in-laws. Therefore, the latter, even if known in advance, would not be taken into account in predicting how much time it takes to write a paper. As another example, focalism involves overconfident prediction of one's own reactions to emotional events (e.g., how sad I am going to be if my team loses the game) due to an underestimation of the diluting effect of low-level, situational factors (i.e., underestimation of the fact that most of life remains unaffected by the game; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000).

In all these lines of research, predictions are compared to actual outcomes (i.e., predicted completion times are compared to actual completion times), whereas CLT focuses on comparing predictions for proximal versus distal entities (e.g., estimations of completion times in the distant future vs. the near future). This difference notwithstanding, we would like to note that explaining the discrepancy between prediction and reality as a consequence of construal is in line with our assumption that distances are anchored on one's direct experience and that any distancing entails construal. Let us now turn to

examine the effects of psychological distance on prediction, which, we believe, are mediated by differences in construal in much the same way as differences between predicted and actual outcomes.

Future Temporal Distance and Confidence in Prediction

Direct evidence for these hypotheses comes from research on the effect of temporal distance on predictions of social and nonsocial events (Nussbaum et al., 2004; see also Nussbaum et al., 2003). One study examined the confidence of advanced psychology students in replicating classic findings in psychology in either the near future or the distant future (Nussbaum et al., 2004, study 1). For example, participants imagined entering a class at the university, either the next day or a year later (depending on the experimental condition), handing the students a list of words to memorize and then testing how well they remember it after moving some of the students to a different room. Participants estimated how likely it is that those tested in the same room would outperform, on average, those who were moved to a different room, thus replicating the encoding specificity effect. Participants were more confident in replicating encoding specificity when they imagined conducting the experiment in the distant future than in the near future. The same pattern of results occurred also with other classic findings in social, cognitive, and developmental psychology.

Several other studies assessed confidence in predicting one's own performance on a general knowledge quiz expected to take place either on the same day or 2 months later (Nussbaum et al., 2004, study 3). These studies used the same questions but in either a relatively easy or hard question format, which we assumed is a low-level aspect of performance. Specifically, in one study, the quiz consisted of either multiple-choice questions (relatively easy format) or open-ended questions (relatively hard format). In another study, the quiz consisted of questions with either two response alternatives (relatively easy) or four response alternatives (relatively hard). We also assessed participants' perceived ability in each knowledge domain (how knowledgeable one is in geography, history, etc.). The results showed that the difficult question format appropriately reduced confidence in near-future performance but failed to reduce confidence in distant-future performance. We think that this was the case because question format was a low-level aspect of the situation that, consistent with CLT, affected confidence in near-future outcomes more than in distant-future outcomes. Our results also indicated that participants' beliefs about their general knowledge in different domains predicted their confidence in that domain in the distant future better than in the near future. We think that this was the case because such beliefs constitute a high-level consideration, as they are perceived to be more central to the quiz.

The Nussbaum and colleagues (2003) studies on the effect of future temporal distance on dispositional attribution, discussed in the first section of this chapter as an example of the effect of future temporal distance on level of construal, are also relevant to prediction. They show

that people base their predictions of others' more distant-future behavior more on high-level, dispositional attributions and less on low-level, situational attributions.

In sum, there seems to be considerable empirical support for the idea that temporal distance increases the impact of high-level information (e.g., theories, prototypes, self-beliefs, and personal dispositions) and decreases the impact of low-level information (e.g., irregular outcomes and specific situational and task characteristics) on prediction. Thus, two complementary processes may contribute to the unwarrantedly higher confidence levels associated with distant-future predictions: underweighting of the uncertainty associated with low-level information and overweighting of the certainty associated with high-level information.

Confidence Regarding Past Events

How does confidence about past events change over time? Bearing on this question is research on the hindsight bias, namely, people's tendency to exaggerate their past estimated likelihoods of an event after it had occurred (Fischhoff, 1975; for a review, see Christensen-Szalanski, & Willham, 1991). For example, after January 1, 2000, people reported that before that date they estimated the likelihood of no disaster happening as quite high, higher than the estimates they actually gave before that date (Pease, McCabe, Brannon, & Tagler, 2003). Interestingly, one of the explanations for the hindsight bias is akin to construal. Specifically, it has been proposed that people construct a theory to explain the outcome (e.g., 9/11 events were foreseeable given the frustration of the Muslim world), which makes the outcome seem inevitable.

CLT predicts that the hindsight bias and, more generally, theory-driven confidence in the inevitability of past events would increase over time. The research literature provides some support for this prediction. For example, Bryant and Guilbault (2002) found that hindsight about President Bill Clinton's acquittal in the Monica Lewinsky case increased from 4 days after the verdict to 11 days after the verdict. Obviously, more research is needed to examine in more detail the hypothesis that hindsight would increase over time and to specify the conditions under which this would occur. For example, it would be interesting to examine whether confidence in the inevitability of historical events increases over large time spans (e.g., how inevitable were the events of September 11, 2001, or how inevitable was World War II) and whether an increase in confidence is associated with holding general theories that explain those events. It would also be interesting to manipulate the salience of local, low-level theories (e.g., 9/11 happened because the FBI failed to arrest the suicide pilots) as opposed to global, high-level theories (the Muslim world felt outraged toward the United States), and examine whether the salience of the latter, more than the former, would make the effect of hindsight increase over temporal distance. Such results, if obtained, would suggest that past events that are explained in global and abstract terms seem increasingly inevitable over the course of time, more than past events that are attributed to specific, low-level events. In other words, the more global one's theories, the more the distant past would seem inevitable and the distant future would seem foreseeable.

Summary: Psychological Distance and Confidence

We reviewed research on the effect of future and past temporal distance on confidence in prediction. This research demonstrates that distancing enhances confidence in predictions that are based on high-level features, such as theories and schemas, but not confidence in predictions that are not based on high-level constructs. Would similar effects obtain also for social distance and for hypothetical versus actual scenarios? We could not find any research that directly addresses this question and could only speculate about such effects. A possible prediction would be, for example, that remote hypothetical models (e.g., a model of the effects of a meteorite hitting earth) would make more confident and clear-cut predictions than models of more realistic and likely events (e.g., a model of the effects of global warming). One could also predict that as a precondition of a theoretical prediction becomes more probable (e.g., as the likelihood of a meteorite hitting earth would increase), confidence in one's predictions would decline.

Preferences

How do people evaluate and make choices about distant future outcomes, as opposed to near future outcomes? How is advice given to others different from one's own choice? How does uncertainty about a situation change the way it is evaluated? And does level of construal change decisions? All these are questions we now address. As before, the most direct evidence bearing on these questions comes from research on the effects of future temporal distance, but findings related to other distances are discussed as well.

Future Temporal Perspective and Preferences

How the value of outcomes changes with temporal distancing has been a question of central importance in psychology (e.g., Ainslie, 1975; Ainslie & Haslam, 1992; Baumeister & Heatherton, 1996; Metcalfe & Mischel, 1999; Rachlin, Brown, & Cross, 2000; Read & Loewenstein, 2000), behavioral economics (e.g., O'Donoghue & Rabin, 2000), and political science (e.g., Elster, 1977; Schelling, 1984). All these behavioral sciences have generally assumed that the value of outcomes is discounted or diminished as temporal distance from the outcomes increases. Indeed, a considerable amount of research suggests that individuals often place higher value on a near-future reward than on a distant-future reward, even when the distant-future reward is larger (e.g., Ainslie & Haslam, 1992; Elster & Loewenstein, 1992; Mischel, Grusec, & Masters, 1969; Mischel, Shoda & Rodriguez, 1989; Read & Loewenstein, 2000).

Contrary to the claim of overall time discounting, CLT proposes that the effect of temporal distance would de-

pend on the level of construal with which value is associated. In this view, temporal changes in the attractiveness of an option depend on the value associated with the high-level construal of the option (high-level value) and the value associated with the low-level construal of the option (low-level value). Temporal distance should increase the weight of high-level value and decrease the weight of low-level value. As a result, temporal distance should shift the overall attractiveness of an option closer to its high-level value than to its low-level value. When the low-level value of an option is more positive than its high-level value, the option should be more attractive in the near future (time discounting). However, when the high-level value of an option is more positive, the option should be more attractive in the distant future (time augmentation).

We examined this hypothesis with different manipulations of high versus low levels of construal: primary, goal-related versus secondary, goal-irrelevant sources of value; feasibility versus desirability and expectancy versus value in gambles; arguments in favor versus arguments against an action; and abstract and primary attitudes and values versus concrete and secondary attitudes and values. This literature was reviewed elsewhere (Liberman & Trope, 1998; Sagristano, Trope, & Liberman, 2002; Trope & Liberman, 2000; see Trope & Liberman, 2003), and is only briefly summarized here. We return to the general question of discounting after we review the effects of different psychological distances on evaluation and choice.

PRIMARY VERSUS SECONDARY ASPECTS OF OBJECTS

Consider an activity consisting of two parts: a main task, which is the goal of the activity, and an unrelated filler task to be performed during a break in the main task. Because the main task is the primary goal of the activity, it is part of a high-level construal of the activity, and because the filler task is a secondary aspect of the activity, it is part of a low-level construal of the activity. CLT therefore predicts that temporal distance will increase the weight of the value of the main task relative to the weight of the value of the filler task in determining the overall attractiveness of the activity. When the main task is more attractive than the filler task, the overall activity would become more attractive over temporal distance. In contrast, when the main task is less attractive than the filler, the overall activity would become less attractive over temporal distance.

Trope and Liberman (2000, study 4) presented participants with activities consisting of either an interesting main task and a boring filler or a boring main task and an interesting filler. Each activity was described as consisting of three sessions of performing the main task, with the filler task performed between these sessions to provide rest and distraction from the main task. For example, an activity titled "Judging Humor" was described as follows: "The main task is judging humor, and will ask you to evaluate the funniness of cartoons. The filler task in between the three sessions is checking data, and will ask you to compare two lists of numbers to check for dis-

crepancies." When the chosen activity was said to take place in a few weeks, participants strongly preferred the activity with an interesting main task to the activity with a boring main task. However, when the chosen activity was said to take place in the same experimental session, this preference was significantly weaker. Thus, as predicted by CLT, temporal distance enhanced the tendency to evaluate activities in terms of goal-relevant rather than incidental aspects, so that with time delay the activity with an interesting main task (but boring filler) became more attractive and the activity with a boring main task (but interesting filler) became less attractive. The same temporal changes in preference were found for evaluations of products with primary and secondary features (Trope & Liberman, 2000, study 3).

FEASIBILITY AND DESIRABILITY

An important difference between high-level and low-level construals of goal-directed action is their emphasis on the desirability versus feasibility of outcomes. Desirability refers to the value of an action's end state, whereas feasibility refers to the ease or difficulty of reaching the end state. For example, desirability concerns the value of receiving a job offer, whereas feasibility concerns the amount of time and effort one has to invest to get the job offer. Given this assumption, CLT predicts that desirability considerations are more likely to guide distant-future preferences, whereas feasibility considerations are more likely to guide near-future preferences.

Liberman and Trope (1998) tested these predictions in a number of studies. One of the studies (study 4) used a realistic choice situation. Tel Aviv University students taking an introductory social psychology course were presented with a choice among several course assignments. The assignments were either easy (based on readings in Hebrew, the students' native language) or difficult (based on readings in English, a foreign language for these students) and either on an interesting topic (e.g., romantic love) or on an uninteresting topic (e.g., history of social psychology). In this situation, the difficulty of the assignment represents a feasibility consideration and the interest level of the assignment represents a desirability consideration. Students had to submit both a near-future and a distant-future assignment. They were told that they would have 1 week to work on each assignment but that the near-future assignment (reading materials and essay questions) would be given immediately whereas the distant-future assignment would be given 9 weeks later. Consistent with CLT, students' preferences showed that time delay decreased the effect of the difficulty of the assignments and increased the effect of the interest level of the topic of the assignments. The preference for the easy but uninteresting assignment decreased over time, whereas the preference for the hard but interesting assignment increased over time. Thus, in selecting a near-future assignment, students were willing to sacrifice interest for the sake of ease. In contrast, in selecting a distant-future assignment, students were willing to sacrifice ease for the sake of interest, thus committing themselves to a desirable but less feasible task. A similar tem-

poral pattern was obtained with various other options (Liberman & Trope, 1998, study 2).

Feedback seeking is another important decision that often pits feasibility against desirability concerns. Freitas, Salovey, and Liberman (2001) reasoned that feedback seeking involves a conflict between the goal of gaining information about oneself (a desirability consideration) and the difficulty of being exposed to self-evaluation (a feasibility consideration). They therefore predicted and actually found that distant future feedback preferences depended on the accuracy of the offered feedback, whereas near-future feedback preferences depended on the evaluative implications of the feedback. Informative but unflattering feedback was preferred for the distant future, whereas uninformative but flattering feedback was preferred for the near future.

An interesting implication of CLT's view on feasibility and desirability concerns the effect of temporal distance on planning. Liberman and Trope (1998, study 5) conceptualized time constraints as a feasibility aspect of an activity and investigated the role of time constraints and desirability of activities in near- and distant-future planning. They showed that plans for the distant future tend to reflect desirability of activities and disregard time constraints, thus creating a tendency to overcommit. It appears that in making distant-future plans, individuals consider each activity in isolation and fail to take into account that each activity they plan comes at the expense of some other activities in which they may want to engage at the same time.

PROBABILITY AND VALUE

The distinction between feasibility and desirability may be extended to games of chance—gambles characterized by probability of winning and the monetary payoff associated with winning. According to CLT, payoff is the superordinate consideration because the payoff determines the desirability of the end state of a gamble. The probability of winning is a subordinate consideration having to do with the properties of the random mechanism, device, or procedure that determines the feasibility of winning. In the normative expected utility model, probability and payoffs combine multiplicatively and therefore have symmetric weight in determining the attractiveness of gambles. However, our studies have demonstrated that people view the probability of winning as subordinated to the payoff; that is, they think that probability is important only if the payoff is high, but that payoff is important regardless of whether the probability of winning is high or low (Sagristano et al., 2002, study 1). This establishes payoffs as pertaining to a higher construal level than probabilities, and entails a prediction by CLT that people would assign more weight to payoffs and less weight to probabilities in deciding for the more distant future.

A series of studies on preference for near- and distant-future gambles tested this prediction (Sagristano et al., 2002). For example, one of the studies assessed monetary bids for gambles to be played on the same day or 2 months later. Participants were presented with a set of 20 bets that varied in probability of winning and expected value and were asked to state the amount of money they were willing to bid to play each gamble. As expected, preference among near-future gambles was primarily based on probability of winning, whereas preference among distant-future gambles was primarily based on the payoffs associated with winning. Thus, for near-future gambles, bids were highest for high-probability–low-payoff bets (i.e., relatively safe bets), whereas for distant-future gambles, bids were highest for low-probability–high-payoff bets (i.e., risky bets). These findings extend CLT to uncontrollable, random outcomes.

ARGUMENTS IN FAVOR OF AND AGAINST AN ACTION

In deciding whether to undertake an action, cons are subordinate to pros. This is because the subjective importance of cons depends on whether or not pros are present more than the subjective importance of pros depends on whether or not cons are present. For example, consider a decision to undergo a medical treatment. If we know that the treatment has some health benefit for us, we would inquire about its potential side effects before making a decision. But if the treatment has no benefits for us, we would decide against taking it without further inquiry about its side effects. In contrast, we would inquire whether a medical treatment has health benefits whether or not it has side effects—when the treatment is known to have no side effect, information about its benefits may tell us whether the treatment is worth taking; when the treatment is known to have some side effects, we may still inquire about the benefits of the treatment in order to determine whether they outweigh its side effects. Thus, the importance of side effects depends on whether the treatment is known to have benefits, but the importance of benefits is independent of whether the treatment is known to have side effects.

After establishing these subordination relations in a series of studies (Eyal, Liberman, Trope, & Walther, 2004, studies 1a and 1b), the authors proceeded to examine an obvious implication that follows from CLT: If cons are subordinate to pros, then pros should become more salient as temporal distance from the action increases, whereas cons should become less salient as temporal distance from the action increases. A series of studies tested this prediction by asking participants to generate arguments in favor and against new (i.e., nonroutine) near-future or distant-future actions. As predicted, participants generated relatively more pro arguments and fewer con arguments when the actions were to take place in the more distant future. The proposed action involved new exam procedures (e.g., switching to open-ended questions instead of multiple-choice questions; study 2), social policies (e.g., restricting private cars in the city center; study 3), and a variety of personal and interpersonal behaviors (e.g., approaching a fellow student and offering to write an assignment together; studies 4–6). In all the studies, participants generated more pros and less cons as temporal distance from the actions increased.

PREDICTING BEHAVIORAL INTENTIONS FROM ATTITUDES AND VALUES

Personal attitudes and values are commonly viewed as transsituational guides (Schwartz & Bilsky, 1987), as abstract structures that provide continuity and meaning under changing environmental circumstances (Feather, 1995), and as stable meaning-producing superordinate cognitive structure (Rohan, 2000). Based on CLT, we propose that attitudes and values, because of their relatively abstract and decontextualized nature, will be more readily applied to and guide choice in psychologically distant situations. As one gets closer to a situation, choices are increasingly more likely to be based on secondary, low-level considerations.

Initial evidence for this analysis was recently obtained by Sagristano, Eyal, Trope, and Liberman (2006). The first session of the study assessed participants' general attitudes toward blood donation, volunteering for psychology experiments, and physical fitness advisement. In the second, purportedly unrelated session, participants were offered an opportunity to actually engage in those activities either in the next 2 days or several weeks later, and their behavioral intentions were assessed. As expected, participants' general attitudes better predicted their intention for the distant future than for the near future.

Another set of studies examined temporal changes in forming value-consistent intentions and applied Schwartz's (1992) value questionnaire to assess the importance participants assign to a wide range of values (e.g., power, benevolence, and hedonism). For example, one study asked participants to imagine 30 behaviors (e.g., rest as much as I can) and to indicate the likelihood of performing each behavior either in the near future or in the distant future. The researchers then correlated the rated importance of each value and the mean likelihood for performing the behaviors corresponding to that value. As predicted, these correlations were higher when the behaviors were planned for the distant future than when they were planned for the near future (Sagristano et al., 2006).

It is also possible to distinguish between values that are central to an individual and more peripheral, secondary values. When a situation is related to a number of different values, the individual's central values are more likely to guide choice from a psychologically distant than a proximal perspective, whereas the individual's secondary values are more likely to guide their choice from the psychologically proximal than the distant perspective. To examine this prediction, Eyal, Liberman, Sagristano, and Trope (2006) measured or manipulated the centrality of values and examined how they predict behavioral intentions. For example, one study assessed the relative centrality of achievement versus altruism values and examined near and distant intentions of solving a dilemma between getting ahead by working extra hours or helping a friend. Results indicated that people who were predominantly achievement oriented planned to be achieving in the distant future more than in the near future, whereas people who were predominantly altruistic planned to be more cooperative in the distant future than in the near future. In other words, participants solved the conflict in favor of the more central value in their own value priorities in the distant future more than in the near future. These results imply that distant-future decisions reflect predominant values whereas in near-future decisions secondary values are also brought into consideration. Interestingly, these results also suggest that individuals with different values are quite similar to each other with respect to their near-future plans but differ considerably in their plans for the distant future. Inasmuch as actual behavior is closer to near-future plans than to distant-future plans, it is also possible to contend that individual differences in values are reflected in distant plans more than in real behavior. This would be reminiscent of the basic notion of social psychology that individual differences often fail to materialize in actual behavior, despite the common intuition that such differences are essential, central, and important. Possibly, people feel that individual differences in values guide behavioral choice because we tend to think of behavioral plans for temporally distant or hypothetical situations.

COMPENSATORY STRATEGIES IN DECISION MAKING

Our choice findings may be interpreted as suggesting that choice for the near future is based on more dimensions and is more compensatory, whereas choice for the distant future gives more weight to a smaller number of dimensions and is, therefore, less compensatory. Decision-making research views looking at attributes within each alternative as the hallmark of compensatory search, because compensation requires summing, within each alternative, the contribution of each attribute to the overall value. Looking at each attribute across alternatives, on the other hand, typically characterizes noncompensatory choice strategies, such as elimination by aspect (Payne, Bettman, & Johnson, 1988). We think that, generally, looking within alternatives across attributes constitutes a lower-level construal than looking at attributes across alternatives, because alternatives are directly experienced entities whereas attributes are abstracted. Moreover, a comparison of alternatives requires construal, or transcending the directly experienced situation (e.g., one cannot experience simultaneously a number of apartments and compare by experience their level of noise), but many attributes of a single alternative may be experienced simultaneously and thus do not require transcendence (see Hsee & Zhang, 2004, for a related distinction between joint and separate evaluation). From this perspective, too, compensatory evaluation of alternatives and within-alternative search would be associated with low-level construal and would characterize processing of proximal decision situations, whereas noncompensatory evaluation and within-attribute search would be associated with high-level construal and therefore would characterize processing of distal decision situations.

This idea was recently tested in an information search study by Borovoy, Liberman, and Trope (2004). Partici-

pants were presented with a matrix of information in which rows represented alternatives (e.g., different apartments), columns represented attributes (e.g., price, location, and noise), and cells represented the standing of each alternative on the corresponding attribute. Participants searched this matrix by exposing the information in each cell, one at the time (see Payne et al., 1988, for a review of this paradigm). We told some participants that they are making a choice for the near future (e.g., that they are choosing an apartment to rent in the next 2 weeks) whereas other participants were told that they are making a decision for the distant future (e.g., that they are choosing an apartment to rent 1 year later). We counted the number of within-alternative steps (i.e., the number of cells that were opened immediately after opening a cell in the same row) and within-attribute steps (i.e., the number of cells that were opened immediately after opening a cell in the same column). As expected, we found more within-alternative steps and less within-attribute steps for more near-future decisions. Moreover, we found that participants opened an equal number of cells and invested a similar amount of time in both temporal distance conditions. Thus, they did not perform a more heuristic search for the distant future sets.

SUMMARY: THE EFFECT OF FUTURE TEMPORAL PERSPECTIVE ON PREFERENCES

Together, the studies reviewed here support the CLT analysis of temporal changes in the effects of high-level and low-level information on preference. A distant-future activity was chosen according to the main task, but a near-future activity was chosen more according to a secondary, filler task. Diagnostic value was the prime determinant of preferences regarding distant-future self-relevant feedback, but pleasantness of the feedback was influential in preferences for the near future. Distant future time planning, but not planning for the near future, was guided by desirability concerns without taking into account time constraints. A future gamble was chosen according to the value of the outcome, whereas a near-future gamble was chosen according to the probability of the outcome. A distant-future action alternative was represented in terms of pro arguments, whereas a near-future action alternative included a representation of more con arguments. Finally, behavioral intentions for the more distant future, more than intentions for the near future, corresponded to abstract and central attitudes and values, whereas more specific and secondary attitudes and values showed the reverse intertemporal pattern.

In many of these cases, an irreversible decision was made at the same point in time regarding near- or distant-future options. Moreover, at the time of the decision, similar low- and high-level information was available for both the near- and distant-future options. Nevertheless, low-level information was more influential in decisions regarding near-future options, whereas high-level information was more influential in decisions regarding distant-future options. Thus, these findings cannot be explained by temporal differences in availability of high- versus low-level information or by the ability to postpone the use of one of these types of information when it pertains to distant-future options.

Notably, our studies do not show more regressive choices for the more distant future. To the contrary, choices for the distant future tend to discriminate more clearly both between alternatives and between individuals. Thus, our participants were not simply uncertain or indifferent in their choices for the distant future, but, to the contrary, exhibited more differentiation and decisiveness regarding distant-future choices. We think that this was the case because they based distant-future preferences on higher-level construals, which are often more simple and schematic than construals of the near future. Our studies on information search and memory explicitly show that participants did not invest less effort and did not engage in shallower processing when thinking about more distant-future decisions. They did, however, apply a less compensatory, more attribute-based, rather than alternative-based, strategy in making decisions for the more distant future. We believe that this was the case because compensation and within-alternative search are based on a lower-level construal of the decision situation than looking within attributes. It is this noncompensatory strategy, we believe, that in many cases gives rise to more clear-cut preferences for the more distant future.

Past Temporal Distance and Preferences

Gilovich and Medvec (1995) found that people regret action in the recent past but inaction in the distant past. For example, when asked about regrets they have about their college years, students would say that they regret majoring in psychology, but older people would say that they regret not taking art classes. One explanation that Gilovich and Medvec proposed for their finding was that reasons for an action are more schematic and essential for the action (e.g., I like art) than reasons against an action (e.g., I am too busy) and therefore are better retained in memory. As a result, with time, the reasons for not taking an action become unclear, and a failure to act becomes less understandable and more regrettable. This explanation is consistent with our findings that pro arguments become more salient over temporal distance whereas con arguments become less salient over distance. Thus, the tendency to increasingly regret inaction over time distance may be explained within CLT as a result of the pro arguments being superordinate to con arguments (see above), and thus more salient at a larger distance.

Social Distance and Preferences

ADVISING VERSUS MAKING DECISIONS: PRIMARY AND SECONDARY ASPECTS

Kray and Gonzalez (1999) and Kray (2000) compared participants' own choices to the advice they gave to close and remote others. They found that in advising others, especially to more socially remote others, participants tended to give more weight to a single attribute that they

designated as the most important and less weight to other, more peripheral attributes. For example, when advising another person about choosing between two jobs, participants gave more weight to personal satisfaction (the most important dimension) and less weight to salary and location (the less important dimensions) than when choosing for themselves (study 2). In two other studies, conceptually similar results were shown to be stronger with a more distant social target (a student in another department) than with a closer target (a student in one's own class). Moreover, it was found that when rating the importance of attributes, advisories to others tend to give more polarized ratings, favoring central attributes and discounting unimportant ones, whereas decisions for oneself tend to have a more balanced view, in which more similar importance ratings are assigned to both important and less important attributes. Kray also found that participants reported more responsibility and potential regret, a larger number of generated decision-relevant attributes, and less regressive choice when making decisions for others than for oneself, from which she concludes that it is implausible that people simply invested less effort in advising others than in deciding for themselves. In our terms, these findings demonstrate choosing according to more central, high-level aspects for more socially remote targets and applying less compensatory strategies for others than for oneself. Both of these results parallel our findings on future temporal perspective and, we believe, may be similarly explained within the framework of CLT.

THE EFFECT OF POWER ON WEIGHTING PRIMARY AND SECONDARY ASPECTS

As discussed earlier, theory and research suggest that powerful individuals feel more independent of others and, therefore, more distinct and distant from others than people with less power (e.g., Hogg, 2001; Hogg & Reid, 2001; Lee & Tiedens, 2001; Snyder & Fromkin, 1980). We also argued that this predisposes the more powerful individuals to adopt a distal perspective on the immediate situation and to form high-level construals of information about the situation (see Smith & Trope, 2006) and therefore to be more attuned to the primary outcomes the situation affords. In contrast, powerless individuals would divide their attention between central and peripheral aspects of the situation and would, therefore, be less likely to form an unequivocal action orientation in line with what the situation affords.

Initial evidence in support of this idea comes from a series of studies by Guinote and Trope (2004). One of these studies assigned participants to the roles of judges or workers. The judges evaluated the performance of the workers and controlled their payment and thus constituted a higher-power role. Before starting this task, in a purportedly unrelated experiment, participants were asked to imagine themselves in two situations, and to describe 1 day in their lives while in those situations. One situation described a friend visiting, whereas the other situation described doing an internship. Both situations provided opportunities to engage in work and social activities, but in the internship situation work activities are more central, whereas in the friend visit situation social activities are more central. The results showed that participants planned more work-related than social-related activities in the internship situation and more social-related than work activities in the friend visit situation. More important here, this difference was stronger for participants in the more powerful role than for participants in the less powerful role. As predicted, then, the more powerful individuals seemed more responsive to primary aspects of the situation and less responsive to secondary aspects of the situation. These findings are consistent with the assumption that greater social power increases the weight of high-level construals in individuals' behavioral choices. Positions of social power often require individuals to take a global orientation, plan ahead, and take decisive action. By promoting high-level construals, social power may naturally support such requirement.

CHOOSING VERSUS PREDICTING OTHERS' CHOICES: PROBABILITY AND VALUE

We argued earlier that in positive bets, payoffs may be conceptualized as being at a higher-level construal than probabilities. If this is true, then the weight of payoffs would increase over distance whereas the weight of probabilities would show the reverse effect. Consequently, people would take more risky (i.e., lower-probability, higher-value) distant bets but more conservative (i.e., high-probability, low-payoff) proximal bets. Self–other differences in choice of positive bets seems to support this prediction. Hsee and Weber (1997) asked participants to make a series of choices between a sure outcome and a risky outcome (e.g., getting $800 for sure vs. a 50% chance to get $2,000) and also to predict the choice that other people would make. They found that people thought that others would take more risky bets than themselves. Moreover, the discrepancy between one's own risk preferences and the predicted risk preferences of others was related to the abstraction of the other person: It was stronger for abstract others, of whom participants had no image, than for concrete individuals, whom participants could see but did not know. Hsee and Weber explained their results in terms of the risk-as-feeling hypothesis according to which people's risk preferences are dependent on their feelings toward risk. They contended that people are more likely to empathize with a more concrete individual and perceive his or her feelings as similar to their own. Consistent with this view, CLT suggests that others, and particularly distant others, are construed more abstractly. Therefore, in predicting their preferences, people give more weight to value and less weight to probability.

Probability and Preference

CLT predicts that increasing the likelihood of an event would decrease the weight of desirability-related features relative to the weight of feasibility-related features in decisions, thereby decreasing (or even reversing) the pref-

erence for a more desirable but less feasible outcome over a less desirable but more feasible outcome. A recent series of studies by Todorov, Goren, and Trope (in press) tested this prediction. In one study, participants were told that a number of companies in New York had started promotional campaigns for their products and services, and that some of them had offered a special promotional plan to NYU students. In the high-probability condition, participants were told that if they signed up for the campaign, they were almost certain to receive a voucher for the company's products. In the low-probability condition, participants were told that they would have about a 1 in 100 chance of receiving a voucher. In both probability conditions, the campaign was described as offering either a highly desirable but less feasible outcome or a less desirable but highly feasible outcome. For example, the highly desirable outcome was receiving 10 CDs from a Tower Record store, but to claim the CDs, the voucher had to be presented at a store at an inconvenient location. The less desirable outcome was receiving one CD that could be claimed at a convenient location.

Participants indicated their willingness to sign up for the campaign. The results showed that whereas under low probability, participants preferred the highly desirable but less feasible alternative to the less desirable but highly feasible alternative, under high probability this preference was reversed. Similar results were obtained with a variety of other prospects pitting desirability and feasibility. In all these cases, participants gave more weight to the outcome's desirability than to its feasibility when the outcome was improbable, but not when it was highly probable. It seems then that that probability changes the weights of outcome feasibility and desirability in much the same way as temporal proximity does.

Construal Level and Preference

As mentioned earlier, we believe that psychological distancing affects preference via construal, by enhancing the weight assigned to high-level value relative to low-level value. It is also possible to examine directly the effects of construal level on preference—although level of construal is not conceptualized as a dimension of distance, CLT predicts the effect of higher-level construal to be similar to that of psychological distancing. Evidence in the domains of risk taking and self-control seems to support this claim. We turn now to examine it.

RISK TAKING

Level of construal appears to have a crucial effect on the relative weight of probabilities and expected utility in risky choice, particularly in natural, real-life decisions. For many decision situations, one can choose whether to view them as one in a series of similar events or rather as unique, one-time occurrences. For example, when facing a decision on whether to invest in developing a new herbal treatment based on dried violets for stress-related headaches, an executive of a drug company may view the decision as one in series of decisions about novel medicines, or, alternatively, as a unique, one-time decision about applying dried violets for treating stress-related headaches. Note that abstraction is needed in order to adopt the former view, as it requires ignoring specific, incidental, and contextual features. From that perspective, viewing the decision as one in a series of similar decisions constitutes a high-level construal of the situation.

Kahneman and Lovallo (1991) proposed that risk avoidance in real life often stems from a narrow categorization of the decision situation as a unique, one-time event. In our example, if the decision is construed with all of its rich specific and contextual details, it will be viewed as unique and would produce risk avoidance. If, on the other hand, the manager would think of the product as one of a series of developments, and one of a series of risks that the company and she personally are to take, then her willingness to take the risk would increase. This is because normatively, the outcome of aggregated gambles is more likely to be close to its expected utility than the outcome of a single gamble. In other words, risk is reduced with repetition (see Lopes, 1996). In our terms, Kahneman and Lovallo's analysis suggests that less risk aversion (in fact, less weight for the dimension of risk and more weight for expected utility) ensues from a higher level of construal, a notion that is compatible with the findings on the enhancing effects of psychological distance on risk taking reviewed earlier. More generally, in series of risky events, a high-level construal may correspond to aggregation, whereas a low-level construal may be related to a viewing each event separately. Because aggregation reduces risk aversion, it follows that high-level construal of risky events would be associated with reduced risk aversion.

SELF-CONTROL

Situations that require self-control involve a conflict between two opposing motivations (e.g., a desire to go out with friends and a need to study for an exam). Recently, Fujita, Trope, Liberman, and Levin-Sagi (2006) proposed an analysis of self-control conflicts as conflicts between behavioral implications of high-level construal (i.e., value that is related to primary, central, goal-relevant, superordinate considerations) and behavioral implications of low-level construal (i.e., value that is related to secondary, incidental, goal-irrelevant, subordinated, features). Failure of self-control, according to this proposal, is succumbing to the motivation implied by the low-level value. For example, if studying for an exam is related to more superordinate goals than going out with friends, then the latter behavior would represent a failure of self-control. Consistent with this analysis, a series of studies by Fujita and colleagues demonstrated that higher-level construals increased self-control. In one study, participants first completed a task in which they indicated either why or how they would maintain good physical health. This task was designed to induce a high- or low-level construal mindset, respectively. Participants were then asked to hold a handgrip while connected to bogus electrodes, os-

tensibly as part of a psychophysiological assessment of personality. Participants were told that the longer they hold the handgrip, the more diagnostic was the information obtained from the apparatus. Thus, the situation presented a conflict between a desire to get diagnostic, self-relevant information (high-level value) and the inconvenience of holding the handgrip (low-level value). The results indicated, as predicted, that participants in the high-level construal condition held the handgrip longer than those in the low-level construal condition.

The finding that high-level construal produces greater self-control than low-level construal is consistent with the conceptualization of self-control as involving a conflict between behavioral implications of low-level versus high-level value. CLT further predicts that temporal, spatial, and social distancing would also enhance self-control. Consistent with this prediction, extant research on temporal distance has shown that people are better able to commit to self-control a long time in advance than a short time in advance (Ainslie & Haslam, 1992; Frederick, Loewenstein, & O'Donoghue, 2003), and when the temptation is physically distance than when it is near (Metcalfe & Mischel, 1999; Mischel et al., 1989).

Summary: Psychological Distance and Preference

It seems ironic that when outcomes become psychologically proximal and likely to materialize, people tend to overweight secondary considerations in their choices. Conversely, it is when outcomes seem unlikely or remote or refer to distant times that people's primary concerns are more likely to guide their preferences. In other words, people appear to be better able to act according to their priorities with respect to distal options than proximal options. This counternormative (and perhaps counterintuitive) conclusion raises interesting self-regulatory questions regarding people's ability to express core aspects of their self-identity in remote versus proximal situations.

We demonstrated effects of psychological distance on evaluation and choice, and we believe that these effects are mediated by the effects of psychological distance on construal. It is important to note, however, that psychological distance may affect not only the perceived value of outcomes but also the motivation to pursue the outcomes, by affecting the outcomes' expectancy. The effects of distance on expectancy are outside the scope of CLT but are worth discussion in order to prevent possible confusion. It is to that discussion that we now turn briefly.

Motivation researchers have related expectancy to uncontrollable probability of outcomes (e.g., in bets; Edwards, 1955), task difficulty (Atkinson, 1957), controllability (Locke & Latham, 1990; Rotter, 1966), and self-efficacy (Bandura, 1982). For our purposes, it is instructive to note that psychological distance may often reduce expectancy in these various forms. For example, people typically have less control over more socially distant individuals. People typically control others less than themselves and strangers less than friends and relatives. Control also diminishes with spatial distance, as it is often

more difficult to do something about things that are farther away. Events in the distant future (e.g., a distant-future rebate) typically involve some uncertainty (e.g., would the company still be around to pay the rebate), which increases with time distance, as noted already by Keren and Roelofsma, 1995 (see also Frederick et al., 2003, and Prelec & Loewenstein, 1991, for a proposed similarity between the effects of time and probability). Proximity also increases efficacy by making action more crucial for achieving an outcome. For example, a long time before an exam, failing to study may be compensated by studying more intensely later, but a short time before the exam, when only a few hours remain, such possibility no longer exists. As another example, failing to help a stranger may be compensated by help by another stranger, but failing to help a close friend is less likely to be compensated by someone else, because people typically have fewer close friends than strangers around them.

Psychological distance may thus decrease motivation due to changes in expectancy. For example, students may be more motivated to study for a close exam than for a distant exam because not studying is easier to compensate for at a distance, or people may be more motivated to help a close friend than a stranger, because in the latter case, their lack of help is more likely to be compensated by others. These changes in motivation do not reflect changes in value (i.e., it is not necessary to assume that the value of a success on the exam increases closer to it) although such changes may, of course, exist. Moreover, changes in expectancy over psychological distance are not mediated by construal but, rather, are real, objective changes (e.g., opportunities for compensation really decrease closer to an outcome) that do not require a psychological theory to explain them. For that reason, they fall outside the scope of CLT (see Liberman & Trope, 2003, for a detailed discussion of that point with respect to temporal perspective).

It is possible however, that motivation contaminates measures of value. For example, it is possible that one's motivation to study for an exam and one's arousal or level of energy would affect one's answer to a question about the importance of succeeding in an exam. If contaminated with motivation, measures of value would show discounting over psychological distance. It is possible that this is one of the reasons for the widespread assumption that value is discounted over psychological distance. CLT proposes, however, that if value is measured independently of motivation and thus is not affected by possible changes in expectancy, then both discounting and augmentation over psychological distance become possible, depending on the construal level with which value is associated.

Affect

How does psychological distance influence affective responses? It is commonly assumed that the intensity of affective reactions decreases with psychological distance. People typically react more strongly to events that are

closer to them in time and space, to events that happen to themselves than to others, and to events that are real more than to hypothetical events. It seems that the effect of the various dimensions of psychological distance is similar and involves reduction in the intensity of affective responses. Theoretically, CLT would predict that affective responses are diminished over temporal distance only if they are low level, and, contrary to that, high-level affective responses may be augmented over time. But is there high-level affect?

Indeed, emotion researchers typically identify affective responses with low-level, concrete processing and contrast them with cognitive responses, which are considered to be more abstract and high level (Loewenstein, 1996; Metcalfe & Mischel, 1999). It is possible, however, that there are different types of affect, and that some may be characterized as more abstract and higher level than others. An interesting possibility would be to distinguish between high- and low-level affect by the extent to which they require going beyond one's direct experience here and now (i.e., the extent to which they necessitate distancing and construal). For example, social emotions such as pride and guilt involve considering the perspective of other people and hope involves considering the future; counterfactual emotions, such as disappointment and regret, involve considering alternatives to reality. Interestingly, it has been suggested that anxiety, contrary to fear, is an apprehension of a potential source of danger at a location other than the immediate (Gray & McNaughton, 2000). All these emotions are distinct from hunger, thirst, pain, anger, happiness, sadness, and fear, which do not necessitate transcending one's direct experience (but, of course, allow for such transcendence, as one can, for example, become angry by thinking of a hypothetical event). We refer to this distinction in terms of level: higher-level emotions are those that require more distancing and construal, whereas lower-level emotions are those that do not necessarily require as much distancing and construal. This distinction between levels of emotion partly overlaps with extant distinctions in the literature between basic emotions and other emotions (Ekman, 1992; Izard, 1977, 1992; but see Ortony & Turner, 1990, for questioning the validity of this distinction). Interestingly, in these theories, emotions are called basic because they are assumed to have innate neural substrates, innate and universal expressions, and unique motivational states. These criteria are very different from the construal versus experience criterion that we propose here. Thus, although the distinctions proposed in these theories overlap to some extent, the theoretical basis that underlies the distinction is different.

The classification of emotions according to level is content based, as it distinguishes between qualitatively different emotions. Besides this distinction, according to CLT, the level of the exact same emotion can differ depending on the situation in which it occurs. For example, as noted before, central aspects of situations constitute higher-level construals than do peripheral aspects. Therefore, affective aspects that are made central (e.g., by virtue of being goal relevant) would be of a higher level than similar aspects that are made peripheral (e.g., irrelevant to the main goal). For exam-

ple, in a funeral, sadness about the transience of life and compassion are central and thus constitute high-level emotions, but happiness upon seeing old friends is peripheral and thus constitutes a low-level, peripheral emotion. In a birthday party, however, the centrality of these emotions reverses.

For both content-based and situationally based variations of level of affective responses, CLT proposes that low-level affective responses, more than high-level affective responses, would be diminished over psychological distance. The latter, we think, may be discounted less or even augmented with distance. Let us now review literature on the effect of various distance dimensions on affective responses.

Future Temporal Perspective and Affect

It is commonly assumed that temporal distance diminishes affective reactions. Research on delay of gratification (Mischel et al., 1989; Mischel, Ayduk, & Mendoza-Denton, 2003) and self-control (Baumeister & Heatherton, 1996; Loewenstein, 1996; Metcalfe & Mischel, 1999) has documented that from a distance, people underestimate their affective reaction and overestimate their ability to make "cold," rational, unemotional decisions. For example, the positive value of watching a funny movie is affective or "hot," whereas the positive value of studying for an exam is cognitive. If the value of watching the movie is discounted more steeply than the value of studying, a temporally inconsistent pattern of preferences emerges wherein studying is preferred a long time in advance but watching the film seems more attractive from a closer perspective. According to this approach, then, temporal distance should always increase the relative weight of cognitive (vs. affective) value in preference.

As noted before, an interesting prediction of CLT is that both affective and cognitive types of value could be either high level or low level and, therefore, could be both augmented and discounted over time perspective. This prediction was tested in a study that independently manipulated the affective–cognitive dimension and level of construal (Trope & Liberman, 2000, study 5). In this study, we assessed desirability ratings of four films varying in affective value (funniness) and cognitive value (informativeness). The films were, thus, funny and informative, funny but uninformative, not funny but informative, or neither funny nor informative. Some of our participants expected to watch the films in the same experimental session, whereas other participants expected to watch them in the second session of the study, 2 months later. The goal of watching the films was also manipulated: It was either affective (getting oneself into a good mood) or cognitive (learning about a topic). We assumed that the features of the film that are related to the goal would be more central than the goal-irrelevant features and thus would constitute a high-level construal of the film. Thus, depending on the goal, either affective features or cognitive features of the films were more central (constituted the high-level construal of the films), whereas the other type of features was rendered goal irrelevant and thus part of the low-level construal of the films.

We examined how time perspective, goal, affective value, and cognitive value influenced the desirability ratings of the films. Consistent with the predictions of CLT, we found that temporal distance increased the influence of the informativeness versus the funniness of the films when the goal was cognitive but decreased the influence of informativeness versus the funniness of the films when the goal was affective. Thus, the effect of high-level, goal-relevant value increased over delay relative to the effect of low-level, goal-irrelevant value.

Social Distance and Affect

A large amount of social psychological research suggests that liking, love, hate, and empathy decrease with social distance. For example, reducing similarity reduces liking and empathy toward a social target (Byrne, 1971; Newcomb, 1961). In terms of CLT, liking and empathy are low level, because they do not require construal or transcending one's direct experiences. We would predict that higher-level interpersonal emotions, such as shame and guilt, would be less discounted over social distance. For example, we would predict that reducing similarity would not reduce and may even increase one's feelings of shame regarding another person.

Interestingly, the literature on empathy distinguishes between cognitive empathy, which involves taking a perspective of another person, and emotional empathy, which involves contagion with another person's emotions and does not involve perspective taking (Davis, Hull, Young, & Warren, 1987; Preston & de Waal, 2002). For example, cognitive empathy involves understanding the difficulties faced by another person, whereas affective empathy involves feeling the other person's pain. In our terms, cognitive empathy pertains to a higher level of construal than emotional empathy. CLT therefore predicts that emotional empathy would be discounted over social distance (and, in fact, any psychological distance) more than cognitive empathy.

Physical Distance and Affect

Closely related to empathy is the ability to experience another person's pain and take it into consideration in one's actions. As discussed previously, experiencing another's pain is an emotional contagion and thus would be classified as emotional empathy, or, in our terms, low-level empathy. We would therefore predict that it would decrease over physical distance. Milgram's (1965) studies on obedience are consistent with this prediction, as they show that physical proximity to the victim reduced the willingness to obey an order to inflict pain on him. It is possible that this was the case because physical proximity enhanced participants' emotional empathy toward the victim. In other situations, too, physical proximity plays a major role in empathic concern. For example, it has been suggested that the cruelty of war has intensified with introduction of weapons that are remotely operated (e.g., aerial bombardment). Indeed, Latané's (1981) social impact theory specifies physical distance as one determinant of social impact in general and emotional impact in particular.

The Effect of Level of Construal on Affect

As with the effects of social distance on evaluation and choice, we believe that the effects of distancing on affect are mediated by its effects on construal. Thus, it would be interesting to examine evidence for the effects of construal level on affect. According to Metcalfe and Mischel (1999), stimuli can be mentally represented either in terms of their emotionally arousing "hot" features or their cognitive, informational, "cool" features. These representations are intrinsically connected to two regulatory systems: hot representations elicit emotionally driven, reflexive, "hot" responses that are predominantly under stimulus control and generate automatic approach–avoidance behaviors. Cool representations, on the other hand, elicit cognitively driven, reflective, "cool" system responses whose functioning requires more effortful and conscious control. Effective self-regulation of reflexive responses associated with the hot system is possible to the extent that individuals have and can access cognitive representations (i.e., distraction, reappraisal, and abstraction) that help cool intrinsically hot stimuli. In this view, then, abstract representations are associated with cool, less affective responses. Mischel's (1974) work on delay of gratification provides support for this idea. It demonstrates that an effective way to overcome immediate temptations and successfully delay gratification is to turn attention away from the concrete qualities of the immediate temptation (e.g., a tasty cookie) and focus on its abstract qualities (Mischel et al., 1989; see also Mischel et al., 2003).

A recent study by Kross, Ayduk, and Mischel (2004) examined the joint effects of construal level and social distance on people's ability to reduce experienced negative emotions. In the study, participants recalled a past interpersonal experience in which they felt overwhelming anger and hostility. Type of perspective was manipulated by making some participants think of what emotions are being felt and other participants on why they are experiencing those emotions. In our terms, the former corresponds to a low-level construal whereas the latter corresponds to a high-construal level. They also manipulated the type of perspective, by asking some people to be immersed in the experience, or taking a first-person perspective, and other participants to take a perspective of a distanced observer. In our terms, this corresponds to a manipulation of social distance. The authors hypothesized that a low-level, "what" focus would activate relatively concrete, hot representations of the specific emotions and thus should be associated with a high level of negative affect whereas a high-level, "why" focus may produce either hot representations or cool representations, depending on the type of people's perspective. Specifically, they predicted that a "why" focus would attenuate negative affect only in a self-distanced perspective but not in a self-immersed perspective.

Emotional response of anger was measured implicitly, by a word completion task, and explicitly, by asking par-

ticipants about their mood. The results showed, as predicted, that participants who focused on the reasons underlying their feelings (a high-level, why construal) and maintained a distanced social perspective (an increased social distance) manifested the lowest levels of anger on both explicit and implicit measures. In our terms, these results indicate an interactive effect of construal level and social distancing, such that emotions decrease when both are high. It would be interesting to repeat this study with higher-level emotions such as guilt. CLT would predict that in that case, distance would not attenuate emotional reaction but might in fact increase it.

Theories on affective reactions to exemplars versus categories also suggest that the former elicit more intense affective reactions than the latter (Sherman, Beike, & Ryalls, 1999). For example, while the idea of "saving whales" leaves people relatively unemotional, "saving Willy" (a concrete whale) often touches them deeply and is reflected in an enhanced willingness to donate money to that end (Kogut & Ritov, 2004; Sherman et al., 1999). As in other cases of distance effects on emotion, we predict this to be the case with low-level emotions more than with high-level emotions. Thus, we predict that compassion toward a single whale may be paradoxically reduced if one thinks about an entire category of suffering whales, but that feelings of guilt and shame (e.g., for polluting water) would not be reduced by shifting one's attention from one concrete victim to an entire category of suffering victims.

Summary: The Effect of Psychological Distance on Affect

A considerable amount of theoretical and empirical work suggests that psychological distance reduces the intensity of affective responses. People react less strongly to temporally and physically distant events than to close events. They feel less for more socially distant individuals and react less strongly to less likely occurrences. CLT not only provides a unifying framework for these various effects but also posits that these effects would be weaker for emotions that are of a higher level of construal either because they are central or because they require transcending the immediate experience. Thus, CLT suggests that the effects of distance on affect depend on level of construal, in the same way that the effect of distance on prediction and evaluation depends on level of construal.

Creativity

The last area in which we examine the implications of CLT is creativity. Creativity has been assumed to profit from abstract thinking (e.g., Finke, 1995; Ward, 1995), and performance on a variety of creativity tasks seems to depend on more abstract construals of problem components. For example, creativity on alternative-uses tests (e.g., generating reasons why to greet somebody; Friedman & Förster, 2002; Schoppe, 1975) should be enhanced by construing the action more abstractly (e.g., as a gesture of communication) rather than more concretely (e.g., as "saying hello"). Therefore, whereas the former might lead to solutions that are more remote and diverse from the actual object, the latter might render common associates accessible, impeding innovation (see Marsh, Ward, & Landau, 1999). If creativity is enhanced by abstract representation of problem elements and if, as discussed earlier, distant perspectives engenders higher-level construals, then psychological distancing may promote creative thinking.

The Effect of Future Temporal Perspective on Creativity

In a series of studies, Förster, Friedman, and Liberman (2004) tested the idea that distancing would enhance creativity by manipulating distant versus near-future time perspective and gauging performance on a variety of creativity tasks. One of the studies assessed performance on insight problems. The following is an example: "A prisoner was attempting to escape from a tower. He found a rope in his cell that was half as long enough to permit him to reach the ground safely. He divided the rope in half, tied the two parts together, and escaped. How could he have done this? [Solution: He unraveled the rope lengthwise and tied the remaining strands together]." The natural way to imagine "cutting the rope in half" does not involve unraveling it. One needs to abandon this concrete image and represent the action more abstractly in order to construe this action in an alternative way. Förster and colleagues asked participants to think about themselves and their lives as they would be a year later (or the next day) and then imagine working on the insight problems at that time. For three such problems, participants displayed more insight in the distant-future condition than in the near-future condition. Interestingly, Förster and colleagues (2004, studies 2 and 3) found that temporal distancing facilitated insight not only in verbal tasks but also in visual tasks, which require abstraction of coherent images from fragmented or "noisy" visual input (e.g., the Snowy Picture Test and the Gestalt Completion Test; see Friedman & Förster, 2000, 2002). Other studies demonstrated that temporal distancing facilitates abstract reasoning but not concrete reasoning (Förster et al., 2004, studies 4 and 5).

Creative Insight and Other Psychological Distances

Anecdotal evidence and some research findings suggest that other dimensions of distance might produce similar effects on creativity. People sometime feel that they had their most ingenious ideas in circumstances quite distant from and dissimilar to their usual working environment. "Incubation" may also testify to a similar connection between distancing and creative insight. Sometimes, after repeated and fruitless attempts to solve a problem, the solution may pop up after one takes some distance from the problem by leaving the problem for a while and doing other things (temporal distancing) or by changing the physical context (spatial distancing). It has been argued that one could become more innovative in solving technological problems by imagining oneself as a little dwarf entering the subject matter of the problem at hand (e.g., entering the computer chip, to solve the problem of their overheating), or to change one's physical location to a

very atypical one (e.g., sitting under one's office desk) in order to "break set" and achieve a fresh look at the problem (Helfman, 1992). Adams (1986) and de Bono (1985; see also Butler & Kline, 1998) suggest that considering the perspective of other people who are involved in a problem might help generate better and more creative solutions. Systematic research on these issues is lacking, and further research is needed to better understand the effect of psychological distancing on creativity and reasoning.

Distancing, decontextualization, and abstraction seem to play a role not only in creative problem solving but also in production and perception of art. In art in general and in modern art in particular, distancing oneself from the concrete shape or material of a piece of art or from its mundane nature appears to be an important prerequisite for appreciating its artistic value. What is true for the perception of art also holds for its production: Deciding that a mundane object is or can be an object of art affords some abstraction. In fact, artistic perception requires abstraction of new meaning from concrete, directly experienced percepts, a meaning that is oftentimes different than the most common and usual abstractions people apply in everyday life. In that, artistic perception is similar to solving insight problems (Arnheim, 1969). Thus, although not empirically tested, distancing and abstraction seem to play a major role in creative problem solving, creative perception, and creative production. We believe that empirically exploring this possibility would be a fruitful avenue for future research.

INTERRELATIONS AMONG DIMENSIONS OF PSYCHOLOGICAL DISTANCE

In the final section of this chapter we examine the idea that all the dimensions of psychological distance are interrelated. We think that this should be the case because they share the feature of being distances from the same thing—one's direct experience—and because, as a consequence of diverging from direct experience, they have similar effects on construal. This section reviews empirical evidence regarding explicit and implicit associations among distance dimensions.

Social Distance and Other Dimensions of Psychological Distance

We now review research testing the interchangeability of social distances dimensions. We examine research on the effects of temporal distance and spatial distance on social distance and then turn to research on the reverse effect, namely, the effect of social distance on temporal and spatial distances.

The Effect of Temporal Distance on Social Distance

A series of studies by Stephan (2004) examined the effect of temporal distance on social distance. We predicted that an increase in temporal distance would produce a corresponding increase in social distance. One study, for

example, used politeness as an indicator of social distance. As mentioned in the first section of this chapter, research on politeness assumes that politeness varies as a function of social distance (Brown & Levinson, 1987). Therefore, the degree of politeness a speaker chooses may be used as an indicator of his or her perceived social distance from the recipient of the message. In our study, participants wrote instructions for a person who was expected to read them either in the near future or in the distant future. For example, participants wrote sightseeing suggestions for a tourist who was supposed to arrive on the following day or 1 year later. Then, participants rated how polite they intended to be in phrasing their suggestions. As predicted, participants indicated higher levels of politeness for the tourist who was expected in the more distant future.

The Effect of Spatial Distance on Social Distance

Spatial distance has been recognized for a long time as a major factor in creating and maintaining social ties (Festinger, 1951). For example, it has been shown that friendship is more likely to develop among spatially close than distant individuals, unless an initial antagonism has existed in the relationships (Festinger, 1951; Schiffenbauer & Schiavo, 1976). This is hardly surprising, as increasing physical proximity may create opportunities for social interaction and expose people to similar experiences (see Festinger, Schachter, & Black, 1950; Priest & Sawyer, 1967). Would a similar relation hold if distance is manipulated independently of these additional factors?

In a study aimed to answer this question, Reichman and Ben Arie (2004) examined the effect of actual sitting distance on politeness. They entered the first meeting of a SAT preparation course. The participants were adults who did not know each other in advance. Participants were asked to write two notes (one explaining how to use the course web site, and the other asking for advice on how to prepare for an exam) to a person sitting next to them, to a person sitting in another class, behind the wall, or to a person in a similar class in another town. After writing the notes, participants rated their own responses for intended politeness. We found that participants addressed more politely others who were in another class or in another city compared to others who were in the same class. Thus, spatial distance affected social distance, which was expressed, in this study, by the chosen level of politeness.

The Effect of Social Distance on Temporal Distance

Stephan (2005) conducted a study to examine the effect of social distance on temporal distance, using politeness to manipulate perceived social distance. Participants were asked to imagine two people conversing and one of them telling about an action he or she intended to perform. Participants then read a statement that was phrased in either normative language or colloquial language. For example participants in the colloquial (normative) condition read: "Sharon decided to keep an eye

on Lisa's baby (to watch Lisa's baby more carefully)." Colloquial utterances are less polite and indicate greater social proximity than normative utterances. Participants indicated how much time later the protagonist would perform the action. Consistent with our prediction, the mean enactment time for the colloquially phrased statements was nearer than for the normatively phrased statements, indicating that closer social distances produced a perception of sooner enactment times.

The Effect of Social Distance on Spatial Distance

Studies on personal space show that social distance affects the spatial distance people prefer to keep from each other. The more familiar and comfortable people are with each other, the closer the physical distance they maintain from each other (for a review, see Hayduk, 1983). People also tend to physically distance themselves from feared or stigmatized others (e.g., AIDS patients—Mooney, Cohn, & Swift, 1992; stereotyped groups—Macrae, Bodenhausen, Milne, & Jetten, 1994; Worthington, 1974). In these studies, however, socially distant targets were also less attractive. To distinguish between these two factors, we manipulated social distance as the degree of politeness used between interlocutors. In our study, participants received a sketch on which the location of the speaker was marked and read a phrase that the speaker said to the addressee. The phrase was either colloquial or normative, for example, "My brother is taking our family car, so the rest of us will stay at home (will be stuck at home)." Participants indicated on the sketch the location of the addressee. As predicted, the use of more polite (normative) language by the speaker produced a perception of a greater physical distance between the interlocutors, as compared to the use of less polite (colloquial) language. Thus, this study demonstrates that people who address each other more politely are expected to communicate across greater spatial distances. Colloquial language, which signified social closeness, produced estimates of smaller spatial distances.

Automatic Associations among Distance Dimensions

We claim that different dimensions of psychological distance share an important aspect of meaning, namely, that they all are distances from direct experience and as such require construal. Moreover, it is possible to assume that psychological distance would be an important aspect of stimuli that would be spontaneously encoded. The Stroop paradigm (Stroop, 1935) offers an apt way to examine these predictions. In a typical Stroop task, participants are faster at naming the ink color of semantically compatible words (the word "blue" or the word "sky" printed in blue ink; the words "green" or the word "grass" printed in green ink) than at naming the ink color of semantically incompatible words (the word "blue" or the word "sky" printed in green ink; the word "green" or the word "grass" printed in blue ink). These results are interpreted as suggesting that reading the words and assessing their semantic meaning is automatic (happens spontaneously, without participants' intention and despite the fact that it does not help in performing the experimental task), and that there is shared meaning between the two dimensions of the task (ink color, the semantic meaning of the words).

A series of experiments by Bar-Anan, Liberman, Trope, and Algom (2005) demonstrated a similar effect with distance-compatible versus distance-incompatible stimuli. The experiments applied a picture–word version of the Stroop task, in which the participants discriminate between cues of one psychological distance dimension while ignoring cues of another psychological distance dimension. We reasoned that if psychological distance is a shared meaning of spatial distance and the other three dimensions, then it would be easier to perform the task when the relevant and the irrelevant cues are congruent in psychological distance than when the relevant and irrelevant cues are incongruent in terms of psychological distance. For example, we predict that participants would identify a stimulus as spatially proximal faster when the irrelevant stimulus is a word that denotes psychological proximity (e.g., the word "we," which represents social proximity, printed on a spatially proximal object) rather than a word that denotes psychological distance (e.g., the word "others," which represents social distance, printed on a spatially proximal object).

Bar-Anan and colleagues (2005) used perspective pictures (e.g., a picture of an alley of trees or a picture of rolling hills). An arrow pointing to either a proximal or a distal point on the landscape was shown on the picture, and a word denoting a psychologically proximal entity ("tomorrow," "friend," "we," or "sure") or a psychologically distal entity ("year," "enemy," "others," or "maybe") was printed on the arrow (see Figure 15.1). In some of the experiments, the task was spatial discrimination, namely, participants indicated whether the arrow pointed to a spatially proximal or distal location. In other experiments, the task was semantic discrimination, namely, participants indicated whether the word on the arrow was, for example, "we" or "others." In both types of tasks, and across all four dimensions of distance, participants were faster in responding to distance-congruent than to distance-incongruent stimuli. These results demonstrate that people assess the psychological distance of stimuli that pertain to spatial distance, temporal distance, social distance, or hypotheticality, even when this information is irrelevant to their current goal and that these various distances share a common aspect of meaning. We interpret these results as initial evidence that psychological distance is a basic, automatically activated aspect of each of these dimensions.

CONCLUSION

As this review illustrates, there is a large amount of research across the behavioral sciences on how people respond to events from the recent versus distant past, near versus distant future, to spatially near versus far objects, to themselves versus others, and to real versus hypothetical, probable versus improbable events. Different theoretical approaches and research paradigms have been
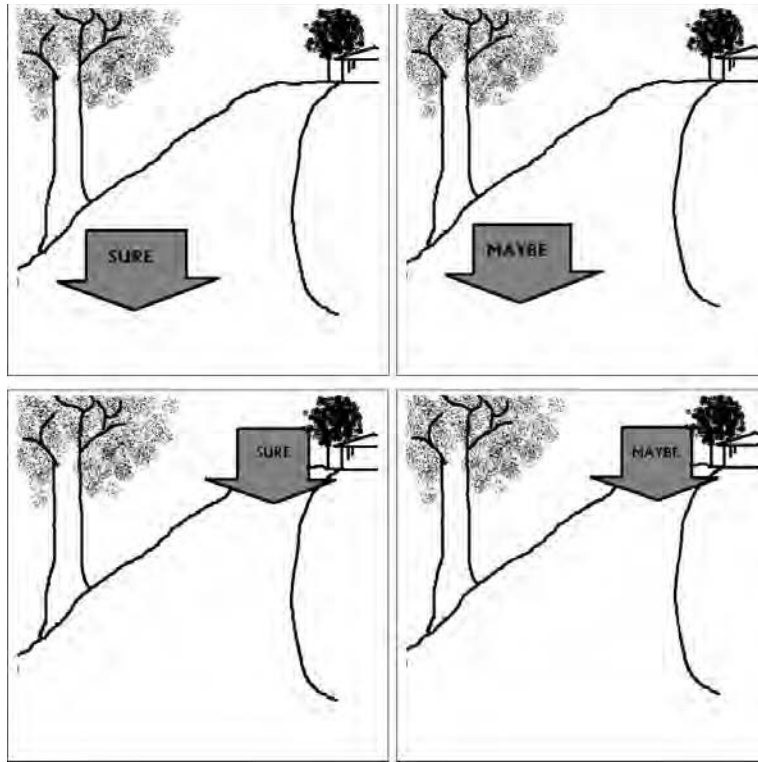
FIGURE 15.1. Stimuli used in the Stroop task, examining associations between spatial distance and hypotheticality. Reaction times in both distance discrimination and word discrimination tasks were faster with congruent pairs (distant with "maybe"; proximal with "sure") than with incongruent pairs (distant with "maybe"; proximal with "sure").

proposed for each of those dimensions. Without denying the uniqueness of each dimension, we propose that they also have something in common, that they all constitute dimensions of psychological distance. At their point of origin is one's direct experience of the "here and now." Transcending this point entails constructing mental models of what is not directly experienced, and the farther removed an object is from direct experience on any distance dimension, the higher (more abstract) the level of construal of that object.

Consistent with this proposal, the research reviewed in this chapter suggests that different distance dimensions are interrelated. For example, distancing an object on one dimension may make it seem more distant on other dimensions, and the psychological distance of objects is assessed spontaneously, across different dimensions. Moreover, a large body of research shows that temporal distance, spatial distance, social distance (e.g., self vs. other and ingroup vs. outgroup), hypotheticality, and (im)probability are all associated with higher levels of construal. That is, the same information about more distant objects is represented more schematically in terms of few superordinate, core features of the object. These construals, in turn, expand one's horizons and guide prediction, evaluation, and action with respect to psychologically more distant entities. Indeed, the present review shows, for example, that distancing an event may increase or decrease one's confidence in predicting the event depending on whether the event is more likely or

less likely under high-level construals than low-level construals. Correspondingly, distancing an event may increase or decrease its attractiveness depending on whether the high-level value of the event is more positive or less positive than its low-level value. Again, these effects appear to hold across different distance dimensions.

The present review suggests, then, that (1) different distance dimensions are related to each other, (2) distancing on any of these dimensions is associated with higher levels of construal, and (3) they are, at least to some extent, interchangeable in their effects on prediction, evaluation, and choice. These three sets of findings suggest that psychological distance, as conceptualized here, may capture a fundamental aspect of meaning and may provide a unifying framework for understanding a wide range of seemingly unrelated social psychological phenomena.

REFERENCES

Adams, J. L. (1986). *Conceptual blockbusting: A guide to better ideas* (3rd ed.). Reading, MA: Addison-Wesley.

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82,* 463–496.

Ainslie, G., & Haslam, N. (1992). Hyperbolic discounting. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 57–92). New York: Russell Sage.

Argyle, M. (1970). *Social interaction.* Oxford, UK: Atherton Press.

Arnheim, R. (1969). *Visual thinking*. Berkeley: University of California Press.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359–372.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147.

Bar-Anan, Y., Liberman, N., & Trope, Y. (in press). The association between psychological distance and construal level: Evidence from an Implicit Association Test. *Journal of Experimental Psychology: General*.

Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2005). *The automatic processing of psychological distance: Evidence from a Stoop Task*. Manuscript submitted for publication.

Bartlett, F. A. (1932). *A study in experimental and social psychology*. New York: Cambridge University Press.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.

Beaman, A. L., Klentz, B., Diener, E., & Svanum, S. (1979). Objective self awareness and transgression in children: A field study. *Journal of Personality and Social Psychology*, *37*, 1835–1846.

Berry, J. W. (1976). *Human ecology and cognitive style: Comparative studies in cultural and psychological adaptation*. New York: Sage.

Borovoy, L., Liberman, N., & Trope, Y. (2004). *Processing near and distant sets of choice alternatives*. Unpublished manuscript, Tel Aviv University.

Brewer, M. B., & Lui, L. (1984). Categorization of the elderly by the elderly: Effects of perceiver's category membership. *Personality and Social Psychology Bulletin*, *10*, 585–595.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.

Bryant, F. B., & Guilbault, R. L. (2002). "I knew it all along" eventually: The development of hindsight bias in reaction to the Clinton impeachment verdict. *Basic and Applied Social Psychology*, *24*, 27–41.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*, 366–381.

Burger, J. M. (1986). Temporal effects on attributions: Actor and observer differences. *Social Cognition, 4*, 377–387.

Butler. D. L. & Kline, M. A. (1998). Good versus creative solutions: A comparison of brainstorming, hierarchical, and perspective changing heuristics. *Creativity Research Journal, 11*, 325–331.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Carver, C. S., & Scheier, M. F. (2000). Autonomy and self regulation. *Psychological Inquiry*, *11*, 284–291.

Choi, I, Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*(1), 47–63.

Christensen-Szalanski, J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*, 147–168.

Davis, M. H., Hull, J. G., Young, R. D., & Warren, G. G. (1987). Emotional reactions to dramatic film stimuli: The influence of cognitive and emotional empathy. *Journal of Personality and Social Psychology*, *52*, 126–133.

de Bono, E. (1985). The CoRT thinking program. In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.), *Thinking and learning skills: Volume 1. Relating instruction to research* (pp. 363–388). Hillsdale, NJ: Erlbaum.

Dunn, E. W., Wilson, T. D., & Gilbert, D. T. (2003). Location, location, location: The misprediction of satisfaction in housing lotteries. *Personality and Social Psychology Bulletin*, *29*, 1421–1432.

Dunning, D., Griffin, D. W., Milojkovic, J., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, *58*, 568–581.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self awareness*. New York: Academic Press.

Edwards, W. (1955). The prediction of decisions among bets. *Journal of Experimental Psychology*, *50*, 201–214.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*, 550–553.

Elster, J. (1977). *Ulysses and the sirens*. Cambridge, UK: Cambridge University Press.

Elster, J., & Loewenstein, G. (1992). Utility from memory and anticipation. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 213-234). New York: Russell Sage.

Eyal, T., Liberman, N., Sagristano, M., & Trope, Y. (2006). *Time dependent effects of primary vs. secondary values on behavioral intentions*. Unpublished manuscript, Tel Aviv University.

Eyal, T., Liberman, N., Trope, Y., & Walther, E. (2004). The pros and cons of temporally near and distant action. *Journal of Personality and Social Psychology*, *86*, 781–795.

Feather, N. T. (1995). Values, valences, and choice: The influence of values on the perceived attractiveness and choice of alternatives. *Journal of Personality and Social Psychology*, *68*, 1135–1151.

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*, 522–527.

Festinger, L. (1951). Architecture and group membership. *Journal of Social Issues*, *7*, 152–163.

Festinger, L., Schachter, S., & Black, K. W. (1950). *Social pressures in informal groups: A study of human factors in housing*. New York: Harper.

Fiedler, K., Semin, G. R., Finkenauer, C., & Berkel, I. (1995). Actorobserver bias in close relationships: The role of self-knowledge and self-related language. *Personality and Social Psychology Bulletin*, *21*, 525–538.

Finke, R. A. (1995). Creative insight and privative forms. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 255–280). Cambridge, MA: MIT Press.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*, 155–194.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Förster, J., Friedman, R. S., & Liberman, N. (2004). Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology*, *87*, 177–189.

Frank, M. G., & Gilovich, T. (1989). Effect of memory perspective on retrospective causal attributions. *Journal of Personality and Social Psychology*, *5*, 399–403.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2003). Time discounting and time preference: A critical review. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 13-86). New York: Russell Sage.

Freitas, A. L., Salovey, P., & Liberman, N. (2001). Abstract and concrete self-evaluative goals. *Journal of Personality and Social Psychology*, *80*, 410–412.

Friedman, R. S., & Förster, J. (2000). The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, *79*, 477–492.

Friedman, R. S., & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology*, *38*, 41–55.

Friedman, R. S., & Förster, J. (2002). The effect of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology*, *38*, 41–55.

Fujita, K. F., Henderson, M. D., Eng, J., Trope, Y., & Liberman, N. (2006). Spatial distance and mental construal of social events. *Psychological Science*, *17*, 278–282.

Fujita, K. F., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, *90*, 351–367.

Funder, D. C., & Van Ness, M. J. (1983). On the nature and accu-

racy of attributions that change over time. *Journal of Personality*, *51*, 17–33.

Gibbons, F. X. (1990). Self-attention and behavior: A review and theoretical update. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 249–303). New York: Academic Press.

Gilbert, D. T. (1998). Ordinary Personology, In D. T. Gilbert, S. T. Fiske, & L. Gardner (Eds.), *The handbook of social psychology* (Vol. 2, pp. 89–150). New York: McGraw-Hill.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science*, *15*, 346–350.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*, 617–638.

Gilbert, D. T., & Wilson, T. D. (2000). Miswanting: Problems in affective forecasting. In J. Forgas (Ed.), *Thinking and feeling: The role of affect in social cognition* (pp. 178–197). New York: Cambridge University Press.

Gilovich, T., & Medvec, V. H. (1995). The experience of regret: What, when, and why. *Psychological Review*, *102*, 379–395.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford, UK: Oxford University Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about self and others. *Journal of Personality and Social Psychology*, *59*, 1128–1139.

Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: Evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology*, *82*, 708–721.

Guinote, A., & Trope, Y. (2004). *Power differences in responsiveness to situational affordances*. Manuscript submitted for publication.

Hampson, S. E., John, O. P., & Goldberg, L. R. (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology*, *51*, 37–54.

Hayduk, L. A.(1983). Personal space: Where we now stand. *Psychological Bulletin*, *94*, 293–335.

Helfman, J. (1992). The analytic inventive thinking model. In R. J. Weber & D. N. Perkins (Eds.), *Inventive minds: Creativity in technology* (pp. 251–270). London: Oxford University Press.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, *5*, 184–200.

Hogg, M. A., & Reid, S. A. (2001). Social identity, leadership, and power. In A. Y. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power: Multiple perspectives on the causes of corruption* (pp. 159–180). Philadelphia: Psychology Press.

Holtgraves, T. (2002). *Language and social action: Social psychology and language use*. Mahwah, NJ: Erlbaum.

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self-others discrepancies in risk preference. *Journal of Experimental Psychology: General*, *126*, 45–53.

Hsee, C. K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology*, *86*, 680–695.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion–cognition relations. *Psychological Review*, *99*, 561–565.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371–376.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107–117.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Jones, E. E., Wood, G. C., & Quattrone, G. A. (1981). Perceived variability of personal characteristics in ingroups and outgroups: The role of knowledge and evaluation. *Personality and Social Psychology Bulletin*, *7*, 523–528.

Kahneman, D., & Lovallo, D. (1991). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*, 17–31.

Kahneman, D., & Snell, J. (1990). Predicting utility. In R. Hogarth (Ed.), *Insight in decision making: A tribute to Hillel J. Einhorn* (pp. 295–310). Chicago: University of Chicago Press.

Kahneman, D., & Snell, J. (1992). Predicting a change in taste: Do people know what they will like? *Journal of Behavioral Decision Making*, *5*, 187–200.

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *Management Science*, *12*, 313–327.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*, 265–284.

Keren, G., & Roelofsma, P. (1995). Immediacy and certainty in intertemporal choice. *Organizational Behavior and Human Decision Processes*, *63*, 287–297.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, *110*, 499–519.

Kogut, S., & Ritov, I. (2004). *The singularity effect of identified victims in separate and joint evaluations*. Manuscript submitted for publication, Hebrew University, Jerusalem.

Kray, L. (2000). Contingent weighting in self-other decision making. *Organizational Behavior and Human Decision Processes*, *83*, 82–106.

Kray, L., & Gonzalez, R. (1999). Differential weighting in choice versus advise: I'll do thins, you do that. *Journal of Behavioral Decision Making*, *12*, 201–217.

Kross, E., Ayduk, O., & Mischel, W. (2004). When asking "*why*" does not hurt: Distinguishing rumination from reflective processing of negative emotions. *Psychological Science*, *16*, 709–715.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*, 343–356.

Latané, B., Liu, J. H., Nowak, A., & Bonevento, M. (1995). Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin*, *21*, 795–805.

Lee, F., & Tiedens, L. Z. (2001). Who's being served? "Self-serving" attributions in social hierarchies. *Organizational Behavior and Human Decision Processes*, *84*, 254–287.

Levi Strauss, C. (1978). *Myth and meaning: Cracking the code of culture*. Toronto: University of Toronto Press.

Libby, L. K., & Eibach, R. P. (2002). Looking back in time: Self-concept change affects visual perspective in autobiographical memory. *Journal of Personality and Social Psychology*, *82*, 167–179.

Liberman, N., Sagristano, M., & Trope, Y. (2002). The effect of temporal distance on level of construal. *Journal of Experimental Social Psychology*, *38*, 523–535.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*, 5–18.

Liberman, N., & Trope, Y. (2003). Temporal construal theory of intertemporal judgment and decision. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 245–276). New York: Russell Sage.

Liberman, N., Trope, Y., McCrae, S. M., & Sherman, S. J. (in press).

The effect of level of construal on the temporal distance of activity enactment. *Journal of Experimental Social Psychology.*

Linville, P. W. (1982). The complexity–extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology*, *42*, 193–211.

Linville, P. W. (1985). Self complexity and affective extremity: Don't put all of your eggs in one cognitive basket. *Social Cognition*, *3*, 94–120.

Linville, P. W., Fischer, G. W., & Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. *Journal of Personality and Social Psychology*, *70*, 421–436.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance.* Upper Saddle River, NJ: Prentice-Hall.

Loewenstein, G. F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*, 272–292.

Loewenstein, G., & Prelec, D. (1992). Anomalies in interpersonal choice: Evidence and an interpretation. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 119–145). New York: Russell Sage.

Lopes, L. L. (1996). When time is of the essence: Averaging, aspiration, and the short run. *Organizational Behavior and Human Decision Processes*, *65*, 179–189.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1998). Saying no to unwanted thoughts: Self-focus and the regulation of mental life. *Journal of Personality and Social Psychology*, *74*, 578–589.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on rebound. *Journal of Personality and Social Psychology*, *67*, 808–817.

Markus, H. R., & Kitayama, S. (1994). The cultural shaping of emotion: A conceptual framework. In S. Kitayama & H. R. Markus (Eds.), *Emotion and culture: Empirical studies of mutual influence* (pp. 339–351). Washington, DC: American Psychological Association Press.

Marsh, R. L., Ward, T. B., & Landau, J. D. (1999). The inadvertent use of prior knowledge in a generative cognitive task. *Memory and Cognition*, *27*, 94–105.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McDonald, H. E., & Hirt, E. R. (1997). When expectancy meets desire: Motivational effects in reconstructive memory. *Journal of Personality and Social Psychology* *72*, 5–23.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations*, *18*, 57–76.

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*, 961–978.

Miller, N. E., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Adams-Bennister-Cox.

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249–292). New York: Academic Press.

Mischel, W., Ayduk, O., & Mendoza-Denton, R. (2003). Sustaining delay of gratification over time: A hot-cool systems perspective. In G. Loewenstein & D. Read (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 175–200). New York: Russell Sage.

Mischel, W., Grusec, J., & Masters, J. C. (1969). Effects of expected delay time on the subjective value of rewards and punishments. *Journal of Personality and Social Psychology*, *11*, 363–373.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989, May 26). Delay of gratification in children. *Science*, *244*, 933–938.

Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997).

Temporal adjustments in the evaluation of events: The "rosy view." *Journal of Experimental Social Psychology*, *33*, 421–448.

Mooney, K. M., Cohn, E. S., & Swift, M. B. (1992). Physical distance and AIDS: Too close for comfort? *Journal of Applied Social Psychology*, *22*, 1442–1452.

Moore, B. S., Sherrod, D. R., Liu, T. J., & Underwood, B. (1979). The dispositional shift in attribution over time. *Journal of Experimental Social Psychology*, *15*, 553–569.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949–971.

Newcomb, T. M. (1961). *The acquaintance process.* New York: Holt, Rinehart and Winston.

Nigro, G., & Neisser, U. (1983). Point of view in personal memories. *Cognitive Psychology*, *15*, 467–482.

Nisbett, R. E. (1993). Violence and U.S. regional culture. *American Psychologist*, *48*, 441–449.

Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, *27*, 154–164.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*, 291–310.

Nussbaum, S., Liberman, N., & Trope, Y. (2004). *Predicting the near and distant future.* Manuscript submitted for publication.

Nussbaum, S., Trope, Y., & Liberman, N. (2003). Creeping dispositionism: The temporal dynamics of behavior prediction. *Journal of Personality and Social Psychology*, *84*, 485–497.

O'Donoghue, T., & Rabin, M. (2000). The economics of immediate gratification. *Journal of Behavioral Decision Making*, *13*, 233–250.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*, 315–331.

Overbeck, J. R., & Park, B. (2001). When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology*, *81*, 549–565.

Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology*, *59*, 173–191.

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, *42*, 1051–1068.

Park, B., Ryan, C. S., & Judd, C. M. (1992). Role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *Journal of Personality and Social Psychology*, *63*, 553–567.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 534–552.

Pease, M. E., McCabe, A. E., Brannon, L. A., & Tagler, M. J. (2003). Memory distortions for pre-Y2K expectancies: A demonstration of hindsight bias. *Journal of Psychology: Interdisciplinary and Applied*, *137*, 397–399.

Peterson, C. (1980). Memory and the "dispositional shift." *Social Psychology Quarterly*, *43*, 372–380.

Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science*, *37*, 770–786.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 1–72.

Priest, R. F., & Sawyer, J. (1967). Proximity and peership: Bases of balance in interpersonal attraction. *American Journal of Sociology*, *72*, 633–649.

Pronin, E., Kruger, J., Savtisky, K., & Ross, L. (2001). You don't know me, but I know you: The illusion of asymmetric insight. *Journal of Personality and Social Psychology*, *81*, 639–656.

Rachlin, H., Brown, J., & Cross, D. (2000). Discounting in judgments of delay and probability. *Journal of Behavioral Decision Making*, *13*, 145–159.

Read, D., & Loewenstein, G. (2000). Time and decision: Introduc-

tion to the special issue. *Journal of Behavioral Decision Making*, *13*, 141–144.

Reichman, P., & Ben Arie, Y. (2004). *The effect of spatial distance on politeness: Evidence for the effect of special distance on social distance*. Unpublished manuscript, Tel Aviv University.

Robins, R. W., Spranca, M. D., & Mendelsohn, G. A. (1996). The actor-observer effect revisited: Effects of individual differences and repeated social interactions on actor and observer attributions. *Journal of Personality and Social Psychology*, *71*, 375–389.

Rohan, M. J. (2000). A rose by any name?: The values construct. *Personality and Social Psychology Review*, *4*, 255–277.

Rosch, E., & Lloyd, B. B. (Eds.). (1978). *Cognition and categorization*. Oxford, UK: Erlbaum.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied 80*, 1–28.

Sagristano, M. D., Trope, Y., Eyal, T., & Liberman, N. (2006). *How temporal distance affects attitude–behavior correspondence*. Manuscript in preparation, New York University.

Sagristano, M. D., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: Odds now, money later. *Journal of Experimental Psychology: General*, *131*, 364–376.

Saulnier, K., & Perlman, D. (1981). The actor- observer bias is alive and well in prison: A sequel to wells. *Personality and Social Psychology Bulletin*, *7*, 559–564.

Schelling, T. (1984). Self command in practice, in theory and in a theory of rational choice. *American Economic Review*, *74*, 1–11.

Schiffenbauer, A., & Schiavo, R. S. (1976). Physical distance and attraction: An intensification effect. *Journal of Experimental Social Psychology*, *12*, 274–282.

Schoppe, U. (1975). *Verbaler Kreativitatstest (verbal creativity test)*. Gottingen, Germany: Hogrefe.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York: Academic Press.

Schwartz, S., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, *53*, 550–562.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558–568.

Semin, G. R., & Fiedler, K. (1989). Relocating attributional phenomena within language- cognition interface: The case of actors' and observers' perspectives. *European Journal of Social Psychology*, *19*, 491–508.

Semin, G. R., & Smith, E. R. (1999). Revisiting the past and back to the future: Memory systems and the linguistic representation of social events. *Journal of Personality and Social Psychology*, *76*, 877–892.

Sherman, S. J., Beike, D. R., & Ryalls, K. R. (1999). Dual-processing accounts of inconsistencies in responses to general versus specific cases. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 203–227). New York: Guilford Press.

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, *11*, 118–127.

Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 391–445). New York: McGraw-Hill.

Smith, P., & Trope, Y. (2006). You focus on the forest when you're in charge of the trees: The effect of power priming on information processing. *Journal of Personality and Social Psychology*, *90*, 578–596.

Snyder, C. R., & Fromkin, H. L. (1980). *Uniqueness: The human pursuit of difference*. New York: Plenum Press.

Stephan, E. (2005). *Social distance and its relation to level of construal, temporal distance and physical distance*. Unpublished manuscript, Tel Aviv University.

Stern, E. R., & Rotello, C. M. (2000). Memory characteristics of recently imagined events and real events experienced previously. *American Journal of Psychology, 113*, 569–590.

Stroop, J. R. (1935). Studies of interference in serial–verbal reaction. *Journal of Experimental Psychology*, *18*, 643–662.

Suengas, A. G., & Johnson, M. K. (1998). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General 117*, 377–389.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. Oxford, UK: Wiley.

Todorov, A., Goren, A., & Trope, Y. (in press). Probability as a psychological distance: Construal and preference. *Journal of Experimental Social Psychology*.

Trope, Y. (1989). Levels of inference in dispositional judgment. *Social Cognition, 7*, 296–314.

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology*, *79*, 876–889.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–421.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3–15.

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology, 57*, 660–671.

Ward, T. B. (1995). What's old about new ideas? In M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 157–178). Cambridge, MA: MIT Press.

Watson, D. (1982). The actor and the observer: How are their perceptions of causality different? *Psychological Bulletin*, *92*, 682–700.

Werkman, W. M., Wigboldus, D. H., & Semin, G. R. (1999). Children's communication of the linguistic intergroup bias and its impact upon cognitive inferences. *European Journal of Social Psychology*, *29*, 95–104.

Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 233–275). New York: Academic Press.

Wilson, T. D., Meyers, J., & Gilbert, D. T. (2001). Lessons from the past: Do people learn from experience that emotional reactions are short-lived? *Personality and Social Psychology Bulletin*, *27*, 1648–1661.

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*, 821–836.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *Group Embedded Figures Test manual*. Palo Alto, CA: Consulting Psychologists Press.

Woodruff, D. S., & Birren, J. E. (1972). Age changes and cohort difference in personality. *Developmental Psychology*, *6*, 252–259.

Worthington, M. R. (1974). Personal space as a function of stigma effect. *Environment and Behavior*, *6*, 289–294.

PART III

PERSONAL MOTIVATIONAL SYSTEM

CHAPTER 16

Feelings and Phenomenal Experiences

NORBERT SCHWARZ
GERALD L. CLORE

Following an initial emphasis on "cold" cognitive processes, which could be conceptualized within the computer metaphor of the information-processing paradigm, social cognition researchers rediscovered "hot" cognition in the 1980s. Two decades later, their interest in the interplay of feeling and thinking is shared by researchers in decision making, cognitive psychology, and related fields. This chapter reviews what has been learned; it focuses on basic theoretical principles and empirical regularities rather than complete coverage of the literature. We first introduce three broad approaches to the interface of feeling and thinking and subsequently evaluate them in light of empirical findings in three key domains, namely, human judgment, strategies of information processing, and memory. Throughout, we emphasize the influence of feelings on cognitive processes; the reverse influence of cognition on emotion is reviewed by Clore, Schwarz, and Conway (1994) and Ellsworth and Scherer (2003).

APPROACHES TO FEELING AND THINKING

Three general approaches to the role of feelings in human cognition focus on the experiential, cognitive, and somatic components of feelings, respectively. The first approach emphasizes the experiential quality of feelings and addresses their informational functions. A second approach emphasizes the thoughts that accompany feelings, whereas a third approach emphasizes hardwired processes, focusing on the somatic components of affective states.

The Experiential Component of Feelings: Feelings as a Source of Information

Central to the experiential approach is the assumption that feelings can serve as a source of information in their own right. This assumption is consistent with traditional theorizing on emotions and has been fruitfully extended to other subjective experiences.

Affective, Bodily, and Cognitive Experiences

Social psychologists often subsume moods and emotions under the generic term "affect." This term, however, can also refer simply to valence—the positive and negative aspect of things. All emotions are affective, but not all affective things are emotions. *Emotions* arise in response to ongoing, implicit appraisals of situations with respect to positive or negative implications for one's goals and concerns (e.g., Ellsworth & Scherer, 2003; Ortony, Clore, & Collins, 1988). They have an identifiable referent (what the emotion is "about"), a sharp rise time, limited duration, and often high intensity. Emotion researchers commonly assume that "emotions exist for the sake of signaling states of the world that have to be responded to, or that no longer need response and action" (Frijda, 1988, p. 354). What exactly emotions signal can be derived from their underlying appraisal patterns. Sadness, for example, signals a loss or lack of reward that is not attributed to the causal action of another agent; when it is attributed to the causal action of another agent, it gives rise to anger. Accordingly, sadness and anger inform us about not only a loss but also about its likely cause, giving

rise to different attributions in judgment studies (e.g., Keltner, Ellsworth, & Edwards, 1993).

Moods, on the other hand, lack a clear referent, may come about gradually, may last for an extended time, and are often of low intensity (Morris, 1989). The experience of a positive or negative emotion may also leave us in a positive or negative mood after the emotion dissipates and its specific cause is no longer attended to (Bollnow, 1956). This difference between moods and emotions is apparent in ordinary language when we say that we are angry "about" something but that we are "in" a bad mood. Hence, moods mostly convey generic valence information that has no clear referent, which accounts for their pervasive influence.

Bodily experiences include feelings such as hunger, pain, and physiological arousal, which inform us about specific states of the organism. These experiences can be informative in their own right, and induced physical arousal (e.g., Zillman, 1978) or proprioceptive feedback from facial expressions (e.g., Strack, Martin, & Stepper, 1988) and arm flexion and extension (Friedman & Förster, 2000) can convey information that influences judgment and information processing.

Finally, *cognitive experiences* such as surprise, amazement, boredom, or feelings of familiarity inform us about knowledge states. Two cognitive experiences that received particular attention are the metacognitive experiences of accessibility and processing fluency. *Accessibility experiences* refer to the ease or difficulty with which information can be brought to mind or thoughts can be generated. They can serve as input into a large variety of judgments; their specific impact is highly malleable and depends on which naive theory of mental processes is brought to bear on the task (Schwarz, 2004).

Processing fluency refers to the ease or difficulty with which new, external information can be processed. Variables like figure–ground contrast, presentation duration, or the amount of previous exposure to the stimulus affect the speed and accuracy of low-level processes concerned with the identification of a stimulus's physical identity and form; they influence *perceptual fluency* (e.g., Jacoby, Kelley, & Dywan, 1989). Variables such as the consistency between the stimulus and its context or the availability of appropriate mental concepts for stimulus classification affect the speed and accuracy of high-level processes concerned with the identification of stimulus meaning and its relation to semantic knowledge structures; they influence *conceptual fluency* (e.g., Whittlesea, 1993). Empirically, both types of fluency show parallel influences on judgment (for a review, see Winkielman, Schwarz, Fazendeiro, & Reber, 2003) and can be subsumed under the general term "processing fluency." Which inferences people draw from experienced processing fluency again depends on the naive theory of mental processes that they bring to bear on the task (Schwarz, 2004). In addition, high fluency is experienced as hedonically positive (as captured by psychophysiological measures; Winkielman & Cacioppo, 2001), and this affective response can itself serve as a basis of judgment (Winkielman et al., 2003).

As our review indicates, all these experiences can influence how we evaluate a stimulus and which strategy of information processing we adopt; they also receive increasing attention in the study of memory.

Judgment

Central to the feelings-as-information approach is the assumption that people draw on their affective, cognitive, and bodily experiences as a source of information. In the case of moods and emotions, people may use their apparent affective response to a target as information in forming an evaluative judgment (e.g., Schwarz & Clore, 1983; Wyer & Carlston, 1979). This possibility is most obvious when the judgment refers, by definition, to one's affective reaction to the stimulus. For example, when asked how much we like a person, we may base the judgment on our feelings toward the person rather than on a review of his or her attributes. Also, when a judgment does not refer directly to our feelings but poses a task that is particularly complex and demanding, we may simplify the task by asking ourselves, "How do I feel about it?" (Schwarz & Clore, 1988). When the apparent affective response is indeed elicited by the target, it provides meaningful information that is relevant to the judgment at hand. Pham, Cohen, Pracejus, and Hughes (2001) observed that people can monitor and report the feelings elicited by moderately complex targets, like pictures or advertisements, very rapidly—in fact, more rapidly than their cognitive responses, consistent with Zajonc's (1980) affective primacy hypothesis.

Because we have only one window on our experience, it is difficult to distinguish *integral* feelings, elicited by the target, from *incidental* feelings that happen to be present at the time. Hence, we may mistake incidental feelings, like a preexisting mood, as part of our reaction to the target. This results in judgments that are congruent with the implications of our feelings, most notably in more positive evaluations under happy rather than sad moods. To disentangle the contributions of the perceiver's feelings from other information about the target, experimental tests of the feelings-as-information hypothesis usually rely on the induction of incidental affect. This gave rise to the erroneous conclusion that "affect can only serve as a heuristic cue due to mistaken inferences," suggesting that reliance on one's feelings "is an ineffective and dysfunctional strategy" (Forgas, 2001, p. 104). This assertion confuses the operational and theoretical level and is at odds with a long tradition of theorizing that emphasizes the signaling functions of affective responses (see Frijda, 1988; Zajonc, 1980). Feelings can serve as a basis of accurate as well as mistaken inferences, depending on the relationship between the feeling and the target. In fact, a growing body of work (see Damasio, 1994; Feldman-Barrett & Salovey, 2002) indicates the adaptive value of attending to one's feelings in judgment and decision making.

However, our feelings only influence judgment when they seem relevant to the task at hand. Hence, their influence is *eliminated* when we, correctly or incorrectly, as-

sume that they are indeed incidental, thus calling their informational value into question (e.g., Schwarz & Clore, 1983). Note that this proposition does *not* imply that mood effects on evaluative judgments require a conscious attribution of one's feelings to the target, in contrast to what some interpreters (e.g., Forgas, 1995a) suggested (for a discussion, see Schwarz, 2001). People usually consider their experiences, ranging from their feelings to the thoughts that come to mind, to be "about" whatever is in the focus of their attention. This observation has been termed the "aboutness" (Higgins, 1998) or "immediacy" principle (Clore et al., 2001); it is reminiscent of the Gestalt observation that stimuli presented in close temporal or spatial proximity are experienced as connected (Heider, 1958). Accordingly, reliance on our thoughts and feelings is the automatic default operation and does not require a conscious attribution, whereas discrediting thoughts and feelings does.

The differences between moods and emotions, noted earlier, suggest that people are more likely to be aware of the cause of their emotions than of the cause of their diffuse moods; this limits the likelihood that they misread their emotions as a response to an unrelated target. Moreover, emotions reflect specific appraisal patterns and hence provide more specific information than global moods, which mostly indicate valence. Finally, the same basic logic applies to bodily sensations as well as accessibility and fluency experiences. As our chapter indicates, people only draw on these experiences as a source of information when their informational value is not discredited.

Processing Style

Moods, emotions, and bodily sensations have also been found to influence people's processing strategies. From the feelings-as-information perspective, these experiences inform us about the benign or problematic nature of the current situation. This, in turn, influences which processing strategy we adopt, consistent with the assumption that human cognition is situated and adaptively tuned to meet situational requirements (see Schwarz, 2002).

Feelings that signal a problematic environment foster systematic, bottom-up processing with considerable attention to detail. As Wegner and Vallacher (1986) noted, this style of reasoning is adaptive when we encounter problems in the pursuit of our goals. We also adopt it when bodily avoidance feedback (e.g., Friedman & Förster, 2000) or emotions that entail a high uncertainty appraisal (e.g., Tiedens & Linton, 2001), correctly or incorrectly, provide an experiential "problem" signal. Conversely, feelings that signal a benign environment are typically associated with more heuristic processing and increased reliance on the top-down use of preexisting knowledge structures, unless otherwise required by the task at hand (e.g., Bless et al., 1996). As predicted by the feelings-as-information approach, these differences in processing style are eliminated when the informational value of the feeling is called into question through attribution manipulations (e.g., Sinclair, Mark, & Clore, 1994).

Empirically, influences of feelings on processing style have been observed across a wide range of tasks, including problem solving (for a review, see Schwarz & Skurnik, 2003), stereotyping (for a review, see Bless, Schwarz, & Kemmelmeier, 1996), and persuasion (for a review, see Schwarz, Bless, & Bohner, 1991). We address distinctions between several related conceptual models in our review of the evidence.

Memory

Experiential considerations played a key role in early treatments of memory processes but have lost popularity and were rediscovered only recently (Brewer, 1992). The accumulating work (for a review, see Kelley & Rhodes, 2002) documents a pervasive role of feelings of familiarity, which arise from processing fluency. For example, Jacoby and Dallas (1981) observed that participants could accurately recognize previously shown rare words but provided numerous false alarms in response to common words. Apparently, they misattributed the familiarity resulting from their frequent exposure to common words to the recency of exposure, erroneously concluding that the word was part of the preceding learning task. Such findings parallel the role of processing fluency in judgment. However, other memorial activities, like retrieving facts from long-term memory, do not involve phenomenal experiences (Brewer, 1992).

Whereas the role of cognitive experiences in memory is increasingly well understood, the influence of moods and emotions on memory has rarely been addressed from an experiential perspective. One approach to this issue (e.g., Bless, 1996) assumes that mood-congruent recall may arise from initial mood effects on evaluative judgment. When asked to recall events from our kindergarten days, for example, we may first wonder what they were like. When in a good mood, we may arrive at a more pleasant assessment, which may then serve as input into reconstructive memory processes (Ross, 1989), resulting in the "recall" of more pleasant events.

The Cognitive Component of Feelings: What Comes to Mind

An alternative approach traces the influence of feelings to the thoughts that accompany moods and emotions. In a pioneering series of studies, Isen, Shalker, Clark, and Karp (1978) observed pronounced mood effects on evaluative judgments. To account for them, they suggested a "cognitive loop" in the form of higher accessibility of mood-congruent information stored in memory. At the same time, Bower (1981; Bower, Monteiro, & Gilligan, 1978) conceptualized the operation of affective states in a general network model of human memory. He represented emotions as central nodes in an associative network, which are linked to related ideas, events of corresponding valence, autonomic activity, and muscular and expressive patterns. When new material is learned, it

is associated with the nodes that are active at the time of learning, including the respective emotion nodes. When an emotion node is stimulated later on, activation spreads along the pathways, increasing the activation of other, connected nodes; activation of a node above a certain threshold brings the represented material into consciousness.

Memory

Bower's (1981) model and related conceptualizations (e.g., Isen et al., 1978; Spies & Hesse, 1986; Wyer & Srull, 1989) generate two key predictions. A *state-dependent learning and recall* hypothesis holds that material learned in one affective or bodily state is more likely to be recalled when we are in the same rather than another state. Whereas this prediction pertains to matching states at learning and recall, a *mood-congruency hypothesis* pertains to matching valences of the affective state and the to-be-recalled material. It holds that positively (negatively) valenced material is more likely to be recalled in positive (negative) moods. Although these hypotheses are conceptually distinct, they are often difficult to distinguish empirically (Morris, 1989). In fact, mood-congruent recall has been most reliably observed in the domain of autobiographical memory, with happy events being more likely to be recalled under happy than sad moods (see Blaney, 1986, for a review). However, happy events are likely to have put one into a happy mood at the time they occurred. Accordingly, mood congruency in autobiographical recall could be due to either state dependency or mood congruency.

Critics of Bower's (1981) model questioned the plausibility of the assumed links between positive (negative) moods and *all* positive (negative) material stored in memory; this structure should reduce the activation of any given node as a function of the overall number of nodes activated. Moreover, an extensive review of relevant experiments (Wyer, Clore, & Isbell, 1999) concluded that reported effects of mood on the recall of valence-congruent semantic concepts in memory may be due to the inadvertent activation of mood-relevant concepts as part of the mood-induction procedure, rather than to mood per se.

Judgment

If some aspect of mood does affect the recall of valenced material, it should also affect tasks to which the material may be relevant. Accordingly, a *mood-congruent encoding* hypothesis suggests that ambiguous material would be encoded in terms of mood-congruent concepts. Moreover, associations that come to mind subsequently (regardless of the ambiguity of the material) may also be mood congruent, resulting in *mood-congruent elaboration*. By the same token, mood-congruent recall of infomation about the target is assumed to result in *mood-congruent judgments.* Finally, any impact of mood-congruent recall and elaboration is predicted to be more pronounced the more the judgment involves substantive thought (Forgas,

1995a), that is, under conditions of high processing motivation and low time pressure.

Processing Style

Mood-congruent recall assumptions can also be used to predict affective influences on styles of information processing. If a given affective state brings a large amount of congruent information into consciousness, it may limit the *cognitive resources* required for working on other tasks. However, it is unclear which mood is most likely to constrain resources. On the one hand, negative affective states are associated with intruding thoughts and ruminations (e.g., Martin & Tesser, 1989) and people in an induced or chronic negative mood have difficulty suppressing mood-congruent material when instructed to do so (e.g., Wenzlaff, Wegner, & Roper, 1988). Hence, Ellis and Ashbrook (1988) suggested in their resource allocation model that *negative* affective states interfere with information processing. On the other hand, Isen (1987) hypothesized that positive material in memory "is more extensive and at the same time better integrated, so that positive affect is able to cue a wide range of thoughts" (p. 217). Accordingly, Mackie and Worth (1989) proposed that *positive* affective states can limit attentional resources due to intruding positive thoughts. Neither proposal can account for the bulk of the available findings, as reviewed below.

Finally, we note that models focusing on *what* comes to mind do not provide an easy way to conceptualize the role of cognitive experiences. Although they can account for affective influences on ease of recall and fluency of processing, the judgment effects of experienced accessibility and fluency cannot be conceptualized without making assumptions extraneous to associative network models.

The Somatic Component of Feelings: A Hard Interface?

The somatic component of feelings has been addressed in two different ways. As Zajonc and Markus (1984) noted, most theories postulate some form of experiential mediation between somatic processes and judgments or other outcomes (e.g., Izard, 1977; Leventhal, 1982; Schachter & Singer, 1962; Tomkins, 1962). Hence, we treat such theories in the context of experiential approaches. As an alternative, Zajonc and Markus suggested that the impact of somatic processes may be hardwired instead of experientially mediated. For example, Zajonc and collaborators emphasized possible representational functions of the motor system (e.g., Adelman & Zajonc, 1989) and suggested a crucial role for the vascular system of the head in emotion regulation (Zajonc, Murphy, & Inglehart, 1989). Much of this work focuses on somatic processes as determinants of emotion and little is known about their role in mediating the consequences of emotions. Reflecting the paucity of research that bears on this mediational issue, and the limits of our expertise, we do not address this perspective in this chap-

ter, which focuses on feelings (i.e., subjective experiences).

FEELINGS AND JUDGMENT

Next, we review research on the impact of different feelings on social judgment and highlight similarities in the informative functions of affective, cognitive, and bodily experiences. Where applicable, we contrast predictions derived from the general approaches discussed previously.

Moods

All models predict more positive judgments under happy than sad moods. This prediction is well supported, with exceptions addressed below. From a feelings-as-information perspective (Schwarz & Clore, 1983, 1988; Wyer & Carlston, 1979), mood-congruent judgments arise because people misread incidental moods as part of their apparent affective reaction to the target; from a mood-congruent recall perspective (Bower, 1981), they arise because moods bring different associations to mind. These perspectives lead to different predictions about the conditions under which mood-congruent judgment should be observed.

Perceived Informational Value

If feelings serve as a source of information, their impact should depend on their perceived informational value. We should discount our feelings as a source of information when there is reason to assume that they may not reflect our reaction to the target but should see them as particularly informative when our apparent reaction to the target contradicts the plausible impact of other influences. Such discounting and augmentation effects (Kelley, 1972) cannot be derived from the assumption that moods or emotions affect judgment through the selective recall of information from memory (e.g., Bower, 1981) or by hardwired processes (e.g., Zajonc & Markus, 1984). These assumptions predict main effects of affective states, whereas the feelings-as-information hypothesis predicts an interaction between affective states and the perception of their likely causes.

Supporting this interaction prediction, Schwarz and Clore (1983) observed that the influence of mood on judgments of life satisfaction was eliminated when participants attributed their current feelings either correctly or incorrectly to a transient source. For example, participants reported higher life satisfaction and a more elated current mood in telephone interviews when called on sunny rather than rainy days. This difference was eliminated when the interviewer mentioned the weather as part of a private aside, thus directing participants' attention to this source of their feelings. Similarly, recalling a sad life event did not influence participants' judgments of life satisfaction when they could misattribute the resulting sad feelings to the alleged impact of the experi-

mental room. In addition, current mood, as assessed at the end of each experiment, was more strongly correlated with judgments of life satisfaction when participants' attention was *not* directed to a transient source of their feelings than when it was. Conceptual replications of these attributional effects have been reported by Gorn, Goldberg, and Basu (1993), Keltner, Locke, and Audrain (1993), Savitsky, Medvec, Charlton, and Gilovich (1998), Schwarz, Servay, and Kumpf (1985), and Siemer and Reisenzein (1998), among others.

Further highlighting that people assess the informational value of their feelings, Avnet and Pham (2004) manipulated participants' perception of the extent to which they can trust their feelings. As expected, induced moods influenced evaluative judgments more when participants were primed to trust rather than distrust their feelings. Finally, augmentation and discounting effects are more reliably obtained under the systematic processing strategy fostered by negative moods (Schwarz & Clore, 1983), which prompts more elaborate causal reasoning (Bohner, Bless, Schwarz, & Strack, 1988).

Perceived Relevance

Even when we perceive our feelings as informative, the information they provide may seem more germane to some judgments than to others. For example, Pham (1998) observed more pronounced mood effects when participants evaluated activities with experiential rather than instrumental goals in mind (see also Adaval, 2001; Yeung & Wyer, 2004). He concluded (Pham, 2004) that we consider our feelings more relevant "for assessing the potential fulfillment of experiential goals (e.g., 'Would I have fun at this movie?') than for assessing the potential fulfillment of instrumental goals (e.g., 'Would seeing this movie help me for the project?')" (p. 366). Similarly, we consider our feelings more relevant when judging our own preferences than when judging the preferences of others. Accordingly, Raghunathan and Pham (1999) found stronger mood effects when individuals made decisions for themselves rather than for others.

Finally, people are less likely to rely on their moods when they have high expertise in the domain of judgment (e.g., Ottati & Isbell, 1996; Sedikides, 1995). High expertise presumably facilitates the assessment of the relevance of one's feelings and renders other, relatively more diagnostic inputs easily accessible. Conversely, being under time pressure increases reliance on one's feelings (Siemer & Reisenzein, 1998), presumably because it interferes with relevance assessments and the search for alternative inputs.

In sum, the use of feelings as information follows the same principles as the use of other information (Feldman & Lynch, 1988): We only rely on them when their informational value is not discredited and when they seem relevant to the judgment at hand. Moreover, their impact decreases when other, more relevant, inputs are accessible.

Feeling versus Content

According to mood-congruent recall models, the impact of moods should increase when the mood induction is thematically related to the target judgment. Suppose, for example, that a depressed mood is induced by thoughts about a serious illness. If so, the valence of the mood *and* the content of the mood induction should facilitate the recall of illness-related material from memory (e.g., Bower, 1981). Hence, illness-related judgments should be more likely to show mood effects than judgments pertaining to other content domains. Empirically, this is not the case (e.g., Johnson & Tversky, 1983; Mayer, Gaschke, Braverman, & Evans, 1992).

For example, Johnson and Tversky (1983) observed that reading descriptions of negative events, which presumably induced a depressed and slightly anxious mood, increased judgments of risk. This effect was independent of the object of judgment or the content by which the mood was induced. Reading about cancer, for example, affected judgments of the risk of cancer but had equally strong effects on judgments of the risk of accidents and divorce. Such generalized effects, undiminished over dissimilar content domains, are incompatible with models of mood-congruent recall. However, they are consistent with the feelings-as-information hypothesis. From this perspective, participants inferred higher risk from their depressed and anxious feelings, thus simplifying an otherwise difficult task (for an extended discussion of the role of feelings in judgments of risk, see Loewenstein, Weber, Hsee, & Welch, 2001).

In addition, other researchers observed mood effects on evaluative judgments in the absence of any evidence for mood effects on the recall of relevant information from memory (e.g., Fiedler, Pampe, & Scherf, 1986). More fundamentally, a general review of the mood and memory literature (Wyer, et al., 1999) suggests that most instances of mood-congruent recall in the literature are ambiguous as to whether they result from mood per se or from the inadvertent activation of mood-relevant concepts. Most mood-induction procedures involve the activation of positive and negative concepts as well as of positive and negative feelings, raising the possibility that mood-congruent recall may often be concept-congruent recall.

Mood-Incongruent Judgments

When we ask ourselves, "How do I feel about it?," our happy or sad moods convey that we feel good or bad about whatever the "it" is. In most cases, this results in mood-congruent judgments, as in the foregoing examples. Nevertheless, mood-incongruent judgments can be observed under specific conditions.

First, suppose you are asked to read a sad story and to evaluate how successful it is at making you feel sad. If you find yourself happy after reading the story, you are likely to conclude that it was a poor sad story—a condition under which positive feelings result in a negative judgment. Empirically this is the case, as Martin, Abend, Sedikides, and Green (1997) demonstrated. In our read-ing, such findings reflect changes in the criterion of judgment rather than changes in the information conveyed by the mood.

Second, mood-incongruent judgments can result from the logic of discounting effects themselves (e.g., Isbell & Wyer, 1999; Ottati & Isbell, 1996). Suppose that you are evaluating a job candidate but are aware that you have been in a miserable mood all day due to an earlier event. How much of your negative affect reflects your reaction to the candidate and how much is due to the earlier event? If you fully attribute your bad feelings to the earlier event, you may arrive at an unduly positive assessment of the candidate. Empirically, attempts to correct for a perceived influence often result in overcorrections of this type (Strack & Hannover, 1996; Wilson & Brekke, 1994), unless the person can draw on other accessible inputs as an alternative route to judgment. Accordingly, (mis)attribution manipulations may eliminate mood effects when diagnostic alternative inputs are available but may foster overcorrection, and hence mood-incongruent judgments, when they are not.

Finally, mood-inducing events can elicit contrast effects in the evaluation of closely related targets by serving as extreme standards of comparison. For example, Schwarz, Strack, Kommer, and Wagner (1987) induced happy or sad moods by conducting an experiment in a very pleasant or unpleasant room. Consistent with the induced moods, their student participants reported higher life satisfaction in the pleasant than unpleasant room. When asked to report their housing satisfaction, however, this pattern reversed, presumably because even modest dorm rooms seemed luxurious compared to the salient standard introduced by the unpleasant room. Theoretically, such comparison-based contrast effects should be limited to judgments for which the mood-inducing event can serve as a highly relevant standard.

Specific Emotions

The general logic outlined for moods also applies to the use of specific emotions as a source of information. For example, Schwarz and colleagues (1985) observed that the impact of a fear-arousing communication on participants' attitudes was eliminated when participants attributed their subjective experience to the arousing side effects of a pill but was enhanced when they assumed the pill would have tranquilizing effects. However, the informational value of specific emotions differs from the informational value of global moods in two important ways.

First, emotions are specific reactions to specific events, whereas moods are of a diffuse and unfocused nature (Morris, 1989). Because of their unfocused nature, moods can be misread as a response to wide range of different targets—but once we attribute the mood to specific cause, its impact on unrelated judgments vanishes (Schwarz & Clore, 1983). In contrast, the source of a specific emotion is more likely to be in the focus of attention from the onset, thus limiting the emotion's informational value for unrelated judgments. Accordingly, Keltner, Locke, and colleagues (1993) observed that hav-

ing participants label their current feelings with specific emotion terms was as efficient in eliminating mood effects on subsequent judgments as a standard misattribution manipulation. Their findings are also an important methodological warning: Using specific emotion terms as manipulation checks invites causal attributions to determine the specific emotion, which can eliminate the expected effect.

Second, emotions reflect an underlying appraisal (e.g., Ellsworth & Scherer, 2003; Ortony et al., 1988) and the experience of a specific emotion implies that a specific set of appraisal criteria has been met. Anger, for example, informs us that somebody did us wrong and hence provides more specific information than a diffuse negative mood. If so, different emotions of the same valence should have differential effects, which can be predicted on the basis of the underlying appraisals (e.g., Lerner & Keltner, 2000). A growing body of literature supports this reasoning (for a review, see Han, Lerner, & Keltner, in press).

In an early study, Gallagher and Clore (1985) showed that feelings of fear affected judgments of risk but not of blame, whereas feelings of anger affected judgments of blame but not of risk. Similarly, Lerner, Gonzalez, Small, and Fischhoff (2003) observed in a national survey during the immediate aftermath of the terrorist attacks of September 11, 2001, that inducing participants to focus on the experienced fear increased risk estimates and plans for precautionary behavior, whereas focusing on the experienced anger did the reverse. Studying attributions of responsibility, Keltner, Ellsworth, and colleagues (1993) observed that angry participants assigned more responsibility to human agents than to impersonal circumstances, whereas sad participants assigned more responsibility to impersonal circumstances than to human agents, again consistent with the underlying appraisal patterns. Using a minimal group paradigm, DeSteno, Dasgupta, Bartlett, and Cajdric (2004) found that anger elicited more negative automatic evaluations of the outgroup, whereas sadness did not affect outgroup evaluations relative to a neutral mood control.

Going beyond judgment effects, several studies documented effects of emotions on goal-oriented behavior. Raghunathan and Pham (1999) suggested that sadness, a common response to the loss or absence of a reward, may prompt the goal of reward acquisition; in contrast, anxiety, a response to threats, may prompt a goal of uncertainty reduction. To test these predictions, they provided sad or anxious participants with a choice that required a trade-off between risk and rewards. As expected, sad individuals pursued reward acquisition and preferred high reward options, even though they came at the cost of high risk. Conversely, anxious individuals pursued uncertainty reduction and preferred low-risk options, even though they came at the cost of low reward. Similarly, Lerner, Small, and Loewenstein (2004) showed that induced disgust and sadness can affect economic decisions. Compared to a neutral emotion condition, disgusted participants sold goods they owned at a lower price and offered less money to acquire new goods, presumably because disgust prompts "expel" and "intake avoidance"

goals. Sad participants, on the other hand, offered more money to acquire goods, consistent with Raghunathan and Pham's (1999) observation that sadness motivates reward acquisition; however, they also sold what they had at a lower price, in contrast to what this account would predict. Lerner and colleagues attribute both findings to participants' desire to change their sad circumstances, which would motivate pricing decisions that facilitate the sale of what one owns and the acquisition of what one wants.

In combination, these findings indicate that specific emotions influence judgments in ways that are consistent with the underlying appraisals. Moreover, emotions give rise to specific goals, which influence subsequent behavior. Attributing the emotion to an unrelated event would presumably eliminate these effects (e.g., Schwarz et al., 1985).

Bodily Sensations

The research on mood and emotions indicates that incidental feelings influence judgment only when the nature of the feelings, the salience of their causes, or other aspects of the situation allow them to be experienced as reactions to the target of judgment. The available research suggests that this conclusion also applies to bodily sensations.

Arousal States

Exploring the impact of heightened excitation levels, Zillman (1978) had participants engage in various forms of exercise. Shortly after the exercise, no impact of increased excitation level was observed, presumably because participants were still aware of its source. After some delay, however, subsequent judgments were affected by the residual arousal. Apparently, participants misinterpreted their arousal as a reaction to the target, once the temporal distance of the exercise rendered this alternative source less accessible and plausible. Similarly, Zanna and Cooper (1976) observed in their classic misattribution experiments that cognitive dissonance effects were eliminated when participants could attribute the resulting arousal state to some other source.

These lines of research indicate that individuals draw on their perceived arousal state as a source of information, unless its informational value is called into question, as we have seen for other feelings. Moreover, apparent "excitation" effects can be obtained in the absence of any actual arousal, based on false feedback (e.g., Valins, 1974). This is incompatible with the assumption that some actual arousal needs to be present to be "transferred" (Zillman, 1978) and highlights the informational value of *perceived* arousal as the crucial ingredient.

Facial Feedback

Darwin (1872/1965) noted that "most of our emotions are so closely connected with their expression that they hardly exist if the body remains passive" (p. 257). Taking a more extreme view, James (1890) maintained that emotions were, in fact, nothing more the awareness of

our expressions: "We feel sorry because we cry, angry because we strike, and afraid because we tremble" (p. 243). Experimental research provided some support for the assumption that emotional expressions may elicit distinctive emotional experiences. For example, Ekman, Levenson, and Friesen (1983) found that posing different facial expressions of emotion triggered different patterns of autonomic activity (e.g., changes in heart rate, skin temperature, and skin conductance). Hence, the impact of facial expressions on judgment may involve autonomic activity as well as proprioceptive feedback.

For example, Keltner, Ellsworth, and colleagues (1993) induced participants to take on a sad or angry facial expression, without labeling the expression in emotion terms. These expressions affected attributions of responsibility in the same way previously discussed for recall-induced sadness and anger. In fact, the influence of facial expressions alone was larger than the influence of facial expressions combined with an emotional recall task, perhaps because the influence of the expression itself was less transparent. Using a particularly subtle manipulation, Strack and colleagues (1988) had participants rate the funniness of cartoons while holding a pen in their mouths in such a manner that a smile was either facilitated or inhibited. Holding a pen between one's teeth requires that one contract the same muscles that are used when one smiles, while holding a pen between one's lips results in pursing the lips, rendering a smile impossible. As expected, participants reported greater amusement at the cartoons when the muscle contractions resembled rather than inhibited a smile. Stepper and Strack (1993) further showed that such processes are not limited to facial feedback. Their participants reported higher pride when they received positive performance feedback in an upright posture rather than a slumped posture.

As observed for moods and emotions, the impact of emotional expressions can be discounted and augmented by suitable attributional manipulations (Olson & Roese, 1995). Moreover, bodily experiences are perceived as a response to whatever is in the focus of one's attention, rendering their influence more context dependent than has often been assumed. For example, Tamir, Robinson, Clore, Martin, and Whitaker (2004) observed that brow (vs. cheek) tension reduced preferences in an easy judgment context but increased preferences in a difficult context. Similarly, head shaking (vs. nodding) either increased or decreased prosocial affect, depending on the context in which the judged character was presented. Finally, a subliminal smile (vs. frown) led to higher self-ratings of performance when paired with one's own actions but to lower self-ratings of performance when paired with a competitor's actions.

Other research documented effects of bodily sensations beyond the affective domain. For example, contracting the corrugator muscle elicits a feeling of effort that can affect judgment in ways that parallel the effect of difficulty of recall or thought generation (e.g., Sanna, Schwarz, & Small, 2002; Stepper & Strack, 1993). Similarly, arm flexion and arm extension can provide approach and avoidance feedback that influences processing style in ways that parallel the influence of happy and sad moods (e.g., Friedman & Förster, 2000). We address these findings in the respective sections that follow.

Cognitive Experiences

The same principles hold for metacognitive experiences, which only influence individuals' judgments if their informational value is not called into question.

Accessibility Experiences

According to most models of judgment, we should evaluate an object more favorably when we bring many rather than few positive attributes to mind; similarly, we should consider an event more likely when we generate many rather than few reasons for its occurrence. Empirically, this is often *not* the case (e.g., Sanna, Schwarz, & Small, 2002; Sanna, Schwarz, & Stocker, 2002; Wänke, Bohner, & Jurkowitsch, 1997). Recalling many attributes or generating many reasons is more difficult than recalling or generating only a few, and these metacognitive accessibility experiences are informative in their own right. What people conclude from them is more malleable than the conclusions drawn from moods and emotions and depends on which of many naive theories of mental functioning they apply (Schwarz, 2004).

THE ACCESSIBILITY–FREQUENCY LINK

One naive theory links recall experiences to characteristics of the external world and holds, "The more exemplars exist, the easier it is to bring some to mind." This correct belief is at the heart of Tversky and Kahneman's (1973) availability heuristic and people infer higher frequency and probability when examples are easy rather than difficult to bring to mind. Because frequent exemplars are also more typical for their category, ease of recall further suggests high typicality. Accordingly, people infer that they use their bicycles more often after recalling few rather than many instances (Aarts & Dijksterhuis, 1999); rate themselves as more assertive after recalling few rather than many of their own assertive behaviors (Schwarz, Bless, Strack, et al., 1991); and hold an attitude with more confidence after generating few rather than many supporting arguments (Haddock, Rothman, Reber, & Schwarz, 1999).

When people apply this naive theory, their inferences are consistent with the implications of *what* comes to mind when recall or thought generation is experienced as easy but opposite to these implications when it is experienced as difficult. These effects cannot be traced to differences in the quality of the recalled examples but are eliminated when the subjective experience is misattributed to an external influence. In the latter case, participants draw on accessible content, reversing the otherwise observed pattern (e.g., Haddock et al., 1999; Sanna, Schwarz, & Small, 2002; Schwarz, Bless, Strack, et al., 1991). Moreover, yoked participants, who merely read the thoughts generated by another and are hence deprived of the generation experience, are more influenced when their partner lists many rather than few argu-

ments, in contrast to the person who lists them (e.g., Wänke, Bless, & Biller, 1996). Finally, the same effect can be observed when all participants list the same number of thoughts and their subjective experience of difficulty is manipulated through facial feedback in the form of corrugator contraction, an expression associated with mental effort (e.g., Sanna, Schwarz, & Small, 2002; Stepper & Strack, 1993). Thus, merely furrowing one's brow can be sufficient to arrive at conclusions that are opposite to the content of one's thoughts.

Experienced difficulty of thought generation can thwart the success of popular debiasing strategies, which encourage people to guard against overconfidence, hindsight bias, and similar fallacies by thinking about counterfactual alternatives. Such strategies only work when generating counterfactuals is experienced as easy but backfire when it is experienced as difficult (Sanna & Schwarz, 2004; Schwarz, Sanna, Skurnik, & Yoon, 2007).

OTHER NAIVE THEORIES

While the accessibility–frequency link has received most attention, people hold a variety of other naive theories about memory and thought generation, rendering the inferences from accessibility experiences highly malleable (for a review, see Schwarz, 2004). For example, they correctly assume that details of recent events are easier to recall than details of distant events, and details of important events easier than details of unimportant ones. Which of these theories they apply depends on the task posed. Xu and Schwarz (2005) had participants recall details of the Oklahoma City bombing. When first asked to date the event, participants inferred that it was more recent after recalling 2 rather than 10 details; when first asked how important the event was to them at the time, they inferred higher importance after recalling 2 rather than 10 details. More important, application of a given theory entails an attribution of the experience to a specific cause (here, recency or importance), which can change the implications of the experience for subsequent judgments (Schwarz, 2004). Hence, participants who initially attributed the difficulty of recalling many details to the event's temporal distance subsequently reported that the event was quite important to them— presumably because they could still recall details even though the event had apparently happened long ago. Conversely, participants who initially attributed difficulty of recall to low personal importance subsequently dated the event as closer in time—presumably because they could still recall details despite the event's low personal importance.

PROCESSING MOTIVATION

Finally, people are more likely to rely on their accessibility experiences under conditions that commonly foster heuristic processing but to turn to accessible content under conditions that commonly foster systematic processing. Rothman and Schwarz (1998) asked men to recall few or many behaviors that increase or decrease their risk of heart disease. Men without a family history of heart disease relied on their accessibility experiences and inferred *higher* risk after recalling few risk-increasing or many risk-decreasing behaviors. In contrast, men with a family history relied on recalled content and inferred *lower* risk under these conditions. Apparently, the personal relevance of the task, indicated by their family history, influenced the processing strategy used (see also Grayson & Schwarz, 1999). Similarly, Ruder and Bless (2003) observed that being in a good mood, a condition that fosters heuristic processing as reviewed below, increased reliance on accessibility experiences. Conversely, being in a sad mood, a condition that fosters systematic processing, increased reliance on recalled content.

In combination, these findings challenge the widely shared assumption that we can predict people's judgments by knowing *what* comes to mind. Unless we take the person's metacognitive experiences, relevant naive theories, and processing motivation into account, our predictions will often be erroneous (Schwarz, 2004).

Processing Fluency

Numerous variables can influence the fluency with which new information can be processed, ranging from visual characteristics of the presentation (like figure–ground contrast or print font) to the semantic relatedness of the material and the frequency and recency of previous exposure. Because these variables result in similar phenomenological experiences of fluent processing, the meaning of the experience is open to interpretation. Which interpretation people choose and which inferences they draw again depend on the naive theory they bring to bear. Applicable theories are recruited by the task and application of one theory usually renders the experience uninformative for inferences that require a different theory (Schwarz, 2004).

Some naive theories pertain to presentation conditions; people assume that material is easier to process when shown for long rather than short durations, with high rather than low clarity, and so on. Other theories pertain to one's state of knowledge; people assume, for example, that familiar material is easier to process than unfamiliar material. When fluency deriving from one's state of knowledge is brought to bear on judgments of presentation characteristics, it results in *illusions of perception*. Thus, people infer that a stimulus was presented for a longer duration, or with higher clarity, when it is easy to process due to earlier exposures (e.g., Whittlesea, Jacoby, & Girard, 1990; Witherspoon & Allan, 1985). Conversely, when fluency deriving from favorable presentation conditions is brought to bear on judgments of one's knowledge, it results in *illusions of memory* and people erroneously infer that the stimulus is familiar (e.g., Whittlesea et al., 1990), resulting in false recognition judgments.

Even when processing fluency is due to previous exposure, and correctly attributed to this source, it may result in erroneous judgments when perceivers misidentify the specific instance of exposure. For example, learning the names of nonfamous individuals in an experimental session may later result in erroneous judgments of fame when one of the names seems familiar and one cannot re-

call the context in which it was initially encountered (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989). As in the case of accessibility experiences, bodily sensations that convey mental effort can mirror low fluency. Strack and Neumann (2000) observed in a conceptual replication of Jacoby and colleagues' (1989) fame study that contracting the corrugator muscle can protect against illusions of fame, presumably by conveying low fluency. Finally, the impact of fluency experiences is eliminated when people (mis)attribute fluency to an unrelated source (for a review, see Kelley & Rhodes, 2002), as observed for other feelings.

Several consequences of fluency are of particular interest to social psychologists. First, fluency affects *judgments of truth*. Presumably, the sense of familiarity that arises from high fluency suggests that one has heard the statement before, which may serve as a consensus cue—if many people believe it, there is probably something to it (Weaver, Garcia, Schwarz, & Miller, in press). This is most apparent when fluency results from repeated exposure and numerous studies obtained robust "illusions of truth" under these conditions (e.g., Begg, Anas, & Farinacci, 1992; Hasher, Goldstein, & Toppino, 1977). After some delay, illusions of truth can even emerge when the statement is explicitly marked as "false" (e.g., Skurnik, Yoon, Park, & Schwarz, 2005). This poses major problems for education campaigns and rumor control, as Allport and Lepkin (1945) noted decades ago: Corrections of false information often entail a reiteration of the false statement, thus increasing its fluency and later acceptance. More surprisingly, simply presenting statements with good figure–ground contrast (Reber & Schwarz, 1999), or in a rhyming form (McGlone & Tofighbakhsh, 2000), is sufficient to increase their acceptance as true. This suggests that any variable that increases fluency may also increase perceived truth.

Second, fluency affects judgments of *liking*, *preference*, and *beauty*. As Zajonc (1968) demonstrated, repeated exposure to an initially neutral stimulus, without any reinforcement, leads to gradual increases in liking (for a review, see Bornstein, 1989). Several authors suggested that this mere exposure effect may be due to increased fluency (e.g., Seamon, Brody, & Kauff, 1983). Supporting this hypothesis, a growing body of findings indicates that *any* variable that facilitates fluent processing increases liking, even under conditions of a single exposure. For example, people like the same stimulus more when it is preceded by a visual (Reber, Winkielman, & Schwarz, 1998) or semantic (Winkielman et al., 2003) prime that facilitates fluent processing and less when it is preceded by primes that impede fluent processing. In fact, the influence of many variables long known to affect liking and aesthetic preference, from figure–ground contrast to symmetry and prototypicality, can be can be traced to increased processing fluency (Reber, Schwarz, & Winkielman, 2004).

The available evidence suggests that fluency itself is hedonically marked and elicits a positive affective response that can be captured with psychophysiological measures (Winkielman & Cacioppo, 2001). This affective response presumably mediates the effect of fluency

on preference judgments. Supporting this assumption, Winkielman and Fazendeiro (reported in Winkielman et al., 2003) observed that the influence of fluency on liking was eliminated when participants misattributed their positive affect to music played in the background. What is less clear is *why* processing fluency is experienced as affectively positive. Relevant proposals range from the adaptive value of a preference for familiar stimuli (Freitas, Azizan, Travers, & Berry, 2005; Zajonc, 1968) to the adaptive value of fast stimulus identification (Winkielman, Schwarz, & Nowak, 2002).

The parallel effects of processing fluency on judgments of truth and aesthetic preference also shed new light on Keats's famous assertion that "beauty is truth, truth is beauty"—both judgments are based, in part, on the same experiential information. Moreover, the relationship between familiarity and affective response is bidirectional. Stimuli that evoke positive affect also seem more familiar, even when processing fluency is controlled for (Monin, 2003), as do novel arguments when people are in elated moods (Garcia-Marques & Mackie, 2001). Once again, the single window we have on our subjective experiences makes it difficult to distinguish between commonly related experiences and facilitates misattributions.

Priming as a Misattribution Process

The same attribution logic holds for our thoughts (Clore, 1992; Clore & Colcombe, 2003). We usually assume that the thoughts that come to mind are "about" whatever we are thinking about (Higgins, 1998)—or why else would they come to mind now? When we are aware that our thoughts may have been prompted by an unrelated influence, we discount them as a source of information. Hence, priming procedures work best when they are subtle and embedded in other tasks (e.g., Higgins, Rholes, & Jones, 1977), allowing us to misread the primed thoughts as a response to the target. But priming manipulations backfire when they are blatant (e.g., Martin, 1986) or participants' attention is drawn to them (e.g., Strack, Schwarz, Bless, Kübler, & Wänke, 1993), making people aware that their thoughts may not be "about" the target. In these cases, people attempt to correct for the perceived influence, often resulting in overcorrection (Strack & Hannover, 1996; Wilson & Brekke, 1994). This experiential perspective on the influence of thought content blurs the lines between the experiential and cognitive accounts described earlier. It suggests, for example, that even if judgment effects were to involve mood-congruent recall from memory, they might be attenuated or eliminated if people became aware that these thoughts had come to mind only because the people were in a good or bad mood.

Summary

In sum, people assume that their thoughts and feelings are "about" what is in the focus of their attention. They draw on their apparent responses to the target as relevant information, unless they are aware that their experi-

ences may be due to another source. In the latter case, they may attempt to correct for their likely influence; this often results in theory-driven overcorrection, unless alternative diagnostic inputs are available. Reliance on one's affective feelings is particularly likely when the judgment is affective in nature (e.g., preference or well-being), which often renders one's feelings the most diagnostic source of information available; when little other information is accessible; when the judgment is overly complex and cumbersome to make on the basis of a piecemeal information processing strategy; or when time constraints or competing task demands limit the attention that may be devoted to evaluating the informational value of one's feelings or to searching for alternative inputs. The same logic applies to specific emotions, bodily sensations, and metacognitive experiences, with the qualifications noted previously. Hence, forming judgments on the basis of one's feelings may sometimes be thought of as a simplifying heuristic strategy. Note, however, that this strategy entails assessments of the informational value of one's feelings and of their relevance to the task at hand, as well as the use of naive theories as inference rules.

The feelings-as-information approach provides a parsimonious and unifying framework for conceptualizing the influence of a wide variety of phenomenological experiences; it is the only available conceptualization for the influence of cognitive experiences. Moreover, its predictions are consistent with the bulk of the evidence. Obviously, this supportive evidence, by itself, does not rule out other pathways of influence. This issue is particularly contentious for the influence of moods, which other researchers are inclined to attribute primarily to mood-congruent recall (for extensive reviews, see Forgas, 1995a, 2001, 2003). Mood-congruent recall could, in principle, provide a plausible pathway for mood effects on judgment even under conditions in which the feeling itself is discredited as a source of information and people engage in an effortful, piece-meal judgment strategy. However, in the absence of (mis)attribution manipulations that discredit the informational value of participants' current feelings, it is difficult to determine which process drives a particular instance of mood effects on evaluative judgment.

FEELINGS AND STRATEGIES OF INFORMATION PROCESSING

Numerous findings indicate that feelings may influence individuals' spontaneous adoption of heuristic or systematic strategies of information processing. While the evidence is uncontroversial, there is less agreement on the underlying process. Moreover, most of the conceptualizations focus on the influence of happy or sad moods. One approach attributes affect-induced differences in processing strategy to differences in attentional resources, although some researchers assume that negative moods limit resources (e.g., Ellis & Ashbrook, 1988), whereas others assume that positive moods do so (e.g., Mackie & Worth, 1989). A second approach attributes

differences in processing style to the informational functions of moods. Variants of this approach assume that our feelings inform us about the state of the environment (Schwarz, 1990; Weary, Marsh, Gleicher, & Edwards, 1993), the contingency of hedonically relevant rewards (Wegener & Petty, 1994), or the progress made in goal pursuit (Clore et al., 2001), or that they serve as input into specific performance decisions (Martin, Ward, Achee, & Wyer, 1993). A third approach focuses on mood maintenance and repair. It assumes that we protect our pleasant affect by avoiding excessive mental effort when we feel happy (mood maintenance) but engage in effortful thought to distract ourselves and to improve our mood when we feel sad (mood repair; e.g., Clark & Isen, 1982; Erber & Erber, 2001). These different accounts gave rise to a lively debate (see the contributions in Martin & Clore, 2001). However, their exclusive focus on the influence of moods captures only part of the accumulating evidence. As reviewed next, proprioceptive approach or avoidance feedback (Friedman & Förster, 2000, 2002) is functionally equivalent to being in a happy or sad mood, as are environmental cues, like smiling or frowning faces (Ottati, Terkildsen, & Hubbard, 1997) and even the color of the paper on which a task is presented (Soldat, Sinclair, & Mark, 1997). We therefore attempt to provide an integrative account that captures these commonalities while maintaining core assumptions of previous conceptualizations (presented in Martin & Clore, 2001).

Situated Cognition: Cognitive Processes Are Tuned to Meet Situational Requirements

"My thinking is first and last and always for the sake of my doing," noted William James (1890, p. 333) more than a century ago. To serve our doing in adaptive ways, our cognitive processes are responsive to the environment in which we pursue our goals, as a growing body of research into situated cognition indicates (see Smith & Semin, 2004). We propose that the adaptively tuned nature of human cognition is at the heart of the observed shifts in processing strategy (Schwarz, 2002). If so, we should observe that external and internal cues that signal a benign and unproblematic environment are functionally equivalent, as are external and internal cues that signal a problematic environment. These signals have cognitive and motivational consequences, which are highly adaptive under most circumstances.

When facing a problematic situation, we are usually motivated to do something about it. Any attempt to change the situation requires a careful assessment of its features, an analysis of their causal links, detailed explorations of possible mechanisms of change, and anticipation of the potential outcomes of any action that might be initiated. Consistent with these conjectures, negative affective states are associated with a narrowed focus of attention (e.g., Broadbent, 1971; Easterbrook, 1959) and a higher level of spontaneous causal reasoning (e.g., Bohner et al., 1988), paralleling the observation that failure to obtain a desired outcome shifts attention to a lower level of abstraction (e.g., Wegner & Vallacher,

1986). These influences foster vigilance and bottom-up, data-driven processing (Bless, 1997; Schwarz, 1990), in particular when the negative feeling entails a high uncertainty appraisal, as is the case for sadness (Tiedens & Linton, 2001). Moreover, it may seem unwise to rely on one's usual routines and preexisting general knowledge structures under these conditions, thus discouraging top-down strategies. Finally, we may be unlikely to take risks in a situation that is already marked problematic and may therefore avoid simple heuristics and uncertain solutions.

Conversely, when we face a benign situation that poses no particular problem, we may see little need to engage in detailed analyses and may rely on our usual routines and preexisting knowledge structures, which have served us well in the past (Bless, 1997; Schwarz, 1990). This encourages less effortful, top-down processing as a default, *unless* current goals require otherwise. In pursuing such goals, we may be willing to take some risk, given that the general situation is considered safe. As a result, we may prefer simple heuristics over more effortful, detail-oriented judgmental strategies, and may explore new procedures and possibilities and pursue unusual, creative associations.

In combination, these conjectures suggest that our cognitive processes are tuned to meet the situational requirements signaled by our feelings. In contrast to earlier conceptualizations, which pertained exclusively to the influence of moods, this cognitive tuning hypothesis does *not* entail that happy individuals are somehow unable (Mackie & Worth, 1989) or unwilling (Schwarz & Bless, 1991) to engage in systematic processing. Rather, it merely entails that the happy mood *itself* does not signal a situation that poses particular processing requirements. Hence, the spontaneously adopted heuristic processing style and reliance on preexisting knowledge structures should be easy to override, rendering processing under happy moods more flexible than processing under sad moods. In contrast, the systematic processing style fostered by negative moods should be difficult to override, reflecting that it would be maladaptive to ignore a potential "problem" signal (Bless & Schwarz, 1999). Moreover, any variable that can signal a benign or problematic situation should have effects that parallel the influence of happy or sad moods. Finally, specific emotions provide information that goes beyond global benign/problem signals and elicit processing strategies that are specifically tuned to meet the requirements entailed in the underlying appraisal pattern.

The feelings-as-information logic further predicts that feelings should exert no influence on processing style when their informational value is called into question. Empirically this is the case, as Sinclair and colleagues (1994) and Gasper (2004) observed for the case of moods, in contrast to what other approaches would predict.

This conceptualization does not preclude other influences. Recent negative events, for example, may temporarily preoccupy the person to an extent that interferes with systematic processing, despite the presence of negative affect (Ellis & Ashbrook, 1988). Similarly, people may indeed engage in mood repair when they feel bad (Erber & Erber, 2001). But it would be maladaptive to do so at the expense of attention to the causes of one's feelings. Hence, we would expect mood maintenance and repair to take priority only after the person concluded that little could be done or when the laboratory task seems irrelevant to begin with. Also, mood maintenance and repair would be expected when pleasant affect (rather than performance) is a primary goal, as might be expected when experimental instructions suggest so (e.g., Wegener, Petty, & Smith, 1995).

Feelings and Cognitive Performance

We first illustrate the interplay of feelings and task requirements in the choice of processing strategies. Subsequently, we review prototypical findings, highlighting the functional equivalence of different feelings and environmental cues.

Feelings and Task Requirements: Moods, Scripts, and Concentration Tests

If happy moods increase, and sad moods decrease, our tendency to rely on the "usual routines," people should be more likely to rely on an applicable script (Schank & Abelson, 1977) when they are in a happy rather than sad mood. Empirically, this is the case. Employing a dual-task paradigm, Bless, Clore, and colleagues (1996) had participants listen to a tape-recorded restaurant story that contained script-consistent and script-inconsistent information. While listening to the story, participants worked on a concentration test that required them to mark certain letters on a work sheet. Good performance on this test requires detail-oriented processing; in contrast, the restaurant story can be understood by engaging *either* in script-driven top-down processing or in data-driven bottom-up processing. As predicted, happy participants were likely to recognize previously heard script-inconsistent information but showed high rates of erroneous recognition of script-consistent information—the classic pattern of schema-guided memory. Neither of these effects was obtained for sad participants, indicating that they were less likely to draw on the script. Given that top-down processing is less taxing than bottom-up processing, we may further expect that happy participants do better on a secondary task. Confirming this prediction, happy participants outperformed sad participants on the concentration test.

In combination, these findings indicate that moods influence the spontaneously adopted processing style under conditions in which different processing styles are compatible with the individual's goals and task demands, as was the case for comprehending the restaurant story. Under these conditions, sad individuals are likely to spontaneously adopt a systematic, bottom-up strategy, whereas happy individuals rely on a less effortful top-down strategy. But when task demands (as in the case of the aforementioned concentration test) or explicit instructions (e.g., Bless, Bohner, Schwarz, & Strack, 1990)

require detail-oriented processing, happy individuals are able and willing to engage in the effort. The latter observation is inconsistent with the proposal that happy moods limit cognitive resources (e.g., Mackie &Worth, 1989), or generally impair processing motivation (e.g., Schwarz & Bless, 1991).

Categorization

The detail-oriented, bottom-up processing style associated with sad moods fosters the formation of fine-grained, narrow categories, whereas the top-down, heuristic processing style associated with happy moods fosters the formation of more inclusive categories (Isen, 1987). Thus, happy participants are more likely to include unusual exemplars in a category than participants in a neutral mood, assigning, for example, "feet" to the category "vehicles" and "cane" to the category "clothing" (Isen & Daubman, 1984). They also list more unusual exemplars when given a category prompt (Hirt, Levine, McDonald, Melton, & Martin, 1997) and match geometric figures on the basis of global rather than local similarities (Gasper & Clore, 2002). Finally, Beukeboom and Semin (2005) observed that sad participants identified behaviors at a lower level of abstractness than happy participants, paralleling the impact of actual obstacles (Wegner & Vallacher, 1986).

Similar effects have been obtained with other manipulations that signal a benign or problematic situation. Using a subtle bodily feedback manipulation, Friedman and Förster (2000, 2002) asked participants to either press their hand downward against the top of the table (arm extension) or upward against the bottom of the table (arm flexion). Arm extension (as in pushing something away) is usually associated with avoidance behavior, whereas arm flexion (as in pulling something closer) is associated with approach behavior. The proprioceptive feedback of the involved muscle activation is sufficient to influence processing style: Relative to a control, participants who flexed their arms provided more inclusive categorizations, whereas those who extended their arms provided less inclusive categorizations.

Creative Problem Solving

As may be expected on the basis of the categorization findings, happy individuals typically outperform sad or neutral-mood individuals on creativity tasks, like the Remote Associates Test (e.g., Isen, Daubman, & Nowicki, 1987) or Duncker's candle problem (Isen & Daubman, 1984). Again, studies using arm flexion or extension to provide bodily approach–avoidance feedback obtained parallel results (Friedman & Förster, 2000). Participants who flexed their arms were more likely to break the set than participants who extended their arms, resulting in better performance on tests that require the identification of targets hidden in complex visual patterns. Performance on such tasks is facilitated by the application of familiar concepts to the hidden targets, while disregarding irrelevant detail and breaking the set imposed by the

distractor—requirements that give top-down processing an advantage over bottom-up processing.

Analytic Reasoning Tasks

If sad moods foster systematic, detail-oriented processing, they should facilitate performance on analytic reasoning tasks. The bulk of the available evidence is consistent with this prediction (see Schwarz & Skurnik, 2003, for a review). For example, Fiedler (1988) reported that sad participants produced fewer inconsistencies in multiattribute decision tasks than happy participants. Specifically, the latter were twice as likely than the former to violate transitivity by producing inconsistent triads of the form A > B and B > C, but A < C. Similarly, Melton (1995) observed that happy participants performed worse on syllogisms than participants in a neutral mood. Again, the influence of bodily approach–avoidance signals parallels these effects. Using analytical reasoning tasks taken from the graduate record exam (GRE), Friedman and Förster (2000) observed that participants who extended their arms (avoidance) solved nearly twice as many problems correctly as participants who flexed their arms (approach).

Finally, external cues can serve the same function. Soldat and colleagues (1997) presented analytic reasoning tasks, also taken from the GRE, on paper of an upbeat red, or a somewhat depressing blue, hue. Across several replications, participants performed better when the tasks were printed on blue rather than red paper, with white paper falling in between. The performance advantage of blue paper was most pronounced for complex tasks, which posed higher processing demands. Paralleling these laboratory findings, Sinclair, Soldat, and Mark (1998) found that students did better on an exam when printed on blue rather than red paper, in particular for difficult questions.

In contrast to the foregoing findings, mostly based on tasks taken from the GRE, other studies revealed performance deficits under depressed affect on complex logic and mathematics tasks (for a review, see Clore et al., 1994). Theoretically, mixed findings are to be expected for such tasks because none of the hypothesized processes will *necessarily* result in improved performance. For example, greater attention to detail per se will not improve performance when the task requires the application of an unknown algorithm, and it may impede performance when the person gets sidetracked by irrelevant details. Similarly, top-down processing strategies may facilitate as well as impede performance, depending on whether the available heuristic is applicable to the current task. It is therefore not surprising that the most consistent findings have been obtained with common social reasoning tasks, like persuasion and impression formation, with which people are highly familiar.

Persuasion

In general, strong arguments are more persuasive than weak arguments when recipients engage in systematic

processing, whereas argument strength exerts little influence under heuristic processing (for a review, see Petty & Cacioppo, 1986). Accordingly, the impact of argument strength can serve as a diagnostic tool for assessing processing strategy. Studies using this strategy consistently found that happy recipients engage in less, and sad recipients in more, elaboration of counterattitudinal messages than recipients in a nonmanipulated mood (e.g., Bless et al., 1990; see Schwarz, Bless, & Bohner, 1991, for a review). Hence, happy recipients are moderately and equally persuaded by strong as well as weak arguments, whereas sad recipients are strongly persuaded by strong arguments and not persuaded by weak arguments. Conversely, Worth and Mackie (1987) observed that happy recipients were more likely than sad recipients to rely on heuristic strategies in assessing the validity of the message, paying attention to cues such as the communicator's status or expertise in forming a judgment. Consistent with the feelings-as-information logic, these effects are eliminated when recipients are aware that their mood is due to an unrelated source (Sinclair et al., 1994).

As noted previously, however, the impact of moods can be overridden by other variables; hence explicit instructions to pay attention to the arguments (e.g., Bless et al., 1990), or the promise that carefully thinking about the message would make one feel good (e.g., Wegener et al., 1995), have been found to elicit systematic message processing in happy recipients. What characterizes the information processing of happy individuals is not a general cognitive or motivational impairment but a tendency to spontaneously rely on heuristic strategies and general knowledge structures in the absence of goals that require otherwise.

Paralleling the effects of recipients' moods, Ottati and colleagues (1997) observed that the same message is less likely to be scrutinized when presented by a communicator with a smiling, happy face than when presented by a communicator with a neutral, somber face. They suggested that the communicator's conveyed affect can serve informative functions that parallel recipients' own affect. Further illustrating the power of environmental affective cues, Soldat and Sinclair (2001) had participants read persuasive messages printed on colored paper. Participants were persuaded by strong arguments, but not by weak arguments, when the message was presented on paper of depressing blue hue. However, both types of arguments were similarly persuasive when the paper had an upbeat red hue.

Stereotyping and Impression Formation

Paralleling the persuasion findings, numerous studies indicate that perceivers in a sad mood are more likely to elaborate individuating information about the target person, whereas perceivers in a happy mood are more likely to draw on the person's category membership as a heuristic cue. This results in more stereotypical judgments under happy than under sad moods (e.g., Bodenhausen, Kramer, & Süsser, 1994; Isbell, 2004; for a review, see Bless, Schwarz, & Kemmelmeier, 1996). Related research on the influence of brand names on product evaluations

similarly shows higher reliance on brand information under happy than sad moods (e.g., Adaval, 2001). Individual differences in chronic affect parallel these findings. Mildly depressed perceivers attend more to individuating information than do nondepressed perceivers (e.g., Edwards & Weary, 1993) and seek more, and more diagnostic, information (Hildebrandt-Saints & Weary, 1989). Finally, happy individuals' reliance on category membership information can again be overridden by manipulations that increase their processing motivation, such as personal accountability for one's judgment (Bodenhausen et al., 1994) or anticipated interaction with the target person (e.g., Hildebrandt-Saints & Weary, 1989).

Summary

In sum, internal and external cues that signal a benign or problematic situation have cognitive and motivational consequences. Human cognition is tuned to meet situational requirements and problem signals foster vigilance and the adoption of a detail-oriented bottom-up processing style, which is usually adaptive. Signals that characterize the situation as benign, on the other hand, are not, by themselves, associated with particular processing requirements. They foster reliance on preexisting knowledge structures and top-down processing, *unless* goals or task demands require otherwise. Which processing strategy facilitates or impedes performance depends on the specific task. The bulk of the evidence is compatible with this framework (Schwarz, 2002), which offers a unified conceptualization of the operation of internal and external signals in the context of situated cognition.

This does not preclude the operation of other pathways under specific conditions. Sometimes people's preoccupation with recent happy or sad events will indeed limit their cognitive resources (Ellis & Ashbrook, 1988; Mackie & Worth, 1991), although it is difficult to separate the impact of affect per se from the impact of event-related thoughts. Other times, people will indeed attempt to focus on something else to improve a bad mood (e.g., Wegener et al., 1995) or avoid mental effort to maintain a good mood (e.g., Clark & Isen, 1982), although dealing with current problems will probably take precedence in naturalistic contexts. Similarly, moods may serve as input into specific performance decisions (Martin, 2001) and may influence which material comes to mind when we work on a problem (Forgas, 2001). None of these mood-related processes, however, is sufficiently general to account for the observed parallel effects of affective, bodily, and cognitive cues that signal a benign or problematic environment.

Finally, we note that the reviewed findings provide little support for assumptions that positive feelings will have mostly "positive" effects (Fredrickson, 2001). Instead, their effect depends on the specific task at hand and positive feelings can facilitate positive outcomes (such as increased creativity or resiliency; see Fredrickson, 2001) as well as negative ones (such as increased stereotyping and impaired logical problem solving), in contrast to what positive psychologists seem to hope for.

FEELINGS AND MEMORY

Historically, conscious experiences that accompany the process of remembering played an important role in the study of memory, until they went out of fashion with the behaviorist revolution (for reviews, see Brewer, 1992; Roediger, 1996). Over the last two decades, cognitive psychologists have begun to correct the resulting pervasive "neglect of conscious experience" (Tulving, 1989, p. 4) by rediscovering some of the historic themes. This rediscovery is part of a shift from a quantity-oriented "storehouse" metaphor of memory to an accuracy-oriented "correspondence" metaphor of memory (Koriat, Goldsmith, & Pansky, 2000). The storehouse metaphor is exemplified by the list-learning paradigm, with a focus on how many previously learned items can be recovered. Within this metaphor, the role of subjective experiences is conceptualized in terms of a storage architecture, as illustrated by Bower's (1981) model that treats moods and emotions as nodes in a network. The correspondence metaphor treats memory as a "perception of the past" and focuses on whether "this perception is veridical or illusory" (Koriat et al., 2000, p. 484). Inference processes play a key role in this approach. They can be intuitive or analytic, paralleling dual-process distinctions in social cognition, and based on phenomenal experiences that accompany the remembering process as well as other information. For example, the source monitoring approach (Johnson, Hashtroudi, & Lindsay, 1993) assumes that remembering includes inferences about the source of the memories that come to mind, which are based on their phenomenal qualities. From this perspective, all memory failures (except omissions) are based on a failure of source monitoring (i.e., an inferential process). Cognitive experiences in the form of perceptual and conceptual fluency are central to Jacoby and Kelley's (e.g., 1998) attributional approach to memory, which "conceives of remembering as a combination of fluent processing of an event with the mental set that attributes the fluency to past experience " (Roediger, 1996, p. 88).

To date, research into the role of cognitive experiences in memory has almost exclusively taken an approach that is consistent with the feelings-as-information logic. In contrast, research into the role of moods and emotions in memory has mostly been guided by the architectural assumptions of Bower's (1981) network model and has paid little attention to inferential processes, which figure more prominently in judgment research. Next, we review prototypical findings.

Cognitive Experiences

In our review of fluency effects in judgment, we noted that fluency due to one's state of knowledge can result in perceptual illusions when applied to judgments about the presentation of material. Conversely, fluency due to the conditions of presentation can result in memory illusions when applied to judgments of one's knowledge (for reviews, see Kelley & Rhodes, 2002; Koriat et al., 2000). For example, in *recognition* tests, people are more likely to identify a new stimulus as old when its processing

is facilitated by a preceding prime (e.g., Jacoby & Whitehouse, 1989), higher visual clarity of the presentation (e.g., Whittlesea et al., 1990), or a highly related semantic context (Whittlesea, 1993). Paralleling effects observed in judgments of truth and beauty, Rhodes and Kelley (2003) also observed erroneous recognition when the test items were preceded by a rhyming rather than nonrhyming prime. Such enhanced fluency from an unrecognized source presumably also underlies *déjà vu* experiences, where something new seems strangely familiar (Jacoby & Whitehouse, 1989). The observation that bodily sensations of effort can affect fluency-based fame judgments (Strack & Neumann, 2000) further suggests that bodily feedback may also influence recognition, although relevant data are not yet available. Finally, fluency effects on recognition and feelings of familiarity are not obtained when participants attribute fluency to a source other than previous exposure—for example, because they are aware of the prime (Jacoby & Whitehouse, 1989) or realize that clarity of presentation is being manipulated (Whittlesea et al., 1990).

Models that focus solely on *what* comes to mind cannot account for the reviewed phenomena, nor can they account for related research into feelings of knowing and judgments of learning (for a review, see Koriat & Levy-Sadot, 1999). At the same time, it is apparent that the experiential approach is relatively silent with regard to *what* comes to mind and instead focuses on the conditions under which we consider whatever comes to mind a reflection of past events.

Mood

Moods may influence memory at the encoding as well as retrieval stage. Extending the investigation of mood effects on processing style, Storbeck and Clore (2005) demonstrated that happy participants are more likely than sad participants to produce false memories. Drawing on Roediger and McDermott's (1995) paradigm, they presented participants with lists of words (bed, pillow, rest, etc.) that are highly associated with a critical lure (e.g., sleep). Participants typically recall having seen the critical lures, even though they were never presented. This effect depends on engaging in gist processing (in addition to item-specific processing) and is hence more pronounced under happy than sad moods, reflecting the differences in processing style discussed previously. Additional manipulations located this effect at the encoding rather than at the retrieval stage, as theoretically expected.

The lion's share of research into mood and memory, however, has addressed mood effects on recall in the context of Bower's (1981) model, reviewed earlier. This model predicts that material learned while in a given affective state should be better recalled when in the same rather than a different state. Support for this *state-dependent recall* hypothesis has been obtained in several studies that used a "two-list interference paradigm," where list A is learned while in a happy mood and list B is learned while in a sad mood. Being in the same mood at the time of recall facilitated recall in several studies, whereas being in the opposite mood inhibited recall

(e.g., Bower et al., 1978; Schare, Lisman, & Spear, 1984). However, other studies (e.g., Bower & Mayer, 1985; Marshall-Garcia & Beck, 1985) failed to replicate this pattern, even when using the same materials. Bower and Mayer (1985) concluded from their own non-replications that "mood-dependent retrieval is an evanescent will-o'-the-wisp, and not the robust outcome suggested by earlier reports" (p. 42).

A second prediction holds that positive moods facilitate recall of positively valenced material and inhibit recall of negatively valenced material; the reverse is expected for negative moods. Note that this *mood-congruency* prediction pertains to the match of mood at recall and valence of the to-be-recalled material, independent of the mood at the time of learning. Unfortunately, the conceptually straightforward distinction between state dependency and mood congruency is difficult to sustain in the domain that produced the most supportive findings, namely, the recall of autobiographical information (for reviews, see Blaney, 1986; Morris, 1989; Singer & Salovey, 1988), as discussed earlier. In addition, mood congruency may be limited to relatively unstructured material and tends to be difficult to find when positive and negative elements are closely interconnected in a narrative (Hasher, Rose, Zack, Sanft, & Doren, 1985; Mecklenbräuker & Hager, 1984), consistent with the logic of network models.

Empirically, participants' recall often shows a marked asymmetry (e.g., Natale & Hantas, 1982; see Blaney, 1986; Singer & Salovey, 1988, for reviews). Participants in a happy mood recall more happy, and fewer sad, memories than do participants in a neutral mood, indicating facilitative as well as inhibitive effects of happy moods. In contrast, sad participants recall fewer happy events, but not more sad events, suggesting inhibitive but not facilitative effects of sad moods. Three different accounts have been offered for this asymmetry. One proposal holds that positive material is more interconnected in memory than negative material (e.g., Isen, 1984; Matlin & Stang, 1979). If so, a given mood-related association would spread to a larger amount of similarly valenced material under positive rather than negative moods. However, data bearing directly on these structural assumptions are not available. Moreover, others (e.g., Higgins, Van Hook, & Dorfman, 1988) proposed that negative events are more likely to be interconnected in memory, given that they elicit more explanatory activity (Bohner et al., 1988). A second proposal attributes the observed asymmetry to mood repair efforts (Clark & Isen, 1982). According to this hypothesis, sad participants attempt to improve their mood by avoiding further negative thoughts. One might expect, however, that recalling happy memories is an even more effective strategy for mood repair, yet such mood-incongruent recall is rarely observed (for an exception, see Parrott & Sabini, 1990, who found mood-incongruent recall under happy as well as sad moods, in contrast to what the mood repair logic would predict).

As a third possibility, Schnall, Clore, and Ryan (2005) suggested that the usually obtained asymmetry may be due to mood induced differences in processing style.

Using the original materials of Bower, Gilligan, and Monteiro (1981), they crossed happy and sad moods with positive and negative conceptual primes. They observed that happy participants used the primes in recall, regardless of whether they were positive or negative, whereas sad participants did not. Given that happy moods are usually induced by positive conceptual content (either as a function of the mood manipulation or a naturally occurring event), reliance on this accessible conceptual content as a recall cue would produce a pattern of content-congruent recall that looks like mood-congruent recall. Because sad moods promote item-specific processing and discourage reliance on other accessible content, less congruency would be observed under this condition, fostering the familiar asymmetry.

Emotions

Research on the impact of specific emotions on memory has also been guided by the assumption that emotions should activate emotion-congruent material (Bower, 1981). Hence, emotion-congruent material should be easier to recall, more readily perceived, and more likely to interfere with competing material when one experiences the respective emotion. However, the available data, mostly pertaining to anxiety, do not provide strong support for the operation of some general form of emotion congruency (see Mathews & MacLeod, 1994, for an extensive review). Instead, emotions seem to elicit a focus on material that is content relevant rather than on material that is simply feeling consistent. For example, Mogg, Mathews, and Eysenck (1992) observed in an attentional paradigm that anxious participants were only faster in responding to threatening words when the word pertained to their specific domain of worry. Similarly, Mathews and Klug (1993) crossed the valence of a set of words with whether the content was or was not related to the concerns of anxious patients. Content-related words interfered more than did content-unrelated words—and did so regardless of their valence. Mathews and MacLeod (1994) therefore concluded, "It is the match with current domain of concern, rather than emotional valence or congruence in a general sense, that determines the information that is given processing priority" (p. 37). This conclusion is compatible with the assumption that feelings inform us about the current situation, directing our attention to features that are likely to make us anxious.

From Memory to Judgment

Social psychologists' interest in affective influences on memory is mostly motivated by the assumption that affect-related differences in memory mediate affect-related differences in judgment. To do so, the information that comes to mind must seem to be "about" the target of judgment (e.g., Higgins, 1998). If people are aware that it may only come to mind due to their current mood, for example, accessible thought content may be likely to be discounted, as discussed earlier. Surprisingly, this conjecture has not received direct testing.

WHAT HOLDS WHEN?

Consistent with goals of this handbook, our review of the interplay of feeling and thinking focused on basic principles that apply to more than one type of feeling. Each of the reviewed mechanisms can account for some, but not all, of the available data. Moreover, many of the process assumptions are not mutually exclusive and each one may hold under some conditions. In a commendable integrative effort, Forgas (1995a, 2001) proposed a multiprocess affect infusion model (AIM) that focuses on mood effects and does not address other feelings. The AIM incorporates the theoretical approaches discussed earlier and can accommodate any mood effect predicted by its component theories. The model's original contribution is an attempt to specify the conditions under which previously identified processes are likely to hold. While we agree with many AIM predictions, a selective discussion of some of its ambiguities suggests that an integrative conceptualization remains a challenging task.

The AIM distinguishes four different processing strategies. If the target is familiar and a previously formed judgment is accessible in memory, people are assumed to rely on a *direct access* strategy, provided that the judgment is not personally relevant. Mood is not assumed to play a role in this case. The prototypical example given (Forgas, 2001, Fig. 5.1) is stereotyping, which is assumed to reflect the recall of a previously formed impression of a group. But as reviewed earlier, people are more likely to rely on stereotypes when they are in a good rather than bad mood (e.g., Bodenhausen et al., 1994). Moreover, they are more likely to draw on a previously formed judgment when in a good mood, following a direct access strategy, but to form a new judgment based on currently accessible details when in a bad mood (Bless, Mackie, & Schwarz, 1992). Hence, moods influence the use of a direct access strategy in the first place.

As a second possibility, the AIM introduces a *motivated processing* strategy, which people may employ when they want to reach a certain conclusion, potentially in the interest of mood management goals. To arrive at the desired conclusion, they may engage in a selective information search, which may override mood effects on judgment. We assume that this strategy is less likely when the situation is perceived as problematic, again introducing affective influences.

A third possibility pertains to a *substantive processing* strategy. It is based on extensive memory search and elaboration, giving rise to the influence of mood-congruent recall. People are assumed to use this strategy primarily under conditions of unconstrained processing capacity and high accuracy motivation to form judgments that are demanding (as exemplified by atypical, unusual, or complex targets) and of some importance to them. Being in a sad mood is assumed to facilitate the adoption of this strategy via increased accuracy motivation but may impede its adoption via decreased cognitive capacity. The mood-congruency component of this strategy suggests that mood effects on judgment should follow the pattern of mood effects on recall. Empirically, this does not appear to be the case. The judgment effects

attributed to substantive processing show a largely symmetrical impact of happy and sad moods (see Forgas, 1992, for a review), whereas mood-congruent recall is mostly limited to happy participants in memory experiments, as reviewed earlier. The strategy further suggests, for example, that cognitive responses to a persuasive message should reflect mood-congruent elaboration. Empirically, this is not the case. Sad recipients generate more negative responses to weak arguments, but more positive responses to strong arguments, than do happy recipients (Schwarz, Bless, & Bohner, 1991), which is consistent with differential accuracy motivation but not with mood-congruent elaboration. More complex judgments, however, have often been observed to show stronger mood effects (e.g., Forgas, 1995b; Schwarz et al., 1987). But in the absence of attributional manipulations we cannot tell whether participants simplified a complex task by relying on their feelings as information or engaged in mood-congruent substantive processing. In support of substantive processing, Forgas (2001) emphasizes that participants spend more time perusing the information when it is complex but are subsequently fast in providing a judgment. This pattern is also compatible with the possibility that they ponder a complex task and opt for a heuristic shortcut once they realize how burdensome it would otherwise be. Similarly, the observation of mood-congruent recall *after* the judgment is made does not necessarily imply that it mediated the judgment. Instead, the previously formed judgment may itself serve as a cue in reconstructive memory (for an example, see Bless, 1996). As these conjectures indicate, process identification is riddled by uncertainties and diagnostic evidence is often unavailable.

Finally, the AIM's fourth possibility pertains to a *heuristic processing* strategy, based on one's current feelings as a source of information (e.g., Schwarz & Clore, 1983). People are assumed to use this strategy under conditions of limited processing capacity and low accuracy motivation to form judgments that are simple and/or of limited importance to them. From this perspective, any judgment that is susceptible to misattribution effects pertains, by definition, to simple targets of low importance, which is difficult to reconcile with the available evidence. Being in a good mood is predicted to facilitate the adoption of this processing strategy via decreased accuracy motivation, consistent with mood effects on processing style. Finally, the strategy is assumed to be "ineffective and dysfunctional" because "affect can only serve as a heuristic cue due to mistaken inferences" (Forgas, 2001, p. 104), a conclusion that ignores integral affect and the advantages conveyed by fast and efficient affective reactions to the environment (Frijda, 1988; Zajonc, 1980).

Despite such shortcomings, the AIM predicts the correct outcomes more often than not, which makes it a useful guide for considering the possible influence of moods in many applied contexts. This the case because the AIM can accommodate any of the results predicted by its component theories and because different process assumptions often converge on the same outcome prediction, as the next section illustrates.

FEELINGS AND BEHAVIOR

To appreciate how different mood-related processes can result in the same outcome, suppose that Jane encounters an opportunity to help Mary. Jane may consider how much she likes Mary, whether she has the resources to help her, whether that experience would be pleasant, or how much good her help might do. If she applies the "How-do-I-feel-about-it?" heuristic to any of these questions, she will arrive at more positive assessments when in a good rather than bad mood, making helping more likely. If her moods influence what comes to mind, her elaborations will be more positive when in a good rather than bad mood, again making helping more likely. If Mary is an unknown other, a good mood may facilitate her inclusion in the ingroup through broader categorization, again making helping more likely. If the helping task is not very demanding, it may also provide Jane with a good opportunity to maintain her pleasant mood. Hence, the feelings-as-information, mood-congruent recall, and mood maintenance approaches make the same prediction for positive moods, unless the helping task is highly demanding, which would render it an unlikely avenue for mood maintenance (Schaller & Cialdini, 1990). Moreover, only the mood maintenance approach differs in the predictions for negative moods, as helping may provide an opportunity to improve a bad mood, provided the benevolent act is not too costly. Not surprisingly, positive moods are, indeed, reliably related to prosocial behavior, whereas "the effect of negative moods on benevolence is less consistent," as Eisenberg, Losoya, and Spinrad (2003, p. 797) concluded after a comprehensive review. This overdetermined nature of many mood effects precludes inferences about the underlying processes in the absence of additional process information.

Consistent with social psychologists' preferred theorizing, all influences of moods on prosocial behavior were mediated by cognition or motivation in the foregoing example. In contrast, many emotion theorists (e.g., Frijda, 1988; Leventhal, 1982) believe that emotions affect behavior in more direct ways. For example, fear is assumed to involve behavioral tendencies to escape, anger to involve activation of aggressive responses, and so on. In our reading, the direct effects of emotions are more likely to be motivational, changing the accessibility and priority of goals. The likely goals of fearful or angry persons, for example, are much easier to predict than the likely behaviors. Fear clearly involves a desire to avoid harm or loss, but from knowing only that they are afraid, we cannot predict whether people will sell their stocks, listen to the weather report, or start running. The immediate effects of emotion, therefore, are more mental than behavioral, emphasizing the importance of the processes that were the focus of this chapter.

CONCLUDING REMARKS

The proliferation of research into the interplay of feeling and thinking has resulted in a multitude of findings. Most of them can be framed in terms of two global approaches: an experiential approach that focuses on the informational value of subjective experiences, which include moods, emotions, bodily sensations, and cognitive experiences; and a cognitive approach that focuses on the impact of moods and emotions on the content of the thoughts that come to mind rather than the experience of having the thoughts. Each of these approaches is supported by a number of unique findings, whereas many other findings are compatible with both. For the latter, diagnostic process evidence that would convince advocates of the respective other approach is often missing (as a comparison of this chapter and Forgas, 2003, will aptly illustrate). While researchers working within each approach are able to produce many of the core effects with considerable reliability, the interplay of the underlying processes is awaiting an encompassing conceptualization. We hope that our accentuation of the principles underlying each approach will help in tackling this formidable task.

REFERENCES

Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica*, *103*, 77–89.

Adaval, R. (2001). Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information. *Journal of Consumer Research*, *28*, 1–17.

Adelmann, P. K., & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, *40*, 249–280.

Allport, F. H., & Lepkin, M. (1945). Wartime rumors of waste and special privilege: Why some people believe them. *Journal of Abnormal and Social Psychology*, *40*, 3–36.

Avnet, T., & Pham, M. T. (2004). *"Should I trust my feelings or not?" The metacognition of affect as information in judgment.* Unpublished manuscript, Columbia University, New York.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*, 446–458.

Beukeboom, C., & Semin, G. R. (2005). Mood and representations of behavior: The how and why. *Cognition and Emotion*, *19*, 1242–1251.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*, 229–246.

Bless, H. (1996). Die informative Funktion von Stimmungen: Auswirkungen auf evaluative Urteile, Verarbeitungsstil und Gedächtnis. [Informative functions of moods: Effects on judgment, processing style, and memory.] In E. H. Witte (Ed.), *Soziale Kognition und empirische Ethikforschung* [Social cognition and empirical ethics research] (pp. 27–45). Lengerich, Germany: Pabst Science.

Bless, H. (1997). *Stimmung und Denken [Mood and reasoning].* Bern, Switzerland: Huber.

Bless, H., Bohner, G., Schwarz, N., & Strack, F. (1990). Mood and persuasion: A cognitive response analysis. *Personality and Social Psychology Bulletin*, *16*, 331–345.

Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wölk, M. (1996). Mood and the use of scripts: Does being in a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology*, *71*, 665–679.

Bless, H., Mackie, D. M., & Schwarz, N. (1992). Mood effects on encoding and judgmental processes in persuasion. *Journal of Personality and Social Psychology*, *63*, 585–595.

Bless, H., & Schwarz, N. (1999). Sufficient and necessary conditions

in dual-process models: The case of mood and information processing. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 423–440). New York: Guilford Press.

Bless, H., Schwarz, N., & Kemmelmeier, M. (1996). Mood and stereotyping: The impact of moods on the use of general knowledge structures. *European Review of Social Psychology*, *7*, 63–93.

Bodenhausen, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology*, *66*, 621–632.

Bohner, G., Bless, H., Schwarz, N., & Strack, F. (1988). What triggers causal attributions? The impact of valence and subjective probability. *European Journal of Social Psychology*, *18*, 335–345.

Bollnow, O. F. (1956). *Das Wesen der Stimmungen*. Frankfurt, Germany: Klostermann.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Bower, G. H., Gilligan, S. G., & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General*, *110*, 451–473.

Bower, G. H., & Mayer, J. D. (1985). Failure to replicate mood congruent retrieval. *Bulletin of the Psychonomic Society*, *23*, 39–42.

Bower, G. H., Monteiro, K. P., & Gilligan, S. G. (1978). Emotional mood as a context of learning and recall. *Journal of Verbal Learning and Verbal Behavior*, *17*, 573–585.

Brewer, W. F. (1992). Phenomenal experience in laboratory and autobiographical memory. In M. A. Conway, D. C. Rubin, H. Spinner, & W. A. Wagenaar (Eds.), *Theoretical perspectives in autobiographical memory* (pp. 31–51). Dordrecht, The Netherlands: Kluwer.

Broadbent, D. E. (1971). *Decision and stress*. London: Academic Press.

Clark, M. S., & Isen, A. M. (1982). Towards understanding the relationship between feeling states and social behavior. In A. H. Hastorf & A. M. Isen (Eds.), *Cognitive social psychology* (pp. 73–108). New York: Elsevier.

Clore, G. L. (1992). Cognitive phenomenology: Feelings and the construction of judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 133–164). Hillsdale, NJ: Erlbaum.

Clore, G. L., & Colcombe, S. (2003). The parallel worlds of affective concepts and feelings. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 335–369). Mahwah, NJ: Erlbaum.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 323–418). Hillsdale, NJ: Erlbaum.

Clore, G. L., Wyer, R. S., Jr., Dienes, B., Gasper, K., Gohm, C., & Isbell, L. (2001). Affective feelings as feedback: Some cognitive consequences. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 63–84). Mahwah, NJ: Erlbaum

Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Grosset/Putnam

Darwin, C. (1965). *The expression of the emotions in man and animals*. London: J. Murray. (Original work published 1872)

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). The effect of emotion on automatic intergroup attitudes. *Psychological Science*, *15*, 319–324.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183–201.

Edwards, J. A., & Weary, G. (1993). Depression and the impression-formation continuum: Piecemeal processing despite the availability of category information. *Journal of Personality and Social Psychology*, *64*, 636–645.

Eisenberg, N., Losoya, S., & Spinrad, T. (2003). Affect and prosocial responding. In R. J. Davidson, K. R. Scherer, & H. H.

Goldsmith (Eds.), *Handbook of affective sciences* (pp. 787–803). New York: Oxford University Press.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*, 1208–1210.

Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 25–42). Toronto: Hogrefe.

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572–595). New York: Oxford University Press.

Erber, R., & Erber, M. W. (2001). Mood and processing: A view from a self-regulation perspective. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 63–84). Mahwah, NJ: Erlbaum.

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*, 421–35.

Feldman Barrett, L., & Salovey, P. (Eds.). (2002). *The wisdom in feeling: Psychological processes in emotional intelligence*. New York: Guilford Press.

Fiedler, K. (1988). Emotional mood, cognitive style, and behavior regulation. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 100–119). Toronto: Hogrefe International.

Fiedler, K., Pampe, H., & Scherf, U. (1986). Mood an memory for tightly organized social information. *European Journal of Social Psychology*, *16*, 149–164.

Forgas, J. P. (1992). Affect in social judgments and decisions: A multi-process model. *Advances in Experimental Social Psychology*, *25*, 227–275.

Forgas, J. P. (1995a). Emotion in social judgments: Review and a new affect infusion model (AIM). *Psychological Bulletin*, *117*, 39–66.

Forgas, J. P. (1995b). Strange couples: Mood effects on judgments and memory about prototypical and atypical targets. *Personality and Social Psychology Bulletin*, *21*, 747–765.

Forgas, J. P. (2001). The affect infusion model (AIM): An integrative theory of mood effects on cognition and judgment. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 99–134). Mahwah, NJ: Erlbaum.

Forgas, J. P. (2003). Affective influences on attitudes and judgments. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 596–618). New York: Oxford University Press.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*, 218–226.

Freitas, A. L., Azizan, A., Travers, S., & Berry, T. A. (2005). The evaluative connotation of processing fluency: Inherently positive or moderated by motivational context? *Journal of Experimental Social Psychology*, *41*, 636–644.

Friedman, R. S., & Förster, J. (2000). The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, *79*, 477–492.

Friedman, R. S., & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology*, *38*(1), 41–55.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*, 349–358.

Gallagher, D., & Clore, G. L. (1985, May). *Effects of fear and anger on judgments of risk and blame*. Paper presented at the meetings of the Midwestern Psychological Association, Chicago.

Garcia-Marques, T., & Mackie, D. M. (2001). The feeling of familiarity as a regulator of persuasive processing. *Social Cognition*, *19*, 9–34.

Gasper, K. (2004). Do you see what I see? Affect and visual information processing. *Cognition and Emotion*, *18*, 405–421.

Gasper, K., & Clore, G. L. (2002). Attending to the big picture:

Mood and global vs. local processing of visual information. *Psychological Science*, *13*, 34–40.

Gorn, G. J., Goldberg, M. E., & Basu, K. (1993). Mood, awareness, and product evaluation. *Journal of Consumer Psychology*, *2*(3), 237–256.

Grayson, C. E., & Schwarz, N. (1999). Beliefs influence information processing strategies: Declarative and experiential information in risk assessment. *Social Cognition*, *17*, 1–18.

Haddock, G., Rothman, A. J., Reber, R., & Schwarz, N. (1999). Forming judgments of attitude certainty, importance, and intensity: The role of subjective experiences. *Personality and Social Psychology Bulletin*, *25*, 771–782.

Han, S., Lerner, J. S., & Keltner, D. (in press). Feelings and consumer decision making: The appraisal-tendency framework. *Journal of Consumer Psychology*.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, *16*, 107–112.

Hasher, L., Rose, K. C., Zacks, R. T., Sanft, H., & Doren, B. (1985). Mood, recall, and selectivity in normal college students. *Journal of Experimental Psychology: General*, *114*, 104–118.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Higgins, E. T. (1998). The aboutness principle: A pervasive influence on human inference. *Social Cognition*, *16*, 173–198.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Higgins, E. T., Van Hook, E., & Dorfman, D. (1988). Do self-attributes form a cognitive structure? *Social Cognition*, *6*, 177–206.

Hildebrand-Saints, L., & Weary, G. (1989). Depression and social information gathering. *Personality and Social Psychology Bulletin*, *15*, 150–160.

Hirt, E. R., Levine, G. M., McDonald, H. E., Melton, R. J., & Martin, L. L. (1997). The role of mood in quantitative and qualitative aspects of performance: Single or multiple mechanisms? *Journal of Experimental Social Psychology*, *33*, 602–629.

Isbell, L. M., & Wyer, R. S., Jr. (1999). Correcting for mood-induced bias in the evaluation of political candidates: The roles of intrinsic and extrinsic motivation. *Personality and Social Psychology Bulletin*, *25*, 237–249.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–236). Hillsdale, NJ: Erlbaum.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. *Advances in Experimental Social Psychology*, *20*, 203–253.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology*, *47*, 1206–1217.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*, 1122–1131.

Isen, A. M., Shalker, T. E., Clark, M. S., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology*, *36*, 1–12.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.

Jacoby, L. L., & Kelley, C. M. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica*, *98*, 127–140.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*, 326–338.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126–135.

James, W. (1890). *Principles of psychology* (Vols. 1 & 2). New York: Dover.

Johnson, E., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*, 20–31.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Kelley, C. M., & Rhodes, M. G. (2002). Making sense and nonsense of experience: Attributions in memory and judgment. *Psychology of Learning and Motivation*, *41*, 293–320.

Kelley, H. H. (1972). *Causal schemata and the attribution process*. Morristown, NJ: General Learning Press.

Keltner, D., Ellsworth, P., & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology*, *64*, 740–752.

Keltner, D., Locke, K. D., & Audrain, P. C. (1993). The influence of attributions on the relevance of negative feelings to satisfaction. *Personality and Social Psychology Bulletin*, *19*, 21–30.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–537.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford Press.

Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science*, *14*, 144–150.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: A model of emotion-specific influences on judgment and choice. *Cognition and Emotion*, *14*, 473–493.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings. Carryover effects of emotions on economic decisions. *Psychological Science*, *15*, 337–341.

Leventhal, H. (1982). The integration of emotion and cognition: A view from the perceptual-motor theory of emotion. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition* (pp. 121–156.) Hillsdale, NJ: Erlbaum.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Mackie, D. M., & Worth, L. T. (1989). Cognitive deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology*, *57*, 27–40.

Mackie, D. M., & Worth, L. T. (1991). Feeling good, but not thinking straight: The impact of positive mood on persuasion. In J. Forgas (Ed.), *Emotion and social judgment* (pp. 201–220). Oxford, UK: Pergamon Press.

Marshall-Garcia, K. A., & Beck, R. C. (1985). Mood and recognition memory: A comparison of two procedures. *Bulletin of the Psychonomic Society*, *23*, 450–452.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493–504.

Martin, L. L., Abend, T., Sedikides, C., & Green, J. D. (1997). How would it feel if . . . ?: Mood as input to a role fulfillment evaluation process. *Journal of Personality and Social Psychology*, *73*, 242–253.

Martin, L. L., & Clore, G. L. (Eds.). (2001). *Theories of mood and cognition: A user's guidebook*. Mahwah, NJ: Erlbaum

Martin, L. L., & Tesser, A. (1989). Toward a motivational and structural theory of ruminative thought. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 306–326). New York: Guilford Press.

Martin, L. L., Ward, D. W., Achée, J. W., & Wyer, R. S., Jr. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, *64*, 317–326.

Mathews, A. M., & Klug, F. (1993). Emotionality and interference

with color-naming in anxiety. *Behavior Research and Therapy*, *31*, 57–62.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology*, *45*, 25–50.

Matlin, M. W., & Stang, D. (1979). *The Pollyanna Principle: Selectivity in language, memory, and thought*. Cambridge, MA: Shenkman.

Mayer, J. D., Gaschke, Y. N., Braverman, D. L., & Evans, T. W. (1992). Mood-congruent recall is a general effect. *Journal of Personality and Social Psychology*, *63*, 119–132.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, *11*, 424–428.

Mecklenbräuker, S., & Hager, W. (1984). Effects of mood on memory: Experimental tests of a mood-state-dependent retrieval hypothesis and of a mood-congruity hypothesis. *Psychological Research*, *46*, 335–376.

Melton, R. J. (1995). The role of positive affect in syllogism performance. *Personality and Social Psychology Bulletin*, *21*, 788–794.

Mogg, K., Mathews, A. M., & Eysenck, M. (1992). Attentional bias to threat in clinical anxiety. *Cognition and Emotion*, *6*, 149–159.

Monin, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality and Social Psychology*, *85*, 1035–1048.

Morris, W. N. (1989). *Mood: The frame of mind*. New York: Springer-Verlag.

Natale, M., & Hantas, M. (1982). Effect of temporary mood states on selective memory about the self. *Journal of Personality and Social Psychology*, *42*, 927–934.

Olson, J. M., & Roese, N. J. (1995). The perceived funniness of humorous stimuli. *Personality and Social Psychology Bulletin*, *21*, 908–918.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.

Ottati, V. C., & Isbell, L. M. (1996). Effects of mood during exposure to target information on subsequently reported judgments: An on-line model of misattribution and correction. *Journal of Personality and Social Psychology*, *71*, 39–53.

Ottati, V., Terkildsen, N., & Hubbard, C. (1997). Happy faces elicit heuristic processing in a televised impression formation task: A cognitive tuning account. *Personality and Social Psychology Bulletin*, *23*, 1144–1156.

Parrott, W. G., & Sabini, J. (1990). Mood and memory under natural conditions: Evidence for mood incongruent recall. *Journal of Personality and Social Psychology*, *59*, 321–336.

Petty, R., & Cacioppo, J. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Pham, M. T. (1998). Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, *25*(2), 144.

Pham, M. T. (2004). The logic of feeling. *Journal of Consumer Psychology*, *14*, 360–369.

Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research*, *28*, 167–188.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, *79*, 56–77.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338–342.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382.

Reber, R., Winkielman P., & Schwarz N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*, 45–48.

Rhodes, M. G., & Kelley, C. M. (2003). The ring of familiarity: False

familiarity due to rhyming primes in item and associative recognition. *Journal of Memory and Language*, *48*, 581–595.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76–100.

Roediger, H., & McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803–814.

Ross, M. (1989). The relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, *24*, 1053–1064.

Ruder, M., & Bless, H. (2003). Mood and the reliance on the ease of retrieval heuristic. *Journal of Personality and Social Psychology*, *85*, 20–32.

Sanna, L., & Schwarz, N. (2004). Integrating temporal biases: The interplay of focal thoughts and accessibility experiences. *Psychological Science*, *17*, 474–481.

Sanna, L., Schwarz, N., & Small, E. (2002). Accessibility experiences and the hindsight bias: I-knew-it-all-along versus It-could-never-have-happened. *Memory and Cognition*, *30*, 1288–1296.

Sanna, L. J., Schwarz, N., & Stocker, S. L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, Cognition*, *28*, 497–502.

Savitsky, K., Medvec, V. H., Charlton, A. E., & Gilovich, T. (1998). What, me worry? Arousal, misattribution, and the effect of temporal distance on confidence. *Personality and Social Psychology Bulletin*, *24*, 529–536.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, a physiological determinants of emotional state. *Psychological Review*, *69*, 379–399.

Schaller, M., & Cialdini, R. (1990). Happiness, sadness, and helping: A motivational integration. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 265–296). New York: Guilford Press.

Schank, R., & Abelson, R. (1977) *Scripts, plans, goals and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Schare, M. L., Lisman, S. A., & Spear, N. E. (1984). The effects of mood variation on state-dependent retention. *Cognitive Therapy and Research*, *8*, 387–408.

Schnall, S., Clore, G. L., & Ryan, K. (2005). *Enacted affect-as-information: How expressive behaviors constrain the recall of affective events*. Unpublished manuscript, University of Virginia.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 527–561). New York: Guilford Press.

Schwarz, N. (2001). Feelings as information. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 159–176). Mahwah, NJ: Erlbaum

Schwarz, N. (2002). Situated cognition and the wisdom in feelings: Cognitive tuning. In L. Feldman Barrett & P. Salovey (Eds.), *The wisdom in feelings: Psychological processes in emotional intelligence* (pp.144–166). New York: Guilford Press.

Schwarz, N. (2004). Meta-cognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*, 332–348.

Schwarz, N., & Bless, B. (1991). Happy and mindless, but sad and smart?: The impact of affective states on analytic reasoning. In J. Forgas (Ed.), *Emotion and social judgment* (pp. 55–71). Oxford, UK: Pergamon Press.

Schwarz, N., Bless, H., & Bohner, G. (1991). Mood and persuasion: Affective states influence the processing of persuasive communications. *Advances in Experimental Social Psychology*, *24*, 161–199.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: An-

other look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195–202.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? Informative functions of affective states. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 44–62). Toronto: Hogrefe International.

Schwarz, N., Sanna, L., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology, 39*, 127–161.

Schwarz, N., Servay, W., & Kumpf, M. (1985). Attribution of arousal as a mediator of the effectiveness of fear-arousing communications. *Journal of Applied Social Psychology*, *15*, 74–78.

Schwarz, N., & Skurnik, I. (2003). Feeling and thinking: Implications for problem solving. In J. Davidson & R. Sternberg (Eds.), *The nature of problem solving* (pp. 263–292). Cambridge, UK: Cambridge University Press.

Schwarz, N., Strack, F., Kommer, D., & Wagner, D. (1987). Soccer, rooms and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific life-domains. *European Journal of Social Psychology*, *17*, 69–79.

Seamon, J. G., Brody, N., & Kauff, D. M. (1983). Affective discrimination of stimuli that are not recognized: Effects of shadowing, masking, and central laterality. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 544–555.

Sedikides, C. (1995). Central and peripheral self-conceptions are differentially influenced by mood. *Journal of Personality and Social Psychology*, *69*, 759–777.

Siemer, M., & Reisenzein, R. (1998). Effects of mood on evaluative judgements: Influence of reduced processing capacity and mood salience. *Cognition and Emotion*, *12*, 783–805.

Sinclair, R. C., Mark, M. M., & Clore, G. L. (1994). Mood-related persuasion depends on misattributions. *Social Cognition*, *12*, 309–326.

Sinclair, R. C., Soldat, A. S., & Mark, M. M. (1998). Affective cues and processing strategy: Color coded forms influence performance. *Teaching of Psychology*, *25*, 130–132.

Singer, J. A., & Salovey, P. (1988). Mood and memory: Evaluating the network theory of affect. *Clinical Psychology Review*, *8*, 211–251.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*, 713–724.

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.

Soldat, A. S., & Sinclair, R. C. (2001). Colors, smiles, and frowns: External affective cues can directly affect responses to persuasive communications in a mood-like manner without affecting mood. *Social Cognition*, *15*, 55–71.

Soldat, A. S., Sinclair, R. C., & Mark, M. M. (1997). Color as an environmental processing cue: External affective cues can directly affect processing strategy without affecting mood. *Social Cognition*, *15*, 55–71.

Spies, K., & Hesse, F. W. (1986). Interaktion von Emotion und Kognition [Interaction of emotion and cognition]. *Psychologische Rundschau*, *37*, 75–90.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, *64*, 211–220.

Storbeck, J., & Clore, G. L. (2005). With sadness comes accuracy, with happiness, false memory: Mood and the false memory effect. *Psychological Science*, *16*, 785–791.

Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 579–595). New York: Guilford Press.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A non-obtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *53*, 768–777.

Strack, F., & Neumann, R. (2000). Furrowing the brow may undermine perceived fame: The role of facial feedback in judgments of celebrity. *Personality and Social Psychology Bulletin*, *26*, 762–768.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology*, *23*, 53–62.

Tamir, M., Robinson, M. D., Clore, G. L., Martin, L. L., & Whitaker, D. J. (2004). Are we puppets on a string?: The contextual meaning of unconscious expressive cues. *Personality and Social Psychology Bulletin*, *30*, 237–249.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*, 973–988.

Tomkins, S. S. (1962). *Affect, imagery, and consciousness* (Vol. 1). New York: Springer-Verlag.

Tulving, E. (1989). Memory: Performance, knowledge, and experience. *European Journal of Cognitive Psychology*, *1*, 3–26.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.

Valins, S. (1974). Persistent effects of information about internal reactions. In H. London & R. E. Nisbett (Eds.), *Thought and feeling. Cognitive alteration of feeling states* (pp. 116–125). Chicago: Aldine.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin*, *22*, 1105–1113.

Wänke, M., Bohner, G., & Jurkowitsch, A. (1997). There are many reasons to drive a BMW—Surely you know one: Ease of argument generation influences brand attitudes. *Journal of Consumer Research*, *24*, 70–77.

Weary, G., Marsh, K. L., Gleicher, F., & Edwards, J. A. (1993). Social-cognitive consequences of depression. In G. Weary, F. Gleicher, & K. L. Marsh (Eds.), *Control motivation and social cognition* (pp. 255–287). New York: Springer-Verlag.

Weaver, K., Garcia, S. M., Schwarz, N., & Miller, D. T. (in press). Inferring the popularity of an opinion from its familiarity. *Journal of Personality and Social Psychology*.

Wegener, D. T., & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality and Social Psychology*, *66*, 1034–1048.

Wegener, D. T., Petty, R. E., & Smith, D. M. (1995). Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology*, *69*, 5–15.

Wegner, D. M., & Vallacher, R. R. (1986). Action identification. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 1, pp. 550–581). New York: Guilford Press.

Wenzlaff, R. M., Wegner, D. M., & Roper, D. (1988). Depression and mental control: The resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology*, *55*, 882–892.

Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1235–1253.

Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, *29*, 716–732.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and

mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117–142.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Erlbaum.

Winkielman, P., Schwarz, N., & Nowak, A. (2002). Affect and processing dynamics. In S. Moore & M. Oaksford (Eds.), *Emotional cognition* (pp. 111–138). Amsterdam: Benjamins.

Witherspoon, D., & Allan, L. G. (1985). The effects of a prior presentation on temporal judgments in a perceptual identification task. *Memory and Cognition*, *13*, 103–111.

Worth, L. T., & Mackie, D. M. (1987). Cognitive mediation of positive mood in persuasion. *Social Cognition*, *5*, 76–94.

Wyer, R. S., Jr., & Carlston, D. (1979). *Social cognition, inference, and attribution*. Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., Clore, G. L., & Isbell, L. (1999). Affect and information processing. *Advances in Experimental Social Psychology*, *31*, 3–78.

Wyer, R. S., Jr., & Srull, T. K. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Erlbaum.

Xu, J., & Schwarz, N. (2005, February). *Was it long ago or unimportant?: Diverging inferences from difficulty of recall*. Paper presented at the meeting of the Society for Consumer Psychology, St. Petersburg Beach, FL.

Yeung, C. W. M., & Wyer, R. S., Jr. (2004). Affect, appraisal, and consumer judgment. *Journal of Consumer Research*, *31*, 412–424.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph*, *9*(2), 1–27.

Zajonc, R. B. (1980). Feeling and thinking. Preferences need no inferences. *American Psychologist*, *35*, 151–175.

Zajonc, R. B., & Markus, H. (1984). Affect and cognition: The hard interface. In C. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotions, cognition and behavior* (pp. 73–102). Cambridge, UK: Cambridge University Press.

Zajonc, R. B., Murphy, S. T., & Inglehart, M. (1989). Feeling and facial efference: Implications of the vascular theory of emotion. *Psychological Review*, *96*, 395–416.

Zanna, M. P., & Cooper, J. (1976). Dissonance and the attribution process. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 199–217). Hillsdale, NJ: Erlbaum.

Zillman, D. (1978). Attribution and misattribution of excitatory reactions. In J. H. Harvey, W. I. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 2, pp. 335–368). Hillsdale, NJ: Erlbaum.

The Role of Impulse in Social Behavior

FRITZ STRACK
ROLAND DEUTSCH

THE TWO HORSES OF HUMAN BEHAVIOR

One of the most famous metaphors in philosophy is Plato's model of the soul in terms of a chariot with his two horses (Plato, 1998). They are vividly described in his dialogue "Phaedrus" where a chariot is drawn by a beautiful and upright white horse that needs no whip and is guided only by the charioteer's commands. The second horse, in contrast, is dark in color, ugly, crooked, and hard of hearing. It is scarcely controlled with a combination of whip and goad. Despite their different character, the two horses are capable of jointly and forcefully moving the vehicle forward. However, it may also be the case that one horse will be dominant and determine the course of the chariot. Of course, this picture aptly symbolizes the two forces of human behavior. The white horse depicts reason and reflection, the dark horse passion and impulse.

The metaphor of the two horses is present in many models of human behavior. In some of them, one of the two horses is portrayed as the *dominant* driving power while the second horse is following suit. Other models, however, focus on a potential *conflict* between the two horses. Plato himself had hoped that the white horse of reason was the most frequent force behind human action. Much later, this dominance was elaborated in the so-called rational model (Hume, 1739/2000) that is still the conceptual basis of economics (e.g., Becker, 1962) and part of the social and psychological sciences (e.g., Fishbein & Ajzen, 1975).

More important for present purposes, the metaphor of the chariot and the horses is a first attempt at describing the dual determination of human behavior. Moreover, it resembles its conceptual successors in that one determinant stands in the foreground and is explained in considerable detail while the other is relegated to a secondary status and is described predominantly by negating the characteristics of the primary determinant. This chapter is meant to focus on a type of behavioral determination that has received little attention from research in social psychology. In particular, we would like to shed some light on the nature of impulse and to identify some of the mechanisms that generate impulsive behaviors. Interestingly, such behaviors are often described by using negations, such as "not caused by intention, judgment or decision," "irresistible," "irrational," "irresponsible," and "undifferentiated" and are often accompanied or driven by strong affect (e.g., Le Bon, 1895). Therefore, we begin by considering those nonimpulsive accounts that serve as a reference point for impulsive behavior and point out their shortcomings.

RATIONAL MODELS OF HUMAN BEHAVIOR

In most of its variants, the rational model holds that behavior is a consequence of people's reflective thoughts and not of their feelings. In addition, there are certain contents that are in its focus. According to the rational model, actors are assumed to anticipate and evaluate the

consequences of what they do. As a result, behavior is construed as a decision that is based on what people expect as a result of what they do and on their beliefs if this is good or bad. For the Greek philosophers, these values were derived from basic virtues (like wisdom, courage, and moderation). Moreover, Socrates assumed that aberrations from the path of virtue were merely due to a lack of knowledge. In other words, if actors knew what is good or bad, their actions would necessarily fall in line.

After more than 2000 years, Socrates's model was extended by Adam Smith (1776/1981), the founder of modern economics. He has argued that there is no need to invoke fundamental virtues in order to derive the values that guide behavior. Instead, he believed that recognizing and pursuing one's own interest will be sufficient to maximize the common good. More recently, economists (e.g., Becker, 1976) have proposed that *all* human behavior is caused by decisions that maximize a person's advantage (or, utility). And as in Socrates's notion, seeming exceptions to this rule were understood as the result of incomplete information.

In modern social psychology, the "rational model" has also influenced theorizing. A prominent example is Fishbein and Ajzen's (1975) theory of reasoned action in which people's attitudes toward a behavioral outcome determine what people decide to do. Once again, reflection is the fundamental mechanism. Elements of the rational model can also be found in Bandura's (1977) theory of social learning, which explains human behavior as a function of its anticipated consequences. To predict what people do, it is necessary to understand how they acquire knowledge about what is good or bad. Bandura has described various mechanisms, among them observing others' behaviors.

ABERRATIONS FROM THE PATH OF REASON: IMPULSIVE DETERMINANTS OF HUMAN BEHAVIOR

The Roots of Impulse in Greek Philosophy and in the Christian Tradition

While for the rational theories of human behavior, the white horse is clearly in charge, its dark companion may either trot along or help to pull the carriage into the predetermined direction. Psychologically speaking, human behavior was seen as the result of reflection, while impulsive mechanisms had no major role beyond energizing the execution of reflective decisions.

However, this Socratic view of human behavior did not go undisputed. In Western philosophy, Plato and Aristotle were the first to recognize the unique quality of impulsive determinants of behavior. In particular, passion was identified as the force that would lead people astray. In marked contrast to the implications of the rational model, knowing what is virtuous or good was not enough. In fact, passion was seen to cause people to do what they knew was wrong. In general, behavior driven by passion was seen to be bad for one's own well-being.

Built on Aristotle's ethics (Bywater, 1984), the Christian tradition also proclaimed virtues as guides for behav-

ior (Aquinas, 2002). Deviations from the path of virtue were understood as violations of God's will. Such sins are caused by an evil force with which humans are endowed. Although sinning was understood to be part of human nature, guilt had to be felt if a behavior was recognized to be inconsistent with religious values. Fortunately, however, the church offers to gracefully remove these unpleasant experiences from its faithful. From a psychological perspective, it is important to note that the Christian doctrine includes effective mechanisms of self-regulation that operate through the recognition of inconsistency, negative affect (guilt, shame, contrition, and regret) and institutions that are capable of resetting the person into his or her previous state. In return, recipients are expected to show faith in the redeemer and fight the forces that led them astray.

In Plato's metaphoric picture, the Christian view is a permanent struggle between reason and passion, or between reflection and impulse. The institutions of the church serve to strengthen the impact of reflective processes within the constraints that are provided by the religious values. Pleasant feelings that are associated with impulsive sinning are framed to elicit guilt and shame. Fortunately, these negative feelings can be prevented by engaging in religious exercises, such as praying and fasting. Alternatively, a pleasant hedonic experience may be purified by embedding it into a religiously accepted institution. This becomes particularly apparent in the case of sex, which must either be entirely abandoned by some representatives of the religion or is only allowed as part of reproduction within marriage.

The Roots of Impulse in the History of Psychology

In the history of psychology, Gustave Le Bon (1895) was perhaps the first to describe the perils of impulse. Moreover, he identified the eliciting circumstances and explained the underlying psychological mechanisms. Specifically, Le Bon asserted two states that cause the person to behave according to different psychological principles. When by themselves, individuals were assumed to act according to reason and responsibility; when in groups, they would lose their "conscious personality," and their behavior was seen as the result of impulses that are suggestively transmitted by other group members. Even more drastically, Le Bon depicted the social group to be the "slave of impulses" (*esclave des impulsions*) that are elicited by hypnotically suggestive influences, which only a "strong personality" would be able to resist. In detail, the major characteristics of group behavior were identified as impulsivity, irritability, lack of logical thinking, deficiency of critical judgment, and an exaggeration of feelings. These determinants of behavior along with the mechanism of hypnotic suggestion were localized in what Le Bon called the "group mind," which would lead to uncivilized manifestations of human behavior. Because they are deprived of the capacity of critical reasoning, individuals in social contexts would be the victims of social influences, and their uninhibited emotions would elicit destructive violence and aggression but also acts of exceptional bravery and sacrifice. According to Le Bon,

this is because social behavior, unlike non-social behavior, is not determined by the actors' personal interest but by the forces previously described.

In this first dual-process model of social behavior, Le Bon has directed the readers' attention to the possibility that under certain conditions, human behavior may not be guided by a reflected evaluation of its outcome but by impulsive forces that obey different rules and regularities. Moreover, such impulsive behaviors were not seen as exceptional deviations but as the normal manifestations of behavior in social situations. Thus, persons' actions were understood to be determined either by the forces of impulse or by those of reason. A joint influence of both operating principles was not part of Le Bon's model, and consequentially, there was no conflict between them.

Le Bon's account has not gone uncriticized. Most prominently, Sigmund Freud (1921) disapproved of Le Bon's concept of a group mind and of hypnotic suggestion as a mechanism of social influence. To explain the behavioral phenomena described by Le Bon, he proposed the model of a mental apparatus in which reflection and impulses were interrelated and in which conflicts between the two may arise. His approach was probably the first attempt at describing mental processes and behavior as the result of an interaction of different systems that operate on the basis of different psychological processes. As a result, human behavior has both reflective and impulsive components, the degree of which depends on the type of interaction between what Freud called the ego, the id, and the superego. In Freud's psychoanalytic model, the id was mainly located in the domain of the unconscious. Even elements of the conscious ego were given the potential to be "repressed" to the unconscious (Freud, 1920). Moreover, the ego uses defense mechanisms to deal with unconscious impulses and much of what the defense mechanisms do is unconscious, thereby producing clinical symptoms. Unlike Le Bon's model, Freud's theory was based on the assumption that inputs from the different aspects of the mental apparatus would *jointly* exert their influence on thinking, feeling and acting. Therefore, conflicts are preprogrammed and the target of therapeutic intervention, which consists of raising unconscious contents, including ego's defense mechanisms, to consciousness. Despite these differences, Freud and Le Bon agreed in their assumption that human behavior is not only determined by a rational anticipation of its outcome. In addition, nonreflective forces were seen to play an important role. Also, both theorists believed that intense affect and emotions would play a central role in undermining reflective control.

The Revival of Impulse in Modern Social Psychology

In many explanations in social psychology, the rational model plays an important role. This seems to be mainly due to the dominant role of attitudes as precursors of behavior. Specifically, the interpretation of attitudes as evaluations led to theoretical models in which behavior was conceived as a decision that was based on what was judged to be good or bad (Ajzen, 1991). Decisions were seen to be reached either by cognitive mechanisms that were modeled after normative procedures (e.g., Kelley, 1967) or by simplified mechanisms that take the bounded rationality (Simon, 1991) of human judgments into account (Nisbett & Ross, 1980; Tversky & Kahneman, 1974). Behavioral deviations from preceding judgments were seen to be punished by negative affect, which was then regulated by various processes of rationalization and justification (e.g., Festinger, 1957).

Despite this predominant orientation, there are a few lines of research in which behavior is explained in a way that is not oriented toward the rational model. One is Berkowitz's theory of aggression (e.g., Berkowitz, 1974) in which aggressive behavior is understood to be determined not by expectations of its outcome but by impulsive mechanisms. Negative affect in combination with the mere accessibility of violent thoughts were found to be sufficient to elicit aggression (see also Berkowitz, 2004).

Another relevant line of research comes from Mischel and his colleagues (e.g., Mischel, Ebbesen, & Zeiss, 1972) who have created an experimental paradigm in which children were required to choose between immediately consuming a small reward and waiting for a much larger reward. While the children typically preferred the larger outcome and started waiting for it, the majority of them interrupted the waiting period and forwent the larger reward for the smaller one. More recently, this inability to delay gratification was explained by Metcalfe and Mischel (1999) within a "hot/cool framework" in which a "cool system" specializes in thoughtful processing of knowledge whereas a "hot system" is in charge of emotional processing and reflexive operations. The authors list a number of features that describe the divergent types of processing in the two systems, identify conditions under which one of the systems assumes a dominant role, and describe ways in which the hot and the cool system interact in different psychological situations. Children's failure to wait for the reward was attributed to the predominance of the hot system that facilitated "reflexive" mechanisms that are driven by stimuli that are tangible in the situation.

Applied to adult behavior, the "irrational" preference for immediate rewards has been studied by economists under the name of "temporal discounting" (e.g., Frederick, Loewenstein, & O'Donoghue, 2002), which follows a hyperbolic utility function (Loewenstein & Elster, 1992). The underlying psychological processes, however, were described as "viscerally" driven and accompanied by the experience of "craving" (Loewenstein, 2001). Recently, Trope and colleagues (Liberman & Trope, 2003) tried to explain the hyperbolic nature of temporal discounting by pointing at the different mental construals of immediate and remote events and at the affective and judgmental consequences.

Most important, however, is a more recent development in the domain of social cognition that advanced the thesis that the controlled judgmental processes that were found to generate behavior had the potential to be executed in an automatic fashion. This notion is based on

the fact that a stimulus can cause behavior not only through generating an evaluative judgment and forming a specific goal but also through a direct and automatic connection to the behavior itself. This "perception–behavior link;" (Bargh, 1997) has its roots in William James's (1890; see also Lotze, 1852) "ideomotor principle" and the concept of "habit" from the behaviorist tradition (Thorndike, 1911; see also Bargh & Ferguson, 2000).

The possibility of a nonreflective determination of behavior has reemerged in modern social psychology mainly as the study of cognitive habituation. Specifically, it has been argued that the repetition of mental processes will allow them to be performed in an automatic fashion that is outside awareness, relatively effortless, unintentional, autonomous, and even involuntary or uncontrollable (Bargh, 1989). Moreover, such automatic processes may directly elicit behaviors without the actor being aware of their specific quality and meaning (Bargh & Ferguson, 2000; Oulette & Wood, 1998; Verplanken & Aarts, 1999) while affect is not seen to play a major role. Most important, perhaps, the reflective and automatically impulsive processes were seen to obey the same underlying mechanisms. In other words, it was deemed possible for *all* reflective operations to be delegated to automaticity if they are executed with a sufficient frequency. Because reflection is a scarce commodity, this flexibility was seen to have adaptive advantages. That is, if a behavior can be performed automatically, reflection may focus on new and more complex behavioral decisions.

VARIANTS OF IMPULSIVE DETERMINATION

This historic review suggests that although reflective determination of social behavior was regarded to be normal or typical, psychological theories have repeatedly pointed at alternative mechanisms. However, it becomes apparent that the nonreflective nature of such determinants provides no common mechanism that would underlie the behavior. For example, nonreflective processes may be driven by passion and affect, as Le Bon and Freud would assume, and would be described as "hot" in Metcalfe and Mischel's (1999) model and as "experiential" in Epstein's (e.g., 2003) conceptualization. Alternatively, they may also be seen as a function of the frequency and the recency with which different pieces of information co-occur (Verplanken, Aarts, & Van Knippenberg, 1997).

Driven by affect or habituation, a stimulus may elicit specific response. The same reaction, however, may also be triggered by a category to which this stimulus belongs. In the social domain, the activation of a stereotype may cause impulsive reactions (Bargh, 1997).

THE ROLE OF IMPULSE IN DIFFERENT AREAS OF PSYCHOLOGY

After having discussed some historic roots of impulsive determinants of social behavior, we turn now to areas of psychology in which both reflection and impulse play a major role in accounting for the phenomena in question.

Attitudes

In social psychology, attitudes are seen as precursors of behavior. Therefore, to know a person's evaluation is an efficient way of predicting this person's behavior. However, it is not clear if the influence of attitudes on behaviors takes a reflective or an impulsive path. This ambiguity is reflected by different definitions. Thurstone (1931), for instance, saw in an attitude the affect for or against a psychological object. A little later, Gordon Allport (1935) defined an attitude as a "mental or neural state of readiness" that exerts a "directive or dynamic influence on the individual's response" (p. 1) while Smith, Bruner, and White (1956) identified attitudes as predispositions that influence experience, motivation, and action toward a class of objects.

While these definitions point to impulsive influences, processes of reflection were in the focus of subsequent definitions. For example, Eagly and Chaiken (1993) described an attitude as "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (p. 1). It is important to note that authors who agree that evaluations (as opposed to experienced affect) are at the core of the attitude concept differ in whether these evaluations should be understood as declarative representations in memory (Pratkanis, 1989). For example, Schwarz (2000) proposed that attitudes should be understood as judgments that are not stored in memory but are generated on demand.

If this is the case, the question arises how the information that provides the basis for attitudinal judgments is accessed. Moreover, the nature of the internal representations of valence and their structure must be understood. Most important, the question arises if those internal evaluative representations may influence human behavior without mediating judgments.

Accessibility as a Determinant of Attitudinal Judgments

An important development in the study of attitudes psychology is the rise of social cognition (e.g., Bless, Fiedler, & Strack, 2004). Through the adoption of the information-processing paradigm, the mechanisms underlying attitude formation came into the focus of research. Among them were mechanisms whose execution occurs in an automatic fashion. One is the recall of relevant information from memory. Under the label of "accessibility" (Higgins, Rholes, & Jones, 1977), it has been demonstrated that the generation of judgments greatly depends on the likelihood with which a particular information comes to mind. This "activation potential" (Higgins, 1996) is influenced by the frequency and recency of a prior activation (see Srull & Wyer, 1979, 1980). Accessibility has also been shown to provide an important link between evaluative concepts that are stored in memory and the generation of evaluative judgments. It is impor-

tant to recognize that the activation of evaluative contents does not imply any endorsement of their application to the target. Still, this basic mechanism affects not only what people believe but also, as a consequence, how they behave (e.g., Fazio, Powell, & Williams, 1989). In other words, behavior is not only a function of an evaluative judgment but also a function of the probability with which the underlying information is brought to mind. Although the latter component occurs without awareness and is not controllable in a direct fashion, it is also subject to psychological laws.

Within the area of social cognition, automatic processes have come into the focus of investigation for the accessibility of information in evaluative (and nonevaluative) judgments. Also, inferential processes were claimed to consist partly of "spontaneous" (e.g., Newman & Uleman, 1989) and automatic mechanisms (Gilbert, 1989). Thus, the generation of attitudinal judgments was no longer understood as a domain of processes that are entirely under the person's conscious and intentional control. Instead, it was recognized that unconscious and automatic mechanisms play an important role. Primarily, experimental interventions to influence automatic processes consisted of priming procedures that were applied to increase the activation potential of a specific information (for a recent review, see DeCoster & Claypool, 2004). Thus, early research on automatic influences was predominantly focused on independent variables that affected judgments in an uncontrolled fashion while the dependent variables remained the same as in previous research. As a result, whereas nonreflective influences were identified as judgmental determinants, the last step to behavior is still reflective in nature. Consistent with this argument is the "judged usability" phase of knowledge activation and use (Higgins, 1996). However, subsequent study of nonreflective components of attitudes led to further discoveries of which two seem particularly important: the rise of "implicit attitudes" and the "direct link to behavior."

Implicit Attitudes

Whereas automaticity was originally studied by focusing on both its determinants and its effects on conventional expressions of attitudes, the focus of research has shifted more recently toward consequences that do not manifest themselves in measures such as self-reports. This was partly due to the insight that attitudes may not manifest themselves only in "explicit" evaluative expressions. Rather, and in line with earlier definitions of attitudes as "mental readiness" (Allport, 1935), they were found to be effective in an indirect or "implicit" fashion. Moreover, explicit/direct and implicit/indirect measures of attitudes were often unrelated and seem to tap into different evaluative representations. This reasoning has its roots in research by Patricia Devine (1989) on prejudice, which was found to be low if people were given the opportunity to explicitly endorse components of the negative stereotype about African Americans. However, if the same people were subtly primed with words referring to the nega-

tively evaluated group, an ambiguously described target person was subsequently more likely to be assigned the negative characteristics (e.g., hostility) of the stereotyped group. Moreover, this indirect influence did not depend on the degree to which the prejudice was endorsed (Devine, 1989).

This finding suggests that beyond prejudice, evaluations in general may have components that are not reflected in self-reports but may be captured by more indirect measures. Examples are "affective priming" (Fazio, 2001) and the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998), which has initiated a new era of attitude research. On a behavioral level, the IAT measures the ease of generating a categorical response as a function of its evaluative compatibility. For example, participants have to indicate if a name appearing on a computer screen is characteristic for Blacks or Whites by pressing the same (or a different) key that they have been using to indicate that a word is pleasant or unpleasant. Replicated across a great number of content domains, it has been found that it is more difficult (as reflected in longer response latencies and subjective reports) to assign a target to a category with a negative valence if the same response has been used to categorize another target as "positive" (and vice versa) than to assign a target to a category using a response that has been used for a judgment of the valence that is also a feature of the category. The difference in the ease of generating a categorical response for evaluatively compatible and incompatible responses became known as the IAT effect. It partly reflects the strength of evaluative associations and serves as perhaps the most widespread "implicit" measure of attitudes.

Because the IAT and other implicit measures (for a review, see Fazio & Olson, 2003) are only moderately correlated with self-reported attitudes (Nosek, Greenwald, & Banaji, 2005), the existence of implicit attitudes has been postulated. They were defined both operationally and in contrast to "explicit" attitudes. That is, implicit attitudes are what implicit tests measure, and this is different from the evaluations people report if they are asked to indicate their attitudes (e.g., Banaji, 2001; Wilson, Lindsey, & Schooler, 2000). More specifically, implicit attitudes are assigned the characteristics of automatic processing in that they involve components that are outside awareness, are unintentionally formed in a relatively effortless fashion, are autonomous, and are difficult to control.

Obviously, this reflects a swing in the conceptualization of attitudes toward Allport's (1935) definition. A "mental state of readiness" that exerts a "directive influence on the individual's response" describes a different link between attitude and behavior than an evaluation of the outcome of a particular behavior. However, in this definition, "attitudes" are defined by their behavioral consequences but not by a characteristic that leads to the behavior. If the behavioral determinants of attitudes are related to valence, the question arises how valence can be represented other than as a belief about what is good or bad. This is discussed at a later point in this chapter.

Concept–Behavior Links

Independent of evaluative components, impulsive mechanisms have been invoked to account for the impact of conceptual representations on behavior. For example, it has been demonstrated that without the person's awareness, activating a particular stereotype facilitates stereotype-consistent behaviors (Bargh, Chen, & Burrows, 1996). Similarly, it has been shown that the imitation of others' behaviors may occur outside the imitating person's awareness (Chartrand & Bargh, 1999). More recent evidence suggests that priming participants with a stereotype facilitates consistent reactions. Specifically, Payne (2001) found that flashing a photo of a Black person was more likely to categorize an object that was presented immediately after the photos as a handgun whereas flashing the picture of a White person facilitated the categorization of objects as tools. Moreover, additional analyses suggest that this tendency was based on automatic processes and not on controlled mechanisms. While the basic finding was replicated, new evidence suggests that this effect was not driven by a global facilitation of evaluatively consistent categorizations but by the specific contents of the stereotype (Correll, Park, Judd, & Wittenbrink, 2002; Judd, Blair, & Chapleau, 2004).

Derived from the work on automatic processes, behavior has become the theme of yet another line of research. In particular, it has been argued that goals may operate not only if they are in the focus of the actor's attention but also without the person's awareness. Under the title of "unconscious goal pursuit" (Bargh & Gollwitzer, 1994), it has been demonstrated that priming a certain goal may influence people's behavior. As a consequence, nonconsciously operating goals are seen to enable people to control behavior "without invoking conscious choice or control processes" (Fitzsimons & Bargh, 2003, p. 152). However, one might suspect that the guidance of behavior by goal pursuit differs systematically from a behavioral control that is mediated by the activation of thought contents or the exposure to specific stimuli in the situation. For example, a goal is typically understood to be a concept that is embedded in a hierarchical structure, which provides means to circumvent obstacles. This is not necessarily the case for a behavior that is elicited by perceptual or imaginal stimuli. When it comes to automatization of goal pursuit, the question arises if those characteristics are maintained. In other words, how will people react to the blocking of a goal under conditions of automatic goal pursuit? These issues are discussed in more detail at a later point.

Taken together, under the influence of the concept of attitudes, social behavior was largely understood as the result of mental operations that can be described as reflective evaluation. More recently, as the focus of research has shifted toward automatic processes, mechanisms beyond reflection were included into the theorizing. In particular, an attitude's accessibility was recognized as an important determinant. Moreover, direct links between cognitive contents and behavior were identified. In a related development, the concept of attitudes as a content of consciousness was supplemented by an "implicit" variant that operates without the person's awareness.

These developments are a move toward an understanding of social behavior that increasingly acknowledges the impact of impulsive components. At the same time, however, it is obvious that automaticity is not the only component that describes impulsive behaviors. In addition, affect seems to be an important ingredient. In fact, some theories (e.g., Frijda, Kuipers, & Ter Schure, 1989) have argued that different emotions are associated with different types of action readiness. Fear, for example, is often related to flight whereas anger is linked to aggression. Interestingly, there is evidence that the causal relationship between behavior and emotion may be bidirectional. On the one hand, negative affect has been found to trigger aggressive behaviors (Berkowitz, 1993); on the other hand, facial or postural actions have been demonstrated to influence emotional experiences (Stepper & Strack, 1993; Strack, Martin, & Stepper, 1988). Interestingly, affective components may be an important ingredient even in experimental situations that seem to be driven by purely associative mechanisms. For example, Cesario, Plaks, and Higgins (2005) have demonstrated that the behavior that is elicited by the exposure to stereotyped category members (e.g., to elderly people) is modified by the affect toward the category. These researchers found that when "elderly" was primed, participants later walked more slowly if they (implicitly) liked the elderly but walked more quickly if they disliked the elderly.

However, in the reflectively dominated spirit of social psychological theorizing, affect has lost some of its zest and was domesticated into the ruling paradigm. Emotions were understood as attributed arousal (Schachter & Singer, 1962) and affect became a "source of information" (Schwarz & Clore, 1983) that would enter into inferences like any other piece of knowledge. Aggressive behaviors were seen to be largely controlled by their anticipated consequences (Bandura, 1973) and helping behavior was construed as a series of decision at various stages of information processing (Latané & Darley, 1970).

BIOLOGICAL ACCOUNTS OF IMPULSIVE BEHAVIOR

What are the biological factors that correspond to impulsive responding? Research on this issue has focused particularly on quick behaviors that are driven by perception, as well as on shortsighted behaviors. More specifically, psychological mechanisms that have been studied extensively include decreased inhibitory control and a strong attraction of rewards (Cardinal, Robbins, & Everitt, 2003; Winstanley, Theobald, Dalley, Glennon, & Robbins, 2004).

Inhibition

One integral aspect of impulsivity is the efficiency or ability to inhibit situationally triggered responses. A labora-

tory example of inhibition is the stop task, in which participants have to respond as quickly as possible to a go signal (e.g., a red light) unless the go signal is immediately followed by a stop signal (e.g., a tone). The ability to stop the response if a go signal is present serves as an index of inhibition and correlates with general tendencies of impulsive responding (Shachar, Tannock, & Logan, 1993). There are many studies linking inhibition to the neurotransmitter *serotonin* (Cardinal, Daw, Robbins, & Everitt, 2002; Soubrie, 1986; Winstanley et al., 2004). For instance, Winstanley and colleagues (2004) destroyed the serotonergic system of rats. As a consequence, these rats showed a starkly diminished inhibitory control in various behavioral tasks (see also Harrison, Everitt, & Robbins, 1999). Comparable results were obtained with human participants. Manipulating central serotonin levels pharmacologically, Walderhaug and colleagues (2002) observed less inhibitory control in a speed-accuracy test when serotonin was reduced (cf. Craen, Richards, & de Wit, 2002). Serotonin has also been linked to disinhibited behavior in a more general sense. For instance, low levels of serotonin facilitate aggressive behavior in humans (e.g., Berman, Tracy, & Coccaro, 1997; Bjork, Dougherty, Moeller, Cherek, & Swann, 1999; Bjork, Dougherty, Moeller, & Swann, 2000; Young & Leyton, 2002), and pharmacologically augmenting serotonin may reduce aggression (Cherek & Lane, 1999). Correlational data indicate that low levels of serotonin are tied to violence and impulsiveness in suicides (e.g., Asberg, Thoren, Traskman, Bertilsson, & Ringberger, 1976) or criminal offenses.

In addition to neurotransmitters, anatomical structures of the brain can also be related to inhibition. In a broad sense, many frontal-lobe functions involve the inhibition of triggered or ongoing responses (Gazzaniga, Ivry, & Mangun, 2002; Hawkins & Bender, 2002). One extreme form of frontal disinhibition is *utilization behavior*, a syndrome in which patients compulsively imitate other persons' behavior and immediately use objects independent of intentions, such as grabbing a cup and leading it to the mouth even if it contains no liquid (Lhermitte, 1983). The anterior cingulate cortex (ACC) seems to be involved in several executive functions such as directing attention, detecting action errors, or resolving conflict between action tendencies (Banfield, Wyland, Macrae, Münte, & Heatherton, 2004; Cardinal, Parkinson, Hall, & Everitt, 2002). Particularly, it is involved in suppressing externally triggered behaviors (Paus, 2001). For instance, lesions to the ACC cause failures to suppress no longer rewarded responses (Cardinal, Winstanley, Robbins, & Everitt, 2004). Likewise, lesions to the orbitofrontal cortex (OFC) induce disinhibited behavior in several behavioral tasks (e.g., Berlin, Rolls, & Kischka, 2004). Supplementing the link between serotonin depletion and impulsive aggression, there are also indications that frontal dysfunction may enhance aggression and antisocial behavior (e.g., Hawkins & Trobst, 2000; Paschall & Fishbein, 2002; Raine, Lencz, Bihrle, LaCasse, & Colletti, 2000). There is, however, growing evidence not only that inhibition is a function of the fron-

tal lobes but also that other brain systems may contribute to this function (for a review, see Andrés, 2003).

Attraction/Repulsion

Action without foresight implies a strong attraction toward intermediate rewards and/or weak inhibition. Many studies have addressed the tolerance of delayed reward as an index of impulsive action. Waiting for a delayed reward presupposes knowledge of the reward contingencies, the ability to inhibit the impulse of immediate consumption, and a motivational preparedness to sustain the waiting. Preferences for small immediate rewards over large but later rewards has been linked to low levels of serotonin (Bizot, Bihan, Puéch, Hamon, & Thiébot, 1999; Bizot, Thiébot, le-Bihan, Soubrie, & Simon, 1988; Liu, Wilkinson, & Robbins, 2004; Wolff & Leander, 2002). For instance, Wolff and Leander (2002) found that central serotonin levels that were pharmacologically elevated in pigeons increased the animals' preference for a larger but delayed reward over a smaller but immediate one. In a study by Cherek and Lane (1999), high doses of a drug mimicking the effect of serotonin increased the readiness of human participants to wait for a large but later reward. Given its strong link to inhibition, serotonin may influence the preparedness to wait by the same route.

In addition to serotonin, the neurotransmitter *dopamine* is also related to impulses of approach and reward (Cardinal, Robbins, & Everitt, 2002; Kelley & Berridge, 2002; Parkinson et al., 2002). Much of the evidence relating dopamine and approach behavior stems from the finding that amphetamines reduce the symptoms of attention-deficit/hyperactivity disorder (ADHD), which is characterized by lack of inhibitory control and strong discounting of delayed rewards. Amphetamines, at the same time, are drugs that increase dopamine in the brain (Cardinal, Robbins, & Everitt, 2002). Interestingly, amphetamine also influences animals' preparedness to wait for larger but delayed rewards. Particularly, Cardinal (2000) had rats choose between small immediate and large delayed rewards, while in one group the delay between choice and reward was bridged by the appearance of a light cue and in the other group no cue was present when the animals had to wait. Subsequently, animals were tested under the influence of methamphetamine, which is an amphetamine derivate. Methamphetamine decreased impulsive choices if a cue was present but increased impulsive choice if no cue was present. While introducing cues during delay generally helps bridge the temporal gap, dopamine presumably increases the power of such incentive cues and thus facilitates waiting (see also Dickinson, Smith, & Mirenowicz, 2000; Richards, Mitchell, De Wit, & Seyden, 1997; Wyvell & Berridge, 2000). Under many conditions, however, there are few consistent incentive cues to larger later rewards, but instead the immediate reward is cued. In the case of drugs, for instance, frequent consumption usually establishes many drug cues, which signal the positive effect of consumption. Hence, incentive cues, and thus dopa-

mine activation, may increase impulsive responding (cf. Evenden, 1999; Robinson & Berridge, 2003). It should be noted that amphetamines also elevate serotonin immediately but may lead to serotonin depletion as a proximate effect and thereby cause impulsivity and aggression in the long run (e.g., Moeller & Dougherty, 2002; Richards, Sabol, & de Wit, 1999; Verheyden, Hadfield, Calin, & Curran, 2002).

A brain structure strongly associated with attraction and repulsion is the nucleus accumbens (Acb) (Cardinal et al., 2004). Stimulation of this structure causes pleasure and is rewarding (Berridge, 2002; Ikemoto & Panksepp, 1999). In rats, destruction of the Acb leads to a reluctance to wait for a delayed reward (Cardinal, Pennicott, Sugathapala, Robbins, & Everitt, 2001). Importantly, it has been linked to reward prediction (McClure, York, & Montague, 2004), particularly to learning the incentive value of Pavlovian conditioned stimuli (Cardinal, Parkinson, et al., 2002). For instance, if an animal has learned to press a lever for food, the presentation of a cue that has been previously paired with the food will enhance pressing. In other words, the cue has an incentive function and motivates instrumental responding. Lesions of the Acb eliminate this incentive learning (Cardinal, Parkinson, et al., 2002), and stimulating the Acb exaggerates the motivating power of incentive cues (Wyvell & Berridge, 2000). At least partially, the Acb may play an important role in addictive behavior, which is often triggered by drug cues that cause an aberrant wanting of the drug (Cardinal, Parkinson, et al., 2002; Robinson & Berridge, 2003). A second important brain structure related to attraction and repulsion is the OFC. The OFC has been associated with learning reward contingencies in a flexible manner (e.g., Cardinal et al., 2004). It may thus be a structure promoting reflected choice adapted to complex and changing contingencies (Cardinal, Parkinson, et al., 2002; Rolls, 2000). Damage to the OFC leads to perseverance of responses, which are no longer rewarded or even punished (e.g., Berlin et al., 2004). It also causes a tendency in animals to prefer small but immediate over larger and later rewards (Mobini et al., 2002). In human subjects, OFC damage has been related to a lack of anticipatory fear in risky choices (e.g., Bechara, Damasio, & Damasio, 2000).

Summary

The selective review indicates that hot impulsive responding and its inhibition can be traced to particular brain systems. Up to now, research has primarily focused on impulsive motor responding and increased temporal discounting. Also, aggressive behavior (e.g., Berman, 1997; Brennan & Raine, 1997) and addiction (e.g., Kelley & Berridge, 2002) and fear-related impulses have undergone extensive research (e.g., Öhman & Mineka, 2001). The present review of biological processes underlying impulsive responding is by no means exhaustive. There are other brain structures involved in the mechanisms discussed previously, such as the amygdala (e.g., Winstanley et al., 2004), and the functions of the de-

scribed structures are more complex (for reviews, see Cardinal, Parkinson, et al., 2002; Rolls, 2000). In addition, we have excluded the biological bases of "cool" impulses (see Bargh, 2005). The endeavor of social cognitive neuroscience has just begun to relate more complex social impulses to their biological roots.

IMPULSIVE MECHANISMS DESCRIBED IN OTHER FIELDS OF PSYCHOLOGY

While nonreflective mechanisms have only recently moved into the focus of social psychology, other areas have long acknowledged that possibility of an impulsive determination of behavior. One example is personality psychology, where the concept of "impulse" is a core component in various dispositional frameworks (see Carver, 2005).

Personality

As mentioned earlier in this chapter, Freud's (1920) dynamic model of the person included impulse as the behavioral principle of the id, which had to be controlled by the forces of the ego. Thus, the dispositional strength of the ego was negatively related to impulsive manifestations of behavior. As a consequence, assessing the functions of the ego would be an indirect way of measuring impulsivity. This approach was adopted by Block (e.g., 2002), for whom ego control was the central organizing construct of personality functioning. This dimensional construct was defined by the poles "ego undercontrol" and "ego overcontrol." In terms of behavioral manifestations, undercontrollers were seen to seek immediate gratification, readily express emotions, and be "distractable, spontaneous, energetic, uninhibited, uncompulsive, gregarious, and easily influenced by environmental contingencies" (Kremen & Block, 1998, p. 1062). In contrast, overcontrollers are expected to overly delay gratifications, to inhibit emotional responses, and to be "constrained, nondistractible, persistent, perseverative, rarely obsessive, uncomfortable with ambiguities and uncertainties" (Kremen & Block, 1998, p. 1062).

A second control system in Block's theory is called ego resilience. It has the function to select the appropriate level of ego control in a given situation. In this system, behavioral adaptiveness does not hinge on the form of ego control but on the flexibility (or rigidity) with which it is applied to a given situation. Empirical evidence suggests that ego control and ego resilience are not entirely independent. In particular, it has been demonstrated that extreme levels of ego control are tied to low ego resilience (e.g., Asendorpf & van Aken, 1999). In fact, ego control is related to various manifestations of behavioral restraint. Among them are alcohol and drug use (Jones, 1971; Shedler & Block, 1990) and the ability to delay gratification (Funder & Block, 1989).

Another dispositional characteristic that is related to impulsive behaviors is the trait of conscientiousness in the five-factor model of personality (e.g., McCrae &

Costa, 1997). Conscientiousness is partly defined by a lack of impulsiveness and spontaneity and a tendency to engage in critical thinking (Hogan & Ones, 1997). The other endpoint of the dimension is impulsivity, which is described as a tendency to act without thoroughly considering possible options and consequences (Ashton & Lee, 2001). This may be dysfunctional if the task requires planning ahead and making decisions or functional if the task calls for quick thinking and taking advantage of unexpected opportunities (Ashton & Lee, 2001). The trait of conscientiousness was found to be related to different behaviors ranging from the preference for negotiations as a means of conflict resolution (hi C) (Jensen-Campbell & Graziano, 2001) to the aggressive use of humor (lo C) (Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003), to fewer arrests among prisoners (hi C) (Clower & Bothwell, 2001). There is even evidence that highly conscientious people live longer, because they may take better care of themselves (Friedman et al., 1995).

In his review of the literature on impulse and constraint in personality psychology, Carver (2005) points out that another trait in the five-factor model that may be relevant to impulsive behavior is agreeableness in its component of inhibiting negative affect. Both conscientiousness and agreeableness contribute to the factor psychoticism in Eysenck's (1970) three-factor theory of personality, whereas in Tellegen's (1985) three-factor model, it is his name for a trait that is related to conscientiousness. Behavioral data converge with the previous findings but also have negative emotional components. Specifically, psychoticism is related to antisocial behaviors (Eysenck, 1992) and respondents with low constraint were more likely to show criminal behavior (Krueger, 2002).

As impulsive behaviors are driven by the stimuli that are present in a given situation (and not by their more remote consequences), it seems plausible to look at people's habitual tendency to seek out stimulation as another dispositional basis. Zuckerman and his colleagues have identified "impulsive sensation seeking (ImpSS)" as a trait dimension that predicts people's willingness to take risks for the sake of having novel, varied, and intense experiences (Zuckerman, 1984). This dispositional characteristic had been frequently linked to risk behaviors such as substance use and risky sex (Wagner, 2001). Horvarth and Zuckerman (1993) found strong relations to risky violations of social and interpersonal norms, particularly in the domain of criminal behavior. Among adjudicated adolescents, alcohol problems and the failure to use condoms were found to be highly correlated with IMPSS (Robbins, & Bryan, 2004).

Impulsive Behavior from a Developmental Perspective

Jerome Kagan (e.g., Kagan, 1965), was probably the first psychologist who suggested a dispositional dimension whose poles were described as *impulsive* and *reflective*. His work on cognitive styles has predominantly focused on children's solutions of intellectual tasks where reflectives tended to show long decision times and low error rates because of their examination of all alternatives in a problem situation, whereas impulsives showed fast decision times and higher error rates. Implications of this construct have also been explored for particular learning disorders, reading disability, and hyperactivity (Messer, 1976).

From a more dynamic developmental perspective, impulsive behaviors have been studied in the context of their regulation and control by the person. Whereas Kagan described reflectivity as a function of both socialization and temperament, the temperamental explanation has subsequently assumed a more dominant role in developmental models. Most prominently, Rothbart and Derryberry (1981) differentiated between two components of temperament, namely, reactivity and self-regulation. In particular, reactivity (described as the more impulsive aspects of behavior) was understood to be modulated by self-regulation (or control), which serves to promote "subdominant" responses at the expense of dominant ones (Rothbart & Bates, 1998). This may occur in either an effortful or a reactive fashion. Eisenberg and colleagues (2003) have linked the latter to Block's (Block & Block, 1980) concept of ego control and have applied this notion together with effortful control in a longitudinal study of the development of social adaptation. Although the results are complex and mediated by the intensity of negative emotions, the findings suggest that both effortful and social control have an impact on social status through their positive effect on resiliency.

Impulsive Behavior from a Self-Regulation Perspective

Most attempts dealing with dynamic aspects of impulsive behavior have focused on the antagonism between spontaneous impulses and more enduring goals of the actor. They have suggested several mechanisms how the unwanted forces can be suppressed, constrained, or at least regulated. Thus, impulsive behavior is often understood as an unwanted outcome that needs to be corrected by appropriate interventions.

Under the label "self-regulation," "efforts by the human self" are studied that attempt to "alter any of its own states or responses" (Baumeister & Vohs, 2004, p. 2). In particular, these efforts focus on people's attentional processes, thoughts, feelings, and behaviors. Typically, the self is seen as either a regulating subject that is endowed with strength and willpower (e.g., Schmeichel & Baumeister, 2004) or an object of regulatory processes that follow cybernetic principles (e.g., Carver, 2004). In the first case, the question arises if this agent acts autonomously under the principle of free will. If this is the case, searching for the determinants of regulation cannot be a scientific endeavor. In the second case, it is not clear why the "self" is included to describe the process. If cybernetic mechanisms operate in the same fashion as other principles of thought, affect, and behavior, the addition of the "self" is at least misleading as it also suggests the existence of an autonomous agent.

In both cases, however, there is the—explicit or implicit—presupposition that there are two types of psy-

chological processes: superordinate processes that regulate and subordinate processes that are regulated. Most important, the execution of the superordinate process is assumed to consume energy in order to keep the subordinate process at bay (Baumeister, 2002). Because the supply for this energy is limited, excessive use will deplete it. However, its source will be replenished as a function of time and rest.

A similar dichotomy has been described on the judgmental dimension. Specifically, uncontrolled or "intuitive" processes are often seen to require correction (Strack & Hannover, 1996). Frequently, two processes are invoked that are specialized to either generate a response or to adjust or recompute it according to some superordinate standards of correctness. For example, the so-called anchoring heuristic (Tversky & Kahneman, 1974) has been explained originally by a more automatic response that is elicited by the anchor value and a more deliberative adjustment process that is often insufficient. In the domain of person perception, the characteristic of an observed act has been assumed to automatically activate a category to form an impression of the actor. Given the necessary capacity, the initial impression may be modified or corrected by taking the situational circumstances of the behavior into account (Gilbert & Malone, 1995; Trope, 1986). Beyond improving the accuracy of factual judgments, norms of political correctness were found to modify the judgments of other people. In particular, it has been demonstrated that persons' need to be unprejudiced reduced the influence of stereotypes on explicit judgments (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Dunton & Fazio, 1997; Plant & Devine, 1998).

It is important to note that these models are based on the assumption that only one of the two processes is operating at the same time. That is, depending on the requirements of the task, a more automatic and intuitive type of processing or a more controlled and systematic type of processing will be invoked (e.g., Chaiken, 1987; for a collection of various dual-process models, see Chaiken & Trope, 1999).

Applied to the general theme of this chapter, it can be argued that dual-process models may contribute to a better understanding of the dynamics of impulse by focusing on the uncontrolled component of information processing. In fact, if impulsivity is defined as a characteristic that includes a loss of control, mechanisms that operate in an uncontrolled fashion may be central components/precursors of impulsive behavior. However, most dual-process models (for a recent summary, see Smith & DeCoster, 2000) have no explicit link to behavior. Also, they do not include motivational mechanisms that may play a role in impulsive behavior. Therefore, we focus the remainder of this chapter on a recent dual-systems model that was particularly designed by the authors of this chapter (Strack & Deutsch, 2004) to explain social behavior and to identify its underlying mechanisms. In particular, this model is meant to explain the behavioral dynamics of impulsivity by focusing on the interaction of a reflective and an impulsive mental system. However, before we start discussing the reflective–impulsive model (RIM) in more detail, we would like to briefly discuss two models that are conceptually related to the RIM. However, despite their similarities, the current conceptualization either differs in important aspects or goes beyond the previous models.

Related Frameworks

Mainly based on research on delay of gratification (e.g., Mischel & Ebbesen, 1970), Metcalfe and Mischel (1999) have offered a "hot/cool-system analysis" of human behavior. As in the RIM, the hot and the cool systems operate in parallel and determine behavior in an interactive fashion. Unlike the RIM, however, the two systems are primarily defined by the presence or absence of affect. Specifically, the hot system is described as emotional in nature while the cool system's major characteristic is "cognitive." Although there are other distinctive features of the two systems, the affective quality of the hot system has a dominant role. As a consequence, impulsive behaviors are understood to be affectively driven and to be regulated by the force of willpower, which resides in the cool system.

Although affect is a possible ingredient of the impulsive system in the Strack/Deutsch model, it is not a necessary component. Because we define impulsive processing by the underlying operating characteristics, it does not hinge on the presence of affect. For example, the nonaffective operation of habits is understood to be a possible manifestation of impulsive behavior.

Despite their similarities, the RIM seems to be more general than the hot–cool model. Beyond phenomena of temptation and self-regulation, it is capable of accounting for habitual behaviors and of cognitive phenomena such as selective accessibility and judgmental heuristics.

Although described by its author as a global theory of personality, Epstein (e.g., 1994, 2003) has formulated a cognitive–experiential self-theory (CEST) that shares many basic characteristics of the RIM. Specifically, it proposes two interacting modes of information processing: a rule-based rational system and an emotionally driven experiential system. Moreover, the two models share the same operating principles for the two systems. That is, associative mechanisms for the "experiential" and the "impulsive" system and inferential mechanisms for the "rational" and the "reflective" system (see also Sloman, 1996).

At the same time, there are some important differences that need to be noted. While CEST assumes that the experiential system represents reality in the form of images, metaphors, or narratives and processes them in a "holistic" fashion, the impulsive system of the RIM adopts the structure of an associative network as it is assumed in many cognitive theories of human memory (e.g., Baddeley & Baddeley, 1990), which includes symbolic and abstract representations. Also, the RIM only invokes the two basic determinants of associative processing, namely, the frequency and recency of prior activation. As a consequence, the interactions between the reflective and the impulsive system are based on narrowly circumscribed mechanisms of cognitive accessibil-

ity that allow a precise identification of the cognitive operations that underlie judgmental phenomena (e.g., Mussweiler & Strack, 1999a; Strack & Mussweiler, 1997). Most important, perhaps, the RIM attempts to specify the mental operations in a sequential order to exactly predict their cognitive and behavioral consequences on a microlevel. For example, the RIM was able to generate predictions about different ways of processing negations in the reflective and the impulsive system (Deutsch, Gawronski, & Strack, 2005). In the domain of behavior, the RIM is endowed with mechanisms that describe specific links between the evaluative quality of information and the locomotive quality of the behavior. On this basis, it is possible to integrate findings about the influence of behavior on the selective processing of information (e.g., Förster & Strack, 1996; Strack et al., 1988) and about the facilitation of approach versus avoidance as a function of the valence of the processed information (see Neumann, Förster, & Strack, 2003).

In summary, there is no doubt that the two models are compatible in many of their components. At the same time, the CEST and its characteristics seem to be linked more to the global aspects of psychodynamics and personality theory while the RIM attempts to harness the findings from cognitive and social psychology to identify the precise mechanisms that are constituent elements of the general model.

THE DYNAMICS OF IMPULSIVE BEHAVIOR: THE REFLECTIVE–IMPULSIVE MODEL

To account for impulsive behavior, the RIM (see Strack & Deutsch, 2004, 2005) deviates from the widespread notion that behaviors may be classified into different types on the basis of the underlying process (see also Kruglanski & Thompson, 1999). Instead, it rests on the assumption that each human behavior is the joint result of two independent systems that have different operational principles. One system is called impulsive, the other reflective. Although their operations are described on a mental level of analysis, it is worth noting that neuroscientists have proposed a similar distinction. For example, Bechara (2005) has identified two neural systems that he describes as "impulsive, amygdala dependent" and "reflective, orbifrontal dependent."

The second assumption that is inherent in this model is that the two systems interact at different stages of the processing. In other words, impulsive and reflective processing occurs in parallel and not in a mutually exclusive fashion. More precisely, the impulsive system is assumed to be permanently active while the parallel operation of the reflective system depends on the cognitive capacity that is available. The following sections describe the operating characteristics of the two systems.

Elements of the Impulsive System

The most basic version of the impulsive system is identical to James's (1890) ideomotor principle (see also Lotze, 1852). That is, a perceived or imagined content may elicit

a behavior without a person's intention or goal. In line with recent neurological evidence on mirror neurons (Gallese, Fadiga, Fogassi, & Rizzolattti, 1996), the ideomotor principle can be extended to concepts that include a motor component. Specifically, it has been found in monkeys that the perception of movements activates neurons in the premotor cortex. From the perspective of the present model, the input goes to a behavioral schema that generates the appropriate motor execution on a muscular level. The activity is not executed if there exist competing influences in the impulsive system or a superordinate regulation.

However, the link between perceptual or conceptual input and the behavior may also operate in a less immediate fashion. Specifically, the behavioral implications of a concept may be activated by a concept (or an image) that is only indirectly connected to a behavioral schema. For example, Bargh and colleagues (1996) found that having been exposed to a stereotype about elderly persons caused people to walk more slowly (see also Cesario et al., 2005). If such behaviors are not assumed to be elicited by a decision or a goal, this requires a more complex structure of the impulsive system. In fact, the RIM assumes that in the impulsive system, the perceptual input is linked to the behavioral output by an associative network that is the result of past operations. Specifically, it represents previous inputs and provides a structure by linking it to other contents. Thereby, this network serves as a memory that connects the past to the present. Its most basic principles are associative in nature. It is assumed that like other associative-network models (see Smith, 1998), a link between elements is created by their joint activation, and its strength is a function of the frequency and the recency with which the activation has occurred. Thus, each element has acquired a specific activation potential (Higgins, 1996). The likelihood that the activation of one element will spread to another therefore depends on the existence of a link between the two and on its strength. The activation of a particular element may originate from one element with which it is strongly associated (like "black" and "white") or from several other elements that are activated at the same time. Thus, in the Bargh and colleagues study, it was not sufficient to activate the elderly stereotype to induce the slow walking; in addition, participants had to be induced by other means to engage in walking. Therefore, it seems likely that it was this joint activation that resulted in the behavioral effect.

Advantages and Disadvantages of Creating Habits

As may be apparent already from the present description, operations of the impulsive system require little effort. As a function of frequency and recency, the mere exposure to the appropriate stimulus seems to suffice to elicit a behavior. And if these principles are employed by practicing more complex sequences, the execution of behavior is greatly facilitated. In other words, a behavior has become habitualized and can be performed without much attention. However, the advantage of effortless execution is accompanied by a serious disadvantage,

namely, the slowness with which stable links are created (Devine, Plant, & Buswell, 2000) and particularly their rigidity and perseverance against change. Thus, the benefits a formerly effortful task enjoys if it is "made a habit" are matched by the effort that is needed to "break it" if it turns out to be undesirable.

The Role of Affect and Valence

While behavior can be effectively generated through a strong link to a behavioral schema, another mechanism is equally powerful. It is based on the valence that is experienced in the course of its execution. Specifically, a behavior that is associated with feeling better is more likely to be elicited than a behavior linked to feeling worse. This hedonic experience obeys the same associative principles. As a result, a stimulus associated with the behavior may be sufficient to elicit the feeling that will have a facilitating or an inhibitory function.

In general, the impulsive system is assumed to be oriented toward approach or avoidance, and this "motivational orientation" (Cacioppo, Priester, & Berntson, 1993) may be triggered by the experience of positive or negative affect, the perception of approach or avoidance, or the execution of such behaviors, as well as by the processing of positive stimuli. In the RIM, approach orientation is defined as a preparedness to decrease the distance between the person and an aspect of the environment. This may occur in an actual locomotion or as a symbolic operation (see Markman & Brendl, 2005) or an imagination. Similarly, avoidance orientation is defined as a preparedness to increase the distance between the person and an aspect of the environment. It can be achieved by either moving away from a target (flight) or by causing the target to be removed (fight). Within both motivational orientations, the specific type of response is determined by other influences.

The mechanism of motivational orientation is further described by two operating principles, namely, compatibility and bidirectionality. The compatibility principle states that the processing of information, the experience of affect, and the execution of behavior are facilitated if they are compatible with the prevailing motivational orientation. The bidirectionality principle states that this influence operates in both directions. Thus, behavior may influence affect and evaluation, and vice versa.

Research on regulatory fit (Higgins, 2000) involving promotion and prevention focus (Higgins, 1997) provides a good example of the importance of both motivational orientation and one kind of compatibility. When people have promotion focus concerns with accomplishment and advancement they prefer eager approach strategies to pursue their goals. In contrast, when people have prevention focus concerns with security and responsibilities they prefer vigilant avoidant strategies to pursue their goals. Regulatory fit occurs when individuals pursue goals in a manner that sustains (vs. disrupts) their current motivation orientation. Thus, goal pursuit with eager approach strategies fits a promotion focus but disrupts a prevention focus, and the opposite is true for vigilant avoidant strategies. The value of activities and outcomes can be intensified when people pursue goals under conditions of regulatory fit (e.g., Higgins, Idson, Freitas, Spiegel, & Molden, 2003).

In a general manner, these principles would probably not be too surprising. For example, self-perception theory (Bem, 1967) has long posited that behavior may influence attitudes. However, the underlying mechanisms are quite different. Specifically, this theory assumes that people use their behavior and the situational context to draw inferences about the underlying attitudes. More colloquially, they ask themselves what their attitudes must be if they behave in a particular manner. In contrast, the impulsive mechanisms of the RIM operate without drawing such inferences. The described facilitative effects are predicted to occur even if the meaning of a behavior is not apparent (e.g., Strack et al., 1988). Thus, the motivational orientation can be understood as a global predisposition of the impulsive system that facilitates the processing of information and the execution of behavior in a specific manner. There exists ample evidence for the postulated influences in both causal directions (Strack & Deutsch, 2004).

Although this general principle is a central element in the architecture of the impulsive system, it is important to acknowledge individual differences. One reason may be found in the divergent experiences to which people are exposed. Another reason, however, may lie in inborn temperamental factors. As Kagan and Snidman (1991) have demonstrated, infants differ in their tendencies to approach or avoid unfamiliar situations. And these different "temperaments" were found to influence people's subsequent development. Although their specific structure may be more complex than initially assumed (see Putnam & Stifter, 2005), the finding that innate tendencies toward approach or avoidance play an important role and shape personality characteristics is undisputed. It should also be noted that individual differences in motivational orientation not only concern the tendency to approach desired end states or avoid undesired end states but also occur at the strategic level, such as the tendency mentioned earlier for promotion-focused individuals to use eager strategic means of goal pursuit and prevention-focused individuals to use vigilant strategic means (Higgins, 1997).

At the same time, it is obvious that an adaptive processing system also needs to be oriented toward the specific requirements for the survival of the organism. This should be true for the survival of individuals and their genetic endowment (i.e., for nutrition and reproduction). For example, the impulsive system of people who are deprived of food should facilitate approaching responses toward stimuli that are food related. Similarly, sexual stimuli should elicit attraction particularly under circumstances that lead to reproduction. In fact, evidence (Seibt, Häfner, & Deutsch, 2005) suggests that people who are hungry react faster to food-related words if the required response consists of moving a lever toward the stimulus and react slower if the response is a movement in the opposite direction.

In sum, the impulsive system is specialized to adjust to the organism to the environment in a fast and automatic

manner. Its architecture is determined by a shared genetic endowment that is shaped by adaptive mechanisms, by an inborn disposition toward approach versus avoidance, and by the idiosyncratic history of impulsive operations.

The Reflective System

The advantages of the impulsive system are also accompanied by distinctive shortcomings. That is, its fast and effortless processing goes hand in hand with a considerable rigidity and a failure to perform certain tasks. Specifically, the principle of frequency impairs influences that come from a single piece of information. Similarly, the recency principle causes a distinct disadvantage for information whose impact lies in the past. Moreover, the process of adaptation would be greatly accelerated if changes did not require being personally exposed to the relevant stimuli but allowed learning from others' experiences. Particularly, learning from others' mistakes without suffering the negative consequences of their errors is a great advantage over having to have the negative experience oneself.

To achieve these benefits, there exists a second, reflective system whose operating principles differ from those of the impulsive system. Specifically, the reflective system is based on the generation and transformation of knowledge. That is, the reflective system assigns truth values to information that is contained in associative links and transforms these truth values through syllogistic operations. For example, the perception of a particular person may elicit the characteristic "old" in the impulsive systems. In addition, the reflective system would create a relationship between the perceptual input and the characteristic and assign it the value "true." On the basis of such a propositional categorization, further transformations can be performed. For example, the truth value can be reversed by applying the operation of a negation. Or inferences about other characteristics of a target may be inferred from information that is contained in the category to which it has been assigned.

The reflective system consists of a series of operations that are involved in this epistemic process. At the outset, a deictic procedure ("pointing and referring") assigns an input that may result from either perception or imagery to a category. Through syllogistic reasoning, the "propositional categorization" may become the basis of a "noetic decision" that has a factual and an evaluative component. That is, the reflective system creates knowledge about what is the case or what is good or bad. Structurally, the two variants do not differ and may often be inextricably intertwined. However, the evaluative component of the noetic decision may become the basis of a "behavioral decision" or a behavioral goal. Unlike decisions about facts and values, a behavioral decision focuses on the reduction of a discrepancy between a current state of the self and a positively evaluated possibility. The link between a behavioral decision to actual execution of the behavior is a process named "intending" (e.g., Gollwitzer, 1999), which is terminated if the behavior is executed or if the receding behavioral decision is already met.

Reflective–Impulsive Interactions

Up to this point, the two systems have been described in the major elements and in operational principles. Under the assumption that they function in an either–or fashion, we could continue describing their selective application to various tasks (e.g., Petty & Cacioppo, 1986). However, the RIM posits that the two do not operate by themselves but in close interaction. That is, at the various stages of processing, both systems influence one another. The degree of this influence, however, depends on the conditions that allow the reflective system to contribute. Specifically, it is assumed that the impulsive system is always operating while the parallel reflective operations depend on sufficient capacity and motivation. If they are not available, the impulsive system will have the upper hand.

In the following paragraphs, the interaction of the two systems is described in its general characteristics and in its manifestation at various stages of processing. First and foremost, the reflective operations that lead to the various decisions need an informational basis. Because this information is typically not part of the perceptual input, it must come from a preexistent store. The RIM assumes that the two systems provide different types of memories. The reflective system has a working memory with a very limited capacity and the possibility to directly address its contents (Baddeley, 1986). In contrast, the associative structures that are slowly formed by the impulsive system form a simple long-term store with an unlimited capacity (e.g., Johnson & Hirst, 1991).

This associative store plays an important role for the functioning of the reflective system. On the one hand, it provides the contents that are used for its syllogistic operations. On the simplest level, a propositional categorization can be performed only if the category is already available. That is, the label "elderly" can only be assigned if it is retrieved from storage. Similarly, inferences from a general knowledge about the elderly can only be drawn if an appropriate schema can be found in memory.

This retrieval may be triggered by the input. That is, if the input is strongly associated with a given category, it will most likely be activated and used for further processing in both the impulsive and the reflective system. As a second possibility, there may be no strong link between the input and a category. Then, the activation that is spreading along the associative links may trickle away. However, operations in the reflective system may require the input to be categorized. For this purpose, categorical information from the associative store must be retrieved. Then, the outcome of the retrieval depends not only on the search cue but also on the accessibility of the information. Specifically, we assume that each piece of information has a specific activation potential (Higgins, 1996) that describes the probability with which the information enters into reflective operations. The activation potential depends on the frequency and recency of prior activa-

tion. Thus, processes in the impulsive system may exert a strong influence on reflective operations. For example, people may categorize a target person as "reckless" simply because they had previously been exposed to this category in a different context (Higgins et al., 1977). However, the use of the activated category for a given target also depends on the category's "judged usability" (Higgins, 1996; see also Strack, 1992), which requires an operation that is part of the reflective system.

However, the activation potential of a piece of information is not only determined by the frequency and recency with which the impulsive system is exposed to a stimulus but also by its prior use in the reflective system. In other words, thinking about content will increase the probability that the same (or related) information will be retrieved at a later time. This may have severe consequences on judgment and decision making. For example, if a preceding judgment leads to a selective search of a particular type of information, a subsequent judgment may become biased by the selectively increased accessibility.

This mechanism of selective accessibility was identified to be a cause of judgmental influences under the name of "anchoring heuristic" (Tversky & Kahneman, 1974). That is, the observation that an absolute judgment will be assimilated toward the standard of a preceding relative judgment has been demonstrated to be a function of an increased semantic accessibility (e.g., Mussweiler & Strack, 1999a; Strack & Mussweiler, 1997). Specifically, it was shown (for reviews, see Mussweiler & Strack, 1999b; Strack & Mussweiler, 2003) that forming a comparative judgment had the characteristics of testing an hypothesis in that the standard is entertained as a possibility and in that information is being sought that is consistent with the standard. Even if the information that is found is not sufficient to warrant the acceptance of the hypothesis, the probability that this information will be activated at a later time has been greatly increased. If an absolute judgment about the target has to be formed after the comparative judgment, the previously used information will be more likely to enter into the judgment. As a consequence, the resulting judgment will be assimilated toward the standard of the comparative task. Thus, this anchoring distortion is not due to a direct influence of the anchor value but to the consequences of a specific cognitive operation that is used to generate a comparative judgment. Most important, this influence is introspectively inaccessible and cannot be prevented (Wilson, Houston, Etling, & Brekke, 1996).

These mechanisms exemplify a second type of interaction between the two systems. While a given activation potential in the associative store determines the likelihood that the information is used for reflective purposes, reflective processes modify the activation potential in the associative store, which may become manifest in subsequent operations. Moreover, reflective operations may change associative structures based on the same principles as direct input to the impulsive system. That is, the frequency and recency of co-occurrence in reflective operations determines the likelihood that activation spreads from one element to another in the same way as the frequency and recency of co-occurrence in perceptual exposure.

Synergisms and Antagonisms in the Determination of Behavior

In the RIM, the two systems converge in the final pathway to behavior. That is, the different operational mechanisms contribute to behavioral outcome. To understand the synergistic and antagonistic interplay between the two systems, a closer look at the immediate precursors of the behavior seems to be appropriate.

In the reflective system, behavior is caused by reasoning that generates a noetic decision about the desirability and feasibility of a particular action (cf. Ajzen, 1991; Bandura, 1977). As discussed before, the impulsive system may have already had its effect on those decisions through the accessibility of the activated information. However, a behavioral decision may not immediately result in a goal-directed behavior. Instead, the impulsive system may activate other behavioral schemas that are incompatible with the behavioral decision. Alternatively, the behavioral decision may refer to a later point in time. These issues are examined in turn.

From the previous discussion, it follows that the effect on the behavior depends on the compatibility of the input from both systems. That is, if both systems contribute to an activation of the same schema, the behavior will be facilitated. The execution may be smoothed even further because the contribution of the impulsive system will ease the execution and create a feeling of fluency and positive affect (Csikszentmihalyi, 1988; Winkielman & Cacioppo, 2001). The motivational implications of the cooperation between the two systems are obvious.

However, the two systems may also stand in competition if they activate incompatible schemas or if the execution of impulsive behaviors is inhibited by the reflective system. To take an example from regulatory focus theory (Higgins, 1997), one could imagine a situation in which the impulsive system experiences threat that instigates a prevention-focus preference for vigilant avoidant means but a strategic analysis in the reflection system concludes that eager approach means are preferable. Such antagonistic activation may elicit feelings of temptation and conflict. To gain the upper hand, the reflective system may apply knowledge about the impulsive mechanisms. In particular, it may divert attention from the tempting stimulus. Finally, it should be noted that although both systems may contribute to the execution of a behavior, the impulsive system can take on primary control if the operating conditions for the reflective system are fulfilled. As a consequence, it will be less likely that the behavior will be determined by assessments of feasibility and future consequences than by its immediate associations and the hedonic quality. Under certain circumstances, this may be of great adaptive value. In other situations, however, such impulsive determination may be disruptive and even damaging.

The second issue is about the delay of the execution of a behavioral decision, which is necessary if certain preconditions need to be fulfilled before the behavior can be performed. As a consequence, the temporal gap between the behavioral decision and the execution must be bridged. One possibility would be the permanent activation of the behavioral schemas. However, this would absorb a large amount of cognitive capacity. In addition, this would increase the risk that the behavior would be prematurely executed as a function of ideomotor processes. Because such a mechanism of rehearsal would be maladaptive, we suggest that the gap is being bridged by a process of intending (e.g., Gollwitzer, 1999) that automatically reactivates the behavioral decision, which in turn activates the behavioral schemas that are conducive in the situation. Following its original conceptualization, we also assume the intending process to be self-terminating in that it will be turned off if the goal of the decision has been fulfilled.

Intending also plays an important role if a current behavioral access to a goal is blocked. While the impulsive system will cause emotions like anger, which may result in destructive behaviors, the mechanism of intending may check the instrumentality of the obstructed goal in relation to a superordinate end and choose a different way to achieve the same objective. As with negation, we assume that the impulsive system may circumvent obstacles only through trial and error but not through circumventing obstacles by initiating a new operation of intending.

The Downside of Automaticity

The described transition from effortful reflective operations to effortless associative processing seems to imply a great potential for saving cognitive capacity. In fact, this possibility greatly facilitates the execution of myriad tasks (e.g., Salvucci, 2005) and allows the allocation of the scarce resource to novel tasks. The adaptive value of practice can therefore not be overestimated.

Recently, some researchers (e.g., Bargh, 1997) have even proposed that *all* "controlled" cognitive activities can be automatized. From the present perspective, this optimism is not shared. Rather, the different representational principles of the two systems suggest that some cannot be delegated to the impulsive system. In particular, this applies to representations of what is not the case.

As explained in the previous sections, the memory store of the impulsive system consists of associative structures in which different contents are linked. The only operation is the spread of activation along those links as a function of their strength. Although syllogistic operations occur in the reflective system, their outcomes may establish new links in the associative store. For example, the simplest operation of negating content may create a new characteristic that will then be linked to the original content. Thus, the information that a person is not guilty may lead to the conclusion that he is innocent. Thus, as a function of the reversal of the truth value of a simple propositional representation, a new associative represen-

tation will be formed. Through frequent association of the term with affirmative concepts, a negated term may acquire its own meaning (e.g., "not guilty" in a court of law). But if this link has not been established yet, the impulsive system cannot generate the implications of the negation. Instead, it may process the affirmative content.

Evidence for this possibility comes from various sources. Under a different theoretical framework, Wegner, Ansfield, and Pilloff (1998) had instructed participants to hold a pendulum and prevent it from moving. In another condition, they were specifically told not to let the pendulum swing along a line that was drawn on a piece of paper. Orthogonally to these conditions, participants had to perform a secondary task that required mental effort, a secondary task that required physical effort, or no secondary task as a control condition. The results of this study suggest that the negation that was implied in the experimenter's request was less likely to be processed if participants' attention was directed to another task. However, the affirmative base of the request remained effective. As a consequence, participants in the distraction conditions were more likely to do exactly what they were instructed to prevent. From the present theoretical perspective, this finding supports the notion that negations are generated in the reflective system while the impulsive system processes their affirmative bases. As a result, conditions that impair the reflective system will prevent negations from being formed and facilitate those behaviors that are implied by their affirmative bases. It is important to note that these affirmative influences are generated by the impulsive system, which is not affected by attentional load.

These mechanisms have important implications for any attempts at influencing people's behavior by using negations. For example, in the domain of advertising (Grant, Malaviya, & Sternthal, 2004) it was found that under limited cognitive resources, a brand was evaluated more positively if it was described by a negation that was negatively valenced ("not easy to use") than by a negation that was positively valenced ("not difficult to use"). In a related vein, written advice for elderly users of prescription drugs was not only less effective if it attempted to communicate what users were *not* supposed to do, it even caused users to remember the opposite implications (Skurnik, Yoon, Park, & Schwarz, 2005).

As a consequence, getting people not only to say "no" (or, "not") but also to think and act accordingly seems to work only under conditions that allow the reflective system to be operating. If these conditions are not met, the impulsive system will take over and the influence attempt will backfire. This should particularly be the case if negations are directed toward behaviors that are driven by needs and desires. As the described mechanisms of approach reside in the impulsive system, the reflective system may be at a disadvantage to begin with. As a consequence, negations should be particularly ineffective and even counterproductive under circumstances in which strong needs and desires are activated.

In fact, evidence suggests that this is the case. Generally, research has established that negated statements

are often (falsely) remembered as being true (e.g., Gilbert, 1991), that understanding negated statements consumes more cognitive resources (e.g., Mayo, Schul, & Burnstein 2004), and that automatic responses to negated information direct cognition and behavior in the opposite direction of what was implied logically (e.g., Deutsch et al., 2005; Wegner et al., 1998). More specifically, there are some studies demonstrating such paradoxical effects of negations in situations in which self-regulation is intended. Farrelly and colleagues (2002) compared the effectiveness of antismoking campaigns that used negations (e.g., *Don't smoke*) versus affirmative expressions (e.g., *Tobacco kills*). Exposed to negations, young adults were even more open to the idea of smoking, whereas the affirmative campaign positively changed attitudes toward tobacco. Similarly, drug-related intentions containing negations (e.g., *I will not touch these cigarettes*) will enhance the activation of the negated concept in memory (Palfai, Monti, Colby, & Rohsenow, 1997), thereby pushing behavior toward consumption.

Finally, evidence exists that governmental programs to promote sexual abstinence among teenagers may backfire. In a study that was reported in the news (e.g., newsobserver.com, January 31, 2005), researchers from Texas A&M University collected data suggesting that juveniles who had participated in abstinence-only programs were subsequently more likely to engage in sexual intercourse than those who had not. Although the original data are not published yet, these are consistent with predictions from the RIM. Specifically, it follows from the model that negating the possibility of sexual behavior increases the activation potential of sexual concepts. In a situation of temptation, those concepts will be more likely to be activated and the their impulsive consequence will determine the behavior.

Another characteristic of impulsive responding (i.e., shortsightedness) may be deeply rooted in the characteristics of associative processing. Particularly, representations of the past and the future require propositions (Roberts, 2002). For instance, thinking about a party that happened last month requires ascribing the temporal qualifier *last month* to the representation of the party. Hence, without reflection, processing stops at activating concepts. For instance, thinking of either a past or a future party is predicted to elicit similar impulses. At the same time, research on associative learning indicates that associations that are acquired without insight have a very short time horizon. For instance, Clark and Squire (1998) demonstrated the automatic acquisition of classically conditioned eye-blink responses only when the conditioned and unconditioned stimuli overlapped but not if there was a gap between the two stimuli. In the latter case, conditioned reflexes were only observable for participants who had gained insight into the contingency between conditioned and unconditioned stimuli. There are, however, some nonreflective mechanisms that support future-oriented behavior (Roberts, 2002). However, they appear to be limited to specific and inherited behaviors (such as hoarding in animals) or to depend on extended learning.

BEYOND "IMPLICIT ATTITUDES": IMPULSIVE INFLUENCES ON PSYCHOLOGICAL PHENOMENA

It is obvious that in many aspects, the dynamics of implicit phenomena resemble the processes of the impulsive system. However, the RIM has some features that are not shared by most accounts of implicit/explicit phenomena. First, reflective and impulsive processes in the RIM are assumed to run in parallel, and not in an either–or fashion. As a consequence, it is possible to identify interactions between the two systems at different stages of the processing. Second, reflective and impulsive processes are oriented toward behavior. Although it has been proposed that implicit and explicit phenomena may use different routes to behavior (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Wilson et al., 2000), few models specify the cognitive and motivational structures of such influences.

Different Types of Processing

Through their direct link to behavior, the impulsive mechanisms that are specified in the RIM are distinct from other dual-process models contrasting two systems that are specialized in either systematic or intuitive/heuristic processing (e.g., Chaiken, 1987; Kahneman & Frederick, 2002). While "systematic" processing is largely compatible with reflective operations, "intuitive/heuristic" processing is not necessarily "impulsive." The influence of affect on behavior exemplifies this point. In the RIM, the relation between affect and behavior is at least twofold. On the one hand, the experienced affect may enter into reflective operations "as information" (Schwarz & Clore, 1983) by being propositionally categorized. Then, its impact may depend on its representativeness (Strack, 1992). At the same time, an affective state may exert a much more direct impact on behavior through the impulsive system by influencing the motivational orientation and facilitating approach or avoidance.

From a RIM perspective, even intuitive or heuristic processing occurs in the reflective system. In other words, although a judgment may be simplified (e.g., by using an affective experience), it is still based on knowledge that obeys syllogistic mechanisms. Therefore, impulsive behaviors and intuitive judgments are conceptually distinct and should not be confused in dual-process models. As a consequence, the contribution of the RIM to a better understanding of heuristics is not that it provides a separate processing system for heuristic judgments but that it describes how input from the impulsive system enters into reflective operations and thereby simplifies judgment and decision making. Thus, judgments that are categorized as "systematic" have impulsive components. This has been demonstrated in the context of the anchoring effect, where the "systematic" generation of a comparative judgment increases the activation potential of standard-consistent information. As a consequence, the equally "systematic" generation of a subse-

quent absolute judgment is influenced in an inadvertent and uncontrollable fashion because the recall of the preactivated information will be "impulsively" facilitated (Strack & Mussweiler, 1997). Similarly, a motivational orientation will exert a facilitating effect on the recall of information that is compatible with a given behavior. The same is expected to be true for the influence of deprivational states (see Seibt et al., 2005). Therefore, the absence of system purity precludes a dichotomy of processes that is tied to characteristics of the judgment task. Rather, the RIM implies that depending on the specific circumstances, both the impulsive and the reflective system contribute to the generation of a judgment. Deeper understanding is reached not by classifying judgments as either "impulsive" or "reflective" but by identifying the relative contributions of both systems and by explaining their specific interactions.

Determinants of Counterattitudinal Behaviors

Although the suggested conceptual integration of impulsive mechanisms adds to a deeper understanding of judgmental processes, its main benefit lies in the explanation of behaviors that are inconsistent with people's attitudinal evaluations. We have mentioned before that such behaviors have primarily served the function of an independent variable in theories of cognitive consistency while its determinants remained largely unexplored. As a consequence, certain types of behavior have been given a marginal status or even been neglected in theorizing of social psychological theorizing. We discuss two examples in the concluding paragraphs.

Aggression

No doubt, aggressive behaviors have an enormous impact on social interaction. Moreover, they do so not only in massive outbreaks of violence but also in more subtle manifestations in everyday encounters. Given the pervasiveness of aggression, the relative scarcity of research may be due to the divergent conceptualizations of the phenomenon (see Anderson & Bushman, 2002). In fact, most accounts of aggressive behavior have been influenced by two theoretical orientations that invoke entirely different psychological mechanisms.

From the perspective of Bandura's (1977) social learning theory, aggression was understood as a behavior that was not counterattitudinal but guided by its anticipated and evaluated consequences. For example, it has been shown that the imitation of aggressive behaviors was reduced if the witnessed aggression was punished (Bandura, 1973), while children who observed assaulting conduct consistently rewarded were most aggressive (Rosekrans & Hartup, 1967). In this perspective, aggressive behaviors are a rational means of gaining a desired result, which do not differ from other instrumental behaviors. This view has specific implications for intervention, namely, to teach people that aggression "does not pay" and is a miscalculation of the expected outcome.

This account stands in stark contrast to an alternative explanation that has its roots in psychoanalytic theory (Freud, 1920). Specifically, aggression is understood as an emotional reaction that is not subject to computations of utility. In experimental psychology, this view has first become known as the frustration–aggression hypothesis (Dollard, Miller, Doob, Mowrer, & Sears, 1939), which has undergone various modifications. In the spirit of this perspective, Berkowitz developed a theory that combines affective antecedents with the accessibility of contents to predict aggressive behaviors. In his "neo-associationist" (Berkowitz, 1990) account, the encounter of negative effect and appropriate behavioral cognitions will automatically elicit aggressive behavior. To reduce aggression, it is not sufficient to influence people's beliefs about behavioral contingencies. Rather, an effective intervention must focus on either the emotional antecedents or the activation potential of aggressive thoughts.

Obviously, the mechanisms described in Berkowitz's theory are akin to impulsive processes whereas Bandura's social learning account is based on reflective principles. There were several attempts at combining these divergent perspectives. For example, a widely accepted distinction exists between "hostile" and "instrumental" aggression (Berkowitz, 1993). Whereas "hostile aggression" was described as emotionally driven, the "instrumental" variant was meant to be a thoughtfully chosen action to reach a goal beyond afflicting harm to the victim. Although this distinction captures the different mechanisms to some extent, it seems possible that hostility may delivered in a deliberate fashion and emotional aggression may not necessarily be driven by an intention to afflict harm. More recently, Anderson and Bushman (2002) have proposed to distinguish between two types of goals that are pursued through aggressive action. Specifically, they suggest that any aggression is based on an intention to harm a victim while some types also have ulterior goals, like the acquisition of some valuables in a robbery.

From the perspective of a two-systems model, it seems more meaningful to use the different qualities of the underlying psychological processes as the major criterion. That is, "impulsive aggression" can be defined as a harmful behavior that is not guided by a particular goal but is driven by a negative emotion, predominantly by anger. In contrast, "reflective aggression" is a goal-directed behavior to either afflict harm or reach an ultimate objective while accepting that harm is being afflicted. Of course, in the execution of the behavior, both types of aggression may blend. For example, anger and impulsive aggression may arise if obstacles prevent the actor from attaining the ultimate goal.

Psychologically, it might be questioned if it is useful to categorize a behavior by its consequences. In this case, if afflicting harm is a goal that is pursued by a person choosing the most appropriate means, why should the operational principles differ from other behaviors that are based on rational choice? Thus, if aggression is meant to express a unique psychological quality, it may make sense to reserve this term for a behavior that is not only harmful but also driven by negative emotion (i.e., driven by the impulsive system).

Temptation

An even clearer type of behavior that does not fit into the general perspective of attitudinal determination is that of temptation. Temptation is often defined as a power that directs behavior into an unwanted direction. In other words, it causes a behavior that is negatively evaluated. Typically, temptation is accompanied by a subjective experience that has been described as "craving" (e.g., Loewenstein, 2001). After having "fallen" into temptation, resuming the attitudinal direction of the behavior is often preceded by emotions of regret and even contrition. Temptation can be counteracted not by focusing on people's evaluative beliefs but by interfering with the impulsive mechanisms. For example, this may be achieved by blocking the perceptual input.

If temptation prevents people from doing what is good for them, it is not surprising that it promotes unhealthy behaviors. This is perhaps most obvious in the domain of drug consumption where people engage in a behavior they know to have severe negative consequences. Other examples are unprotected sex and overeating. Some temptations may eventually turn into addictions that are difficult to reverse. Still, research on these important phenomena is rare in social psychology or psychology in general. However, there are some exceptions.

One line of research that has addressed such issues was Walter Mischel's work on delay of gratification (e.g., Mischel, 1973). Using children as experimental participants, Mischel and his collaborators demonstrated the difficulty and often their failure to forgo a smaller but immediate reward for a larger but delayed gratification. Moreover, these researchers identified some conditions that prevent or foster the capability to delay. In particular, it has been found that directing the children's attention toward the larger delayed reward does not necessarily increase their ability to wait. Instead, it was found that children were even more likely to fall for the immediate reward (Mischel, 1973). This result flies in the face of theorists of rational choice (e.g., Frederick et al., 2002), who have argued that preferences for a smaller immediate gratification are due to a disproportionate discounting of the future outcome. Drawing people's attention to this outcome certainly reminds them of the value and reduces tendencies of discounting. At the same time, however, it increases impulsive tendencies of approaching the positively valued outcome, which is only possible for the immediate outcome. Thus, to understand participants' "irrationality" in intertemporal choice, it is not sufficient to assess people's evaluation of the outcome. Rather, it is necessary to take impulsive mechanisms that are not guided by knowledge or beliefs into account.

The second line of research that addresses the issue of temptation is self-regulation (e.g., Baumeister & Vohs, 2004). Here, temptation is seen as a force that has to be counteracted by the "power of the will" (Baumeister & Vohs, 2003). Specifically, willpower is fueled by a source of energy that will be depleted if too much of it is spent. This notion of energy depletion goes beyond mere attentional models in that it also postulates sequential effects. That is, the power of the will continues to be weakened after it has been employed to ward off tempting forces. While this notion of energy is certainly useful for a better understanding of self-regulation, it says very little about the mental processes that are involved if people resist or yield to tempting influences. It seems that recent research on implicit mechanisms may open a new avenue of research that has important clinical implications for the study of addiction (e.g., Deutsch & Strack, 2006; Wiers, Van de Luitgaarden, Van den Wildenberg, & Smulders, 2005).

CONCLUSION

This chapter has attempted to direct readers' attention to behavioral influences that are acknowledged less frequently in current social psychology than in other domains. In particular, the possibility that a behavior may be determined not only by the evaluation of its anticipated consequences has been recognized more often in the areas of personality and developmental and biological psychology than by social psychologists. Human behaviors that are influenced by factors exerting a more direct influence, and sometimes even standing in conflict with what people believe is good for them, have often been called impulsive, and the same label has been used to describe stable personality characteristics. Moreover, specific neuronal mechanisms have been identified that are in charge of impulsive (as opposed to reflective) processing.

In this contribution, we tried to reintegrate impulsive variants of social behavior into social psychology and used our own two-systems theory as a conceptual frame. In closing, we would like to suggest a taxonomy to classify impulsive behaviors. It has two dimensions that seem to be relevant. First, an impulsive behavior may be a manifestation of a *general* behavioral tendency or may be *specifically* directed at a more narrow class of stimuli. At the same time, the impulsive behavior may be *hot* or *cold*; that is, it may be accompanied by affect or valence or not. The resulting table consists of four cells that can be described as follows.

Impulsive behaviors that fall into the category of "general" and "hot" are illustrated by the results of a motivational orientation. The facilitation of a behavior that is accompanied by positive affect would be an example. The "specific" and "hot" combination is best illustrated by behaviors that originate from a homeostatic dysregulation, like hunger and thirst. Habits, in contrast, can be described as "specific" and "cold." Finally, the cell resulting from "general" and "cold" describes a behavioral disposition toward approach and avoidance. As mentioned earlier, it is also important to distinguish between behavioral dispositions toward approaching something desirable versus avoiding something undesirable and motivational orientations toward using eager approach strategies versus using vigilant avoidance strategies (Higgins, 1997).

The RIM that we have proposed as a conceptual framework to understand impulsive mechanisms is only a start. Beyond integrating aspects of human actions that are not

captured by studying the actors' beliefs about facts and values, the RIM attempts to shed light on phenomena that have been largely neglected by social psychologists. Because some of these behaviors (e.g., addictive or compulsive behaviors) may transcend the traditional borders of the field, both social and abnormal psychology may profit from a closer collaboration. On the one hand, insights from social psychology may provide a more basic understanding of the mechanisms underlying behavioral disorders. Equally important, abnormal behavioral manifestations may shed light on the processes that produce regular social phenomena. Just as visual illusions led to a deeper understanding of human perception, and judgmental biases led to identifying basic mechanisms of judgment and decision making, behavioral disorders have the potential to help us discover important processes that underlie social behavior.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–884). Worcester, MA: Clark University Press.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, *53*, 27–51.

Andrés, P. (2003). Frontal cortex as the central executive of working memory: Time to revise our view. *Cortex*, *39*, 871–895.

Aquinas, Saint Thomas. (2002). *Aquinas's Shorter Summa*. Manchester, NH: Sophia.

Asberg, M., Thoren, P., Traskman, L., Bertilsson, L., & Ringberger, V. (1976). "Serotonin depression": A biochemical subgroup within the affective disorders? *Science*, *191*, 478–480.

Asendorpf, J. B., & Van Aken, M. A. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology*, *77*, 815–832.

Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality*, *15*, 327–353.

Baddeley, A. (1986). *Working memory* (Oxford Psychology Series, No. 11). New York: Clarendon Press/Oxford University Press.

Baddeley, A. D., & Baddeley, A. (1990). *Human memory: Theory and practice*. Needham Heights, MA: Allyn & Bacon.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, J. S. Nairne, I. Neath, & A. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.

Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). *Social learning theory*. Oxford, UK: Prentice-Hall.

Banfield, J. F., Wyland, C. L., Macrae, C. N., Munte, T. F., & Heatherton, T. F. (2004). The cognitive neuroscience of self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 62–83). New York: Guilford Press.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. A. Bargh & J. S. Uleman (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (pp. 1–61). Mahwah, NJ: Erlbaum.

Bargh, J. A. (2005). Bypassing the will: Towards demystifying the nonconscious control of social behavior. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 37–58). New York: Oxford University Press.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of everyday life. *Psychological Bulletin*, *126*, 925–945.

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. In W. D. Spaulding (Ed.), *Integrative views of motivation, cognition, and emotion* (Vol. 41). Lincoln: University of Nebraska Press.

Baumeister, R. F. (2002). Ego depletion and self-control failure: An energy model of the self's executive function. *Self and Identity*, *1*, 129–136.

Baumeister, R. F., & Vohs, K. D. (2003). Willpower, choice, and self-control. In G. Loewenstein, D. Read, & R. F. Baumeister (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice*. New York: Russell Sage.

Baumeister, R. F., & Vohs, K. D. (2004). *Handbook of self-regulation: Research, theory, and applications*. New York: Guilford Press.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, *8*, 1458–1463.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.

Becker, G. (1962). Investment in human capital: A theoratical analysis. *Journal of Political Economy*, *70*, 9–49.

Becker, G. (1976). *The economic approach to human behavior*. Chicago: University of Chicago Press.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*, 183–200.

Berkowitz, A. D. (2004). Men's role in preventing violence against women. *Applied Research Forum of VAWNet*.

Berkowitz, L. (1974). Some determinants of impulsive aggression: Role of mediated associations with reinforcements for aggression. *Psychological Review*, *81*, 165–176.

Berkowitz L. (1990). On the formation and regulation of anger and aggression: A cognitiveneoassociationistic analysis. *American Psychologist*, *45*, 494–503.

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control (McGraw-Hill series in social psychology)*. New York: McGraw-Hill.

Berlin, H. A., Rolls, E. T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain*, *127*, 1108–1126.

Berman, M. E. (1997). Biopsychosocial approaches to understanding human aggression: The first 30 years. *Clinical Psychology Review*, *17*, 585–588.

Berman, M. E., Tracy, J. I., & Coccaro, E. F. (1997). The serotonin hypothesis of aggression revisited. *Clinical Psychology Review*, *17*, 651–665.

Berridge, V. (2002). Witness seminar: The black report and the health divide. *Contemporary British History*, *16*, 131–172.

Bizot, J. C., Bihan, C. L., Puéch, A. J., Hamon, M., & Thiébot, M. H. (1999). Serotonin and tolerance to delay of reward in rats. *Psychopharmacology*, *146*, 400–412.

Bizot, J. C., Thiébot, M. H., le-Bihan, C., Soubrie, P., & Simon, P. (1988). Effects of imipramine-like drugs and serotonin uptake blockers on delay of reward in rats. Possible implication in the behavioral mechanism of action of antidepressants. *Journal of Pharmacology and Experimental Therapeutics*, *246*, 1144–1151.

Bjork, J. M., Dougherty, D. M., Moeller, F. G., Cherek, D. R., & Swann, A. C. (1999). The effects of tryptophan depletion and loading on laboratory aggression in men: Time course and a food-restricted control. *Psychopharmacology*, *142*, 24–30.

Bjork, J. M., Dougherty, D. M., Moeller, F. G., & Swann, A. C. (2000). Differential behavioral effects of plasma tryptophan depletion and loading in aggressive and nonaggressive men. *Neuropsychopharmacology*, *22*, 357–369.

Bless, H., Fiedler, K., & Strack, F. (2004). *Social cognition: How individuals construct social reality*. Hove, UK: Psychology Press.

Block, J. (2002). *Personality as an affect-processing system: Toward an integrative theory*. Mahwah, NJ: Erlbaum.

Block, J. H., & Block, J. (1980). Ego control and ego resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 13, pp. 39–101). Hillsdale, NJ: Erlbaum.

Brennan, P. A., & Raine, A. (1997). Biosocial bases of antisocial behavior: Psychophysiological, neurological, and cognitive factors. *Clinical Psychology Review*, *17*, 589–604.

Bywater, I. (Ed.). (1984). *Aristotelis ethica nichomachea* [Scriptorum Classicorum Bibliothea Oxoniensis]. Oxford, UK: Clarendon Press.

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, *65*, 5–17.

Cardinal R. N. (2000). Whisker (version 1.0, computer software). Available online: http://www. pobox.com/users/rudolf/whisker.

Cardinal, R. N., Daw, N., Robbins T. W., & Everitt B. J. (2002). Local analysis of behaviour in the adjusting-delay task for assessing choice of delayed reinforcement. *Neural Networks*, *15*, 617–634.

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews*, *26*, 321–352.

Cardinal, R. N., Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., & Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, *292*, 2499–2501.

Cardinal, R. N., Robbins, T. W., & Everitt, B. J. (2003). Choosing delayed rewards. Perspectives from learning theory, neurochemistry, and neuroanatomy. In R. E. Vuchinich & N. Heather (Eds.), *Choice, behavioral economics and addiction* (pp. 183–218). Amsterdam: Elsevier.

Cardinal, R. N., Winstanley C. A., Robbins T. W., & Everitt B. J. (2004). Limbic corticostriatal systems and delayed reinforcement. *Annals of the New York Academy of Sciences*, *1021*, 33–50.

Carver, C. S. (2004). Self-regulation of action and affect. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 13–39). New York: Guilford Press.

Carver, C. S. (2005). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review*, *9*(4), 312–333.

Cesario, J., Plaks, J. E., & Higgins, E. T. (2005). *Automatic social behavior as motivated preparation to interact*. Unpublished manuscript, Columbia University, New York.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Cherek, D. R., & Lane, S. D. (1999). Effects of d,l-fenfluramine on aggressive and impulsive responding in adult males with history of conduct disorder. *Psychopharmacology*, *146*, 473–481.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science*, *280*, 77–81.

Clower, C. E., & Bothwell, R. K. (2001). An exploratory study of the relationship between the Big Five and inmate recidivism. *Journal of Research in Personality*, *35*, 231–237.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*, 1314–1329.

Craen, J., Richards, J. B., & De Wit, H. (2002). Effect of tryptophan depletion on impulsive behavior in men with or without a family history of alcoholism. *Behavioural Brain Research*, *136*, 349–357.

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In I. S. Csikszentmihalyi & M. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15–35). New York: Cambridge University Press.

DeCoster, J., & Claypool, H. M. (2004). A meta-analysis of priming effects on impression formation supporting a general model of informational biases. *Personality and Social Psychology Review*, *8*, 2–27.

Deutsch, R., Gawronski, B., & Strack, F. (2005). *At the boundaries of automaticity: Negation as reflective operation*. Manuscript submitted for publication.

Deutsch, R., & Strack, F. (2006). Reflective and impulsive determinants of addictive behavior. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*, 835–848.

Devine, P. G., Plant, E. A., & Buswell, B. N. (2000). Breaking the prejudice habit: Progress and obstacles. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 185–208). Mahwah, NJ: Erlbaum.

Dickinson, A., Smith, J., & Mirenowicz, J. (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral-Neuroscience*, *114*, 468–483.

Dollard, J., Miller, N. E., Doob, L. W., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Eisenberg, N., Valiente, C., Fabes, R. A., Smith, C. L., Reiser, M., Shepard, S. A., et al. (2003). The relations of effortful control and ego control to children's resiliency and social functioning. *Developmental Psychology*, *39*, 761–776.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*, 709–724.

Epstein, S. (2003). Cognitive–experiential self-theory of personality. In T. Millon & M. J. Lerner (Eds.), *Comprehensive handbook of psychology, Vol. 5. Personality and social psychology* (pp. 159–184). Hoboken, NJ: Wiley.

Evenden, J. (1999). Varieties of impulsivity. *Psychopharmacology*, *146*, 348–361.

Eysenck, H. J. (1970). *The structure of human personality*. New York: Methuen.

Eysenck, H. J. (1992). Personality and education: The influence of extraversion, neuroticism and psychoticism. *Zeitschrift für Pädagogische Psychologie*, *6*, 133–144.

Farrelly, M. C., Healton, C. G., Davis, K. C., Messeri, P., Hersey, J. C., & Haviland, M. L. (2002). Getting to the truth: Evaluating national tobacco countermarketing campaigns. *American Journal of Public Health*, *92*, 901–907.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*, 115–141.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology*, *54*, 297–327.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research*, *16*, 280–288.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior*. Reading, MA: Addison-Wesley.

Fitzsimons, G. M., & Bargh, J. A. (2003). Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology*, *84*, 148–164.

Förster, J., & Strack, F. (1996). Influence of over head movements on memory for valenced words: A case of conceptual-motor compatibility. *Journal of Personality and Social Psychology*, *71*, 421–430.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review, *Journal of Economic Literature*, *40*, 351–401.

Freud, S. (1920). *Beyond the pleasure principle*. London: Hogarth.

Freud, S. (1921). *Group psychology and the analysis of the ego*. London: Hogarth.

Friedman, H. S., Tucker, J. S., Schwarz, J., E., Martin, L. R., Tomlinson-Keasey, C., et al. (1995). Childhood conscientiousness and longevity: Health behaviors and causes of death. *Journal of Personality and Social Psychology*, *68*, 696–703.

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212–228.

Funder, D. C., & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology*, *57*, 1041–1050.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). *Cognitive neuroscience: The biology of the mind* (2nd ed.). New York: Norton.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Grant, S. J., Malaviya, P., & Sternthal, B. (2004). The influence of negation on product evaluations. *Journal of Consumer Research*, *31*, 583–591.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Harrison, A. A., Everitt, B. J., & Robbins, T. W. (1999). Central serotonin depletion impairs both the acquisition and performance of a symmetrically reinforced go/no-go conditional visual discrimination. *Behavioral Brain Research*, *1000*, 99–112.

Hawkins, K. A., & Bender, S. (2002). Norms and the relationship of Boston naming test performance to vocabulary and education: A review. *Aphasiology*, *16*, 1143–1153.

Hawkins, K. A., & Trobst, K. K. (2000). Frontal lobe dysfunction and aggression: Conceptual issues and research findings. *Aggression and Violent Behavior*, *5*, 147–157.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*, 1140–1153.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology*, *13*, 141–154.

Hogan, J., & Ones, D. S. (1997). Conscientiousness and integrity at work. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 849–870). San Diego, CA: Academic Press.

Horvarth, P., & Zuckerman, M. (1993). Sensation seeking, risk appraisal, and risky behavior. *Personality and Individual Differences*, *14*, 41–52.

Hume, D. (2000). *A treatise on human nature*. Oxford, UK: Oxford University Press. (Original work published 1739)

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. *Brain Research Reviews*, *31*, 6–41.

James, W. (1890). *The principles of psychology* (Vol. 2). New York: Holt.

Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality*, *69*, 323–362.

Johnson, M. K., & Hirst, W. (1991). Processing subsystems of memory. In R. G. Lister & H. J. Weingartner (Eds.), *Perspectives on cognitive neuroscience* (pp. 197–217). New York: Oxford University Press.

Jones, M. C. (1971). Personality antecedents and correlates of drinking patterns in women. *Journal of Consulting and Clinical Psychology*, *36*, 61–69.

Judd, C. M., Blair, I. V., & Chapleau, K. M. (2004). Automatic stereotypes vs. automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social Psychology*, *40*, 75–81.

Kagan, J. (1965). Impulsive and reflective children. In J. Krumboltz (Ed.), *Learning and the educational process* (pp. 133–161). Chicago: Rand McNally.

Kagan, J., & Snidman, N. (1991). Temperamental factors in human development. *American Psychologist*, *46*, 856–862.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In D. Griffin & T. Gilovich (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.

Kelley, A. E., & Berridge, K. C. (2002). The neuroscience of natural rewards: Relevance to addictive drugs. *Journal of Neuroscience*, *22*, 3306–3311.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–239). Lincoln: University of Nebraska Press.

Kremen, A. M., & Block, J. (1998). The roots of ego-control in young adulthood: Links with parenting in early childhood. *Journal of Personality and Social Psychology*, *75*, 1062–1075.

Krueger, R. F. (2002). Personality from a realist's perspective: Personality traits, criminal behaviors, and the externalizing spectrum. *Journal of Research in Personality*, *36*, 564–572.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystanders: Why doesn't he help*. New York: Appleton-Century-Crofts.

Le Bon, G. (1895). *The crowd*. New York: Viking Press.

Lhermitte, F. (1983). "Utilization behavior" and its relation to lesions of the frontal lobes. *Brain*, *106*, 237–255.

Liberman, N., & Trope, Y. (2003). Temporal construal theory of intertemporal judgment and decision. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and choice: Economic and psychological perspectives on intertemporal choice* (pp. 245–276). New York: Sage.

Liu, Y. P., Wilkinson, L. S., & Robbins, T. W. (2004). Effects of acute chronic buspirone on impulsive choice and efflux of 5-HT and dopamine in hippocampus, nucleus accumbens and prefrontal cortex. *Psychopharmacology*, *173*, 175–185.

Loewenstein, G. F. (2001). A visceral account of addiction. In P. Slovic (Ed.), *Smoking: Risk, perception, and policy* (pp. 188–215). Thousand Oaks, CA: Sage.

Loewenstein, G. F., & Elster, J. (1992). *Choice over time*. New York: Russell Sage.

Lotze, H. R. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmann.

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science*, *16*, 6–10.

Martin, R. A., Phulik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, *37*, 48–75.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*, 433–449.

McClure, S. M., York, M. K., & Montague, P. R. (2004). The neural substrates of reward processing in humans: The modern role of functional magnetic resonance imaging. *The Neuroscientist*, *10*, 260–268.

McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509–516.

Messer, S. B. (1976). Reflection-impulsivity: A review. *Psychological Bulletin*, *83*, 1026–1052.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.

Mischel, W. (1973). Cognition in delay of gratification. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium* (pp. 213–229). Oxford, UK: Winston.

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*, 329–337.

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*, 204–218.

Mobini, S., Body, S., Ho, M. Y., Bradshaw, C. M., Szabadi, E., Deakin, J. F. W., et al. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology*, *160*, 290–298.

Moeller, F. G., & Dougherty, D. M. (2002). Impulsivity and substance abuse: What is the connection? *Addictive Disorders and Their Treatment*, *1*, 3–10.

Mussweiler, T., & Strack, F. (1999a). Comparing is believing: A selective accessibility model of judgmental anchoring. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 10, pp. 135–167). Chichester, UK: Wiley.

Mussweiler, T., & Strack, F. (1999b). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, *35*, 136–164.

Neumann, R., Förster, J., & Strack, F. (2003). Motor compatibility: The bidirectional link between behavior and evaluation. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–49). Mahwah, NJ: Erlbaum.

Newman, L. S., & Uleman, J. S. (1989). Spontaneous trait inference. In J. A. Bargh & J. S. Uleman (Eds.), *Unintended thought* (pp. 155–188). New York: Guilford Press.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and social shortcomings of social judgment*. London: Prentice Hall.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*, 166–180.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.

Ouellette, J., & Wood, W. (1998). Habit and intention in everyday life. The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*, 54–74.

Palfai, T. B., Monti, P. M., Colby, S. M., & Rohsenow, D. J. (1997). Effects of suppressing the urge to drink on the accessibility of alcohol outcome expectancies. *Behavior Research and Therapy*, *35*, 59–65.

Parkinson, J. A., Dalley, J. W., Cardinal, R. N., Bamford, A., Fehnert, B., Lachenal, G., et al. (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: Implications for meso-accumbens dopamine function. *Behavioral Brain Research*, *137*, 149–163.

Paschall, M. J., & Fishbein, D. H. (2002). Executive cognitive functioning and aggression: A public health perspective. *Aggression and Violent Behavior*, *7*, 215–235.

Paus, T. (2001) Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Natural Reviews: Neuroscience*, *2*, 417–424.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181–192.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Plato. (1998). *Phaedrus* (J. H. Nichols, Jr., Trans.). Ithaca, NY: Cornell University Press.

Pratkanis, A. R. (1989). The cognitive representation of attitudes. In R. Pratkanis, S. Breckler, & S. Greenwald (Eds.), *Attitude structure and function* (pp. 70–98). Hillsdale, NJ: Erlbaum.

Putnam, S. P., & Stifter, C. A. (2005). Behavioral approach-inhibition in toddlers: Prediction from infancy, positive and negative affective components, and relations with behavior problems. *Child Development*, *76*, 212–226.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, *57*, 119–127.

Richards, J. B., Mitchell, S. H., De Wit, H., & Seyden, L. (1997). Determination of discount functions with an adjusting amount procedure in rats. *Journal of Experimental Analysis of Behavior*, *67*, 353–366.

Richards, J. B., Sabol, K. E., & De Wit, H. (1999). Effects of methamphetamine on the adjusting amount procedure, a model of impulsive behavior in rats. *Psychopharmacology*, *146*, 432–439.

Robbins, R., & Bryan, A. (2004). Relationships between future orientation, impulsive sensation seeking, and risk behavior among adjudicated adolescents. *Journal of Adolescent Research*, *19*, 428–445.

Roberts, C. (2002). Demonstratives as definites. In K. van Deemter & R. Kibble (Eds.), *Information sharing* (pp. 89–136). Stanford, CA: CSLI.

Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin*, *128*, 473–489.

Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annual Review of Psychology*, *54* , 25–53.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, *10*, 284–294.

Rosekrans, M. A., & Hartup, W. W. (1967). Imitative influences of consistent and inconsistent response consequences to a model

on aggressive behavior in children. *Journal of Personality and Social Psychology*, 7, 429–434.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In N. Eisenberg & W. Damin (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (5th ed., Vol. 3, pp. 105–176). New York: Wiley.

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (pp. 37–86). Hillsdale, NJ: Erlbaum.

Salvucci, D. D. (2005). A multitasking general executive for compound continuous tasks. *Cognitive Science*, 29, 457–492.

Schachter, S., & Singer, J. E. (1962). Cognitive, social and psychological determinants of emotional state. *Psychological Review*, 69, 379–399.

Schmeichel, B. J., & Baumeister, R. F. (2004). Self-regulatory strength. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 84–98). New York: Guilford Press.

Schwarz, N. (2000). Social judgment and attitudes: Warmer, more social, and less conscious. *European Journal of Social Psychology*, 30, 149–176.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513–523.

Seibt, B., Häfner, M., & Deutsch, R. (2005). *Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation*. Manuscript under review.

Shachar, R., Tannock, R., & Logan, G. (1993). Inhibitory control, impulsivity, and attention deficit hyperactivity disorder. *Clinical Psychology Review*, 13, 721–739.

Shedler, J., & Block, J. (1990). Adolescent drug use and psychological health: A longitudinal inquiry. *American Psychologist*, 45, 612–630.

Simon, H. A. (1991). Bounded rationality and organizational learning. *Organization Science*, 2, 125–134.

Skumik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, 31, 713–724.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.

Smith, A. (1981). *An inquiry into the nature and causes of the wealth of nations*. Indianapolis, IN: Liberty Press. (Original work published 1776)

Smith, E. R. (1998). Mental representation and memory. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 391–445). New York: McGraw-Hill.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108–131.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. Oxford, UK: Wiley.

Soubrie, P. (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behavioral and Brain Sciences*, 9, 319–335.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, 37, 1660–1672.

Srull, T. K., & Wyer, R. S., Jr. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology*, 38, 841–856.

Stepper, S., & Strack, F. (1993). Proprioceptive determinants of emotional and nonemotional feelings. *Journal of Personality and Social Psychology*, 64, 211–220.

Strack, F. (1992). The different routes to social judgments: Experiential vs. informational strategies. In L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.

Strack, F., & Deutsch, R. (2005). Reflection and impulse as determinants of conscious and unconscious motivation. In J. Forgas, K. Williams, & S. Laham (Eds.), *Social motivation: Conscious and unconscious processes* (pp. 91–112). Cambridge, UK: Cambridge University Press.

Strack, F., & Hannover, B. (1996). Awareness of the influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking motivation and cognition to behavior* (pp. 579–596). New York: Guilford Press.

Strack, F., Martin, L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A non-obtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54, 768–777.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73, 437–446.

Strack, F., & Mussweiler, T. (2003). Heuristic strategies for estimation under uncertainty: The enigmatic case of anchoring. In G. Bodenhausen & A. Lambert (Eds.), *Foundations of social cognition* (pp.79–95). Mahwah, NJ: Erlbaum.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). New York: Guilford Press.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Thurstone, L. L. (1931). The measurement of social attitudes. *Journal of Abnormal and Social Psychology*, 26, 249–269.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, 93, 239–257.

Tversky, A., & Kahnemann, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.

Verheyden, S. L., Hadfield, J., Calin, T., & Curran, H. V. (2002). Sub-acute effects of MDMA (3,4-methylenedioxymethamphetamine, "ecstasy") on mood: Evidence of gender differences. *Psychopharmacology*, 161, 23–31.

Verplanken, B., & Aarts, H. (1999). Habit, attitude and planned behavior: Is habit an empty construct or an interesting case of goal-directed automatic? *European Review of Social Psychology*, 10, 101–134.

Verplanken, B., Aarts, H., & Van Knippenberg, A. (1997). Habit, information acquisition, and the process of making travel mode choices. *European Journal of Social Psychology*, 27, 539–560.

Wagner, M. K. (2001). Behavioral characteristics related tov substance abuse and risk-taking, sensation- seeking, anxiety sensitivity and self-reinforcement. *Addictive Behaviors*, 26, 115–120.

Walderhaug, E., Lunde, H., Nordvik, J. E., Landro, N. I., Refsum, H., & Magnusson, A. (2002). Lowering of serotonin by rapid tryptophan depletion increases impulsiveness in normal individuals. *Psychopharmacology*, 164, 385–391.

Wegner, D. M., Ansfield, M., & Pilloff, D. (1998). The putt and the pendulum: Ironic effects of the mental control of action. *Psychological Science*, 9, 196–199.

Wiers, R. W., Van de Luitgaarden, J., Van den Wildenberg, E., & Smulders, F. T. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction*, 100, 806–819.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Experimental Psychology General*, 125, 387–402.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107, 101–126.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facili-

tation elicits positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winstanley, C. A., Theobald, D. E., Dalley, J. W., Glennon, J. C., & Robbins, T. W. (2004). 5-HT-sub(2A) and 5-HT-sub(2C) receptor antagonists have opposing effects on a measure of impulsivity: Interactions with global 5-HT depletion. *Psychopharmacology*, *176*, 376–385.

Wolff, M. C., & Leander, J. D. (2002). Selective serotonin reuptake inhibitors decrease impulsive behavior as measured by an adjusting delay procedure in the pigeon. *Neuropsychopharmacology*, *27*, 421–429.

Wyvell, C. L., & Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: Enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *Journal of Neuroscience*, *20*, 8122–8130.

Young, S. N., & Leyton, M. (2002). The role of serotonin in human mood and social interaction: Insight from altered tryptophan levels. *Pharmacology, Biochemistry and Behavior*, *71*, 857–865.

Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences*, *7*, 413–471.

CHAPTER 18

Social Identity and Self-Regulation

DAPHNA OYSERMAN

More than simply a store of autobiographical knowledge, self-concept is one's theory about oneself (Brown, 1998). It functions to organize past and present experience, illuminate one's future possibilities, sustain motivation, and control behavior in pursuit of the selves one might become. It provides answers to the basic self questions "Who am I?" and "How do I fit in?" and functions as a roadmap detailing how one goes about being oneself. Self-concept both feels stable, allowing one to answer the "Who am I" question by responding "Me," but is also fluid. Fluidity is experienced both as open potential—allowing one to believe in one's ability to grow, improve, and change—and as the result of automatic responsivity to situational cues. In this sense, who one is depends on what is relevant in the situation and what people who are like oneself seem to be doing.

A basic premise of this chapter is that motivation is identity based. Situational cues about how to be a self are assimilated into one's working self-concept except when these cues set up a contrasting standard of things "they" but not "we" do, feel, or strive to achieve. Individuals are motivated to pursue the goals ingroup members pursue using the means ingroup members use. What these goals and strategies are is something that is contextually cued.

This chapter focuses on an aspect of self-concept—social identity—and an aspect of the process of being a self—self-regulation. Self-regulation is the coordination of affective and behavioral processes to maintain a reasonably positive sense of oneself while behaving in a socially appropriate manner and working toward one's goals. Whether conceptualized in terms of action or inac-

tion, self-regulation links the present, one's current self and current behavior, with the future, one's possible selves and longer-term goal pursuit (Oyserman, Bybee, Terry, & Hart-Johnson, 2004). Self-regulation involves individuals engaging in or refraining from behavior in the immediate or ongoing present to increase the odds of attaining self-relevant goals later. Thus, self-regulation evokes both behavioral inhibition and behavioral activation systems (Avila, 2001). Individuals are motivated to do what ingroup members do and to avoid doing that which ingroup members do not do. In that sense social identity is central to self-regulation.

By using a social identity perspective (Abrams, 1999; Onorato & Turner, 2002) and explicitly connecting social identity and cultural psychology perspectives (Triandis, 1989) a basic convergent outline of social identity emerges. From both cultural and social identity perspectives, the self-concept is conceived of as fundamentally social. Social contexts influence content of self-concept, and one of the major goals of self-concept is to provide a sense of fit with and integration into a larger social whole. While social identity theories emphasize momentary shifts in situation and cross-situation differences in salience of personal versus social role-based identities, cultural psychology theories place more emphasis on chronic or stable situations and cross-national differences in salience of personal identities versus social role-based identities. By integrating these separate but compatible theories, their complementary theoretical assumptions about the social nature of self-concept can be joined. However, because neither cultural nor social

identity frameworks were intended as process models of how social identities influence self-regulation, an integration of these two models alone is insufficient as a process model of how social identity influences self-regulation. To create a process model, an identity-based motivation model is proposed. It is based in a self-schema framework (Markus, 1977) and links cultural and social identity perspectives to a broader social cognition framework (Higgins, 1996; Schwarz & Bless, in press) that outlines how social contexts influence social identities in ways likely to shift motivation and self-regulatory success.

The basic principles that guide this chapter are (1) that individuals are influenced by what comes to mind when making a judgment and (2) that what it is that comes to mind can be contextually or chronically cued. All things being equal, individuals assume that what comes to mind is relevant; in the case of social identity and self-regulation, what comes to mind is assumed relevant to the things "we" do, feel, or believe. What this "we" is, is perceived as stable and even central to identity but may shift over time. Images of what "we" do provide an outline of one's possible future, sketching out both the possible selves "we" can become and the kinds of strategies "we" use to attain these self-relevant goals. When possible selves thus articulated are linked with effective strategies they improve self-regulatory success. Conversely, when the possible selves thus articulated are linked with ineffective strategies they undermine self-regulatory success. In this way, social identities turn on self-regulation by turning on motivation to act like an ingroup member and engage in the pursuits that characterize ingroup members. Social identities provide both reasons to act and reasons not to act, and also ways to act or avoid action to attain goals. They not only cue us to try but also suggest standards for what trying looks like—what we do, what constitutes sufficient effort for us, and so on.

This basic perspective is congruent with current social identity models that suggest that social identities incorporate both positively valenced feelings of connection and specific group-defining attributes. It is also congruent with parallel arguments presented from a cultural perspective that cuing social connection makes salient social aspects of identity. Integrating these perspectives with a social cognition model allows for new predictions. Using an identity-based motivation perspective provides a mechanism to begin to explore otherwise puzzling discrepancies between espoused goals and self-regulatory behavior. While a social identity perspective proposes that all individuals have chronic propensity to make and maintain social connections and to define themselves in terms of these connections, a cultural psychology perspective emphasizes between-person and between-culture chronic differences in the propensity to focus on social connections. Neither perspective alone provides an articulation of how social situations cue self-judgments, when cued information will be included in or excluded from self-judgments, and how this influences self-relevant action over time.

Cultural psychology has proposed relatively stable differences (based in history, socialization, and social institutions) in the propensity to define the self and the social world in terms of groups and embeddedness within groups as well as relatively stable differences in the kinds of groups that are self-defining (e.g., friendship, family, religious, and tribal). Accumulating evidence suggests that chronic differences do exist but within a more malleable context than a stable differences perspective would allow. Moreover, while a cultural perspective suggests a dichotomized perspective in which the self is defined as either social or personal, an identity-based motivation perspective fills out what is meant by a "social" identity—suggesting that these identities also contain traits, propensities, and characteristics that motivate action. It seems likely, as suggested by social identity perspectives, that when social identities are cued, self-defining traits, propensities, and characteristics are those assumed to be ingroup defining.

This is an important advance because it suggests that motivation is not either personal or social but rather simultaneously socially based and personalized. Unfortunately, cultural psychology has not moved much beyond documenting that social identities may be more central to self-definitions in some cultures (and situations) than others. While, as outlined in the following sections, this lack of progress in cultural psychology may be due to the nature of the tasks cultural psychologists use to study self-concept, it is clear that when socially primed, social identities are evoked and these social identities are likely to contain attributes that feel ingroup congruent.

Perhaps most important, while socially based ingroup defining attributes feel distinct, they may or may not be different from attributes characterizing other groups. In some cases, these ingroup defining attributes may be defined explicitly as the opposite of or in direct contrast to the attributes of another social identity group. However, this is likely to be a special case rather than the norm. Indeed, when self-definition requires contrast with another social group, it can be undermining of self-regulatory ability if the other social group has control of important social goals or of effective strategies to attain these goals.

Just as ingroup defining attributes and valued goals may actually be common across groups, so may be strategies to attain them. Of particular interest are situations in which goals are common but ingroups differ in the extent that various strategies are seen as ingroup-relevant ways to attain these goals. For example, both men and women may be able to claim leadership goals as ingroup defining. However, to the extent that effective assertive or aggressive strategies to pursue this goal are "male," women may be more likely to use less effective strategies—and fall short of their leadership goal. Similarly, both boys and girls may be able to claim academic success as an ingroup defining goal. However, to the extent that effective strategies to attain this goal—studying, paying attention in class, following teacher instructions, handing in assignments that are neat and tidy, asking for help—are considered "female" things to do, then boys may be more likely to use less effective strategies—and fall short of their successful-in-school goal.

These are issues that can only be studied by thinking about the power of social identities. Individual women may want to lead and individual boys may want to do well

in school and may in fact be aware of appropriate strategies to effectively attain their goals. However, once social identities are contextually evoked, effective strategies may no longer feel appropriate. In this sense, social identities can be considered the most basic way in which we define ourselves. Once cued, evoked, or turned on, they override individual goals and aspirations unless individual goals and aspirations are sensed as congruent with social identities. In much the same way, social identities seen as more basic can override other social identities unless the various social identities are construed as compatible. Rather than being a woman or a leader, a boy or a scholar, compound social identities—female leader, athlete scholar, or future leader of one's community— allow for integration of goals and otherwise group-incompatible strategies. In each of the following sections, I outline how social identity has been approached, the additional utility of incorporating a cultural psychology framework, and advances made by integrating social identity, cultural psychology, and social cognition perspectives into an identity-based motivation model.

SOCIAL IDENTITY

Social identity theories have historically assumed a distinction between the self as defined by group memberships (the collective or social self) and the self as defined individually (the private self) (Hogg, 2003). All individuals can and do define themselves in both ways, switching between levels of self-definition depending on social contextual cues as to which level is relevant or useful in the moment (Turner & Oakes, 1986). Rather than describing identities as simply social as opposed to personal or private, social identities can be separated into those that focus on memberships in larger groups—collective identities—and those that focus on specific, face-to-face or personal relationships—termed "relational identities" (Brewer, 1991; Brewer & Gardner, 1996).

Collective social self-concepts contain information about the social categories to which one belongs, one's group memberships, as well as information about what members of one's groups are like, how they act, what they care about, and what their goals and values are (Abrams, 1994). Relational social self-concepts contain information about the specific relationships one is part of as well as how one is defined in relation to these specific others (Cross, Gore, & Morris, 2003; Cross, Morris, & Gore, 2002). Collective social identities focus on larger group identities such as those connected to memberships in a gender, racial–ethnic, nationality, religious, tribal, social class, or regional group. Relational identities focus on memberships in particular relationships— friendships, family, marital, peer, or work groups. Some social identities such as sports fan, fraternity member, student, or employee highlight the ambiguity of these distinctions in that any social identity could define both a particular relational identity and a general collective identity.

According to social identity theorists, social identities are at the heart of self-concept (Tajfel, 1972; Turner & Oakes, 1989). From a social identity perspective, we first attempt to make sense of the social world in terms of social categories and social category memberships and use individuating information only if category membership does not apply. For example, men are faster at responding "not me" to words previously rated as feminine after being primed to think of "I," "me," or "my" in a lexical decision task, presumably because thinking of themselves brings to mind their belongingness to the social category "men" and carries with it all the things that men are and are not (Mussweiler & Bodenhausen, 2002).

Social identities are hot social categories. They include a positive feeling of being included in some groups, a valenced affective response to being excluded from other groups, and concomitant positive feelings about ingroup defining attributes and negative feelings about outgroup defining attributes. That which is included in one's social identity is rated more positively than that which is excluded from it. For example, Reed (2004) showed that being primed to think of "we" increased liking for objects associated with ingroup. After using the "we" priming paradigm, a palm pilot described as a way to stay connected with family was more liked than when the palm pilot was not linked to family. Similarly, participants primed with "we" rated ambiguous statements as more similar to their own beliefs than participants primed with "they" (Brewer & Gardner, 1996).

In addition to gender and family, racial–ethnic, religious, and other social groups or categories may be incorporated into self-concept as social identities. Research has demonstrated influences on self-esteem, motivation, and self-regulation from categorizing oneself in terms of membership in a diverse array of groups including racial–ethnic groups (Oyserman, Gant, & Ager, 1995), gender (Schmader, 2002), cultural groups (Seeley & Gardner, 2003), and other culturally meaningful groups, such as blondes or athletes (Seibt & Förster, 2004). In each of these cases, positive and negative social stereotypes about in- and outgroup members exist.

In this way, social identities can be thought of as self-stereotypes in that they are generalizations about groups to which one belongs that influence the sense one makes of who one is and can become and one's place in the social world (Sherman, Judd, & Parke, 1989; Wilson & Dunn, 2004). Because social identities are part of self-concept, they can be used to make predictions about how others will respond to the self as well as what is likely to feel good and what one is likely to do well at. To be useful as the basis of predictions, social identities have to feel stable just as personal identities do (Swann, 1990). This preference for stability of social identity content was demonstrated by Chen, Chen, and Shaw (2004), following Swann's (1990) self-verification model. Chen and colleagues created social group identities in the lab. They demonstrated that participants preferred to interact with others who viewed their ingroup as they did, even if the social identity of their ingroup was negative.

Although research on social identities typically focuses on a particular social identity, self-concepts are assumed to contain multiple social and personal identities. Individuals can categorize themselves at various levels of ab-

stractness and can define themselves in terms of multiple social identities that connect and intersect in different ways (Burke, 2003; Roccas & Brewer, 2002). Moreover, social identities are not simply self-definitions in terms of social category memberships (e.g., "I am a girl" or "I am a Midwestern democrat"). Social identities also include the traits that come with the categories of gender, social class, political affiliation, and so on (Oyserman, Kemmelmeier, Fryberg, Brosh, & Hart-Johnson, 2003). Thus if being Black or African American is defined in terms of academic engagement then the social identity "I am Black" includes academic engagement so that behaviors such as studying, asking questions after class, or persisting at difficult schoolwork are part of one's self-definition. Generally, traits and characteristics seen as ingroup defining are more likely to be accepted as potential self-definitions as well.

In addition to highlighting the importance of one's group memberships in self-concept, social identity theory clarifies the contingent nature of self-concept content (Hogg, 2003). That is, depending on the situation, the self can be seen as separate, unique, and distinct from others, as part of a single social identity, as part of multiple, overlapping, or conflicting identities, or as part of a merged and connected set of identities. Group memberships provide not only a sense of what or who one is but also a way of locating oneself in relation to in- and outgroups (Hogg, 2003). Who one is includes the totality of self-definitions one has, including traits one has or may acquire because they are ingroup defining and traits one does not have or cannot acquire because they are outgroup defining. The totality of one's group memberships creates a distinct self (Hogg, 2003).

Some traits and characteristics are part of multiple ingroup definitions; for example, doing well in school is part of ingroup definition of a number of racial–ethnic groups. However, sometimes social groups vie to claim the same positive domains as defining their ingroup. A number of social identity theorists have noted that given unequal social power, majority groups are likely to be more successful in claiming valued domains as ingroup defining than minority groups so that minorities must develop alternative means of maintaining positive ingroup identity (for reviews, see Blanton, Christie, & Dye 2002; Branscombe & Ellemers, 1998). This between-group tension is likely to be particularly intense in areas that are critical for social advancement and social power. A group may be willing to concede some domains as defining an outgroup more than an ingroup (e.g., athleticism or prowess in sports, rhythm or talent in music) but not others (e.g., intelligence or academic performance). Thus, by highlighting the between-group tension or struggle to define one's group in terms that are both positive and sufficiently distinct from other groups, social identity theory clarifies that a socially constructed self is constructed from those ways of being that have been claimed by one's ingroup.

To maintain positive identities and avoid incorporating negative outgroup appraisals or stereotypes into concept of social identity, minority groups can reframe their ingroup identity in a number of ways. They can devalue the domains that define the outgroup, discount negative feedback about performance in outgroup defining domains, or take a more blanket approach and use ingroup rather than outgroup both for definitions of success and for feedback about progress toward self-relevant goals (e.g., Branscombe, Schmitt, & Harvey, 1999; Crocker & Major, 1989; Crocker, Voelkl, Testa, & Major, 1991; Osborne, 1995; Tajfel & Turner, 1986). All these ways of defining one's social groups are likely to have consequences for how one sees oneself and the goals one is likely to pursue, though social identity theory does not itself provide a process model of how content of social identity is likely to influence ongoing self-regulation.

Social Identity from the Perspective of Cultural Psychology

Compared with social identity-based descriptions, with some exceptions, cultural psychological models of the self have paid less attention to the traits contained within an interdependent conceptualization of the self. Rather, cultural psychology has emphasized the impact of cultural milieu on propensity to define the self in terms of the private and personal as compared to the social and collective (Markus & Kitayama, 1991). While social identity theories do not explicitly connect content of identity to cultural milieu, social identity reasoning is clearly relevant to differences in cultural milieu. Because cultures are shared systems of meaning that are intergenerational, they are likely to shape which groups are meaningful and how they are characterized, and in that sense culture is basic to social identities. Cultures provide standards of meaning so that members of a culture share not only a common language and location but also shared beliefs, perceptions, evaluations, and ways of acting (Oyserman & Lee, in press).

Although there are likely to be multiple dimensions on which cultures differ that are relevant to content of social identity and the process of self-regulation, the two organizing dimensions that have received the most attention are individualism and collectivism. Individualism has been defined as a focus on rights above duties, concern for oneself and immediate family, emphasis on personal autonomy and self-fulfillment, and basing one's identity on one's personal accomplishments (Hofstede, 1980). It is a worldview that centralizes the personal—personal goals, personal uniqueness, and personal control—and peripheralizes the social (e.g., Kâgitçibasi, 1994; Oyserman, Kemmelmeier, & Coon, 2002; Triandis, 1995). Individualism is contrasted with collectivism; whereas individualism focuses on the personal, collectivism focuses on groups and relations that bind and mutually obligate individuals.

According to a cultural perspective on the self, cultures can be divided into those that highlight values of individualism and those that highlight values of collectivism in socialization practices (see Oyserman, Coon, & Kemmelmeier, 2002, for a review). Societies that centralize individualism in socialization practices are more likely to promote parenting and other social institutional practices that bolster an individual or personal identity-

focused form of self-concept in which the self is seen as an independent, separate, and causal agent. Societies that centralize collectivism in socialization practices are more likely promote parenting and other practices that bolster a related, social or collective identity-focused form of self-concept in which the self is seen as part of social groups and having meaning and agency through group memberships (Kâgitçibasi, 2002).

Like social identity theorists, cultural psychologists have assumed that the self can be defined in terms of both social and personal identities. However, cultural psychologists have focused on between-culture differences in the likelihood that the self is social or personal in focus (Markus & Kitayama, 1991). Individualism implies a personal self-focus, that feeling good about oneself, being unique or distinctive (Oyserman & Markus, 1993; Triandis, 1995), and defining the self with abstract traits as opposed to social or situational descriptors are central to self-definition (Fiske, Kitayama, Markus, & Nisbett, 1998). Conversely, collectivism implies a group or collective self-focus, that group membership is a central aspect of identity (Hofstede, 1980) and that the valued personal traits contained in self-concept reflect the goals of collectivism, such as sacrifice for the common good and maintaining harmonious relationships with close others (Markus & Kitayama, 1991; Oyserman, 1993; Triandis, 1995).

As articulated by Markus and Kitayama (1991), this distinction has been described as differing models of the self, the self as "independent" and as "interdependent." Whereas the initial independent–interdependent self model drew from examples of differences between Japanese and Americans and has been criticized by researchers unable to empirically validate this particular cross-national difference in content of self-concept (e.g., Matsumoto, 1999), follow-up use of the model has moved well beyond a particular cross-national comparison. Indeed, the idea that self-concepts differ in foci parallels earlier work on gender differences in self-concept that documented differences in tendency to define the self in terms of agency versus connection (for reviews, see Cross & Madson, 1997; Markus & Oyserman, 1989). Whether due to gender- or culture-based socialization, it seems reasonable to assume between-person differences in the likelihood that one will conceive of oneself as an agentic entity that is separate or independent of others or as a part embedded within a relational web (Markus & Kitayama, 1991; Oyserman & Markus, 1993).

Although sometimes simplified as if to describe the self as either independent or interdependent, cultural psychologists do explicitly describe the self as defined in both ways (e.g., Trafimow, Triandis, & Goto, 1991; Triandis, 1989). That is, across all societies, both independent and interdependent elements are incorporated into self-concepts (Oyserman, Kemmelmeier, & Coon, 2002). What is likely to differ cross-culturally and across societies is the number of everyday moment-to-moment contexts that cue or turn on one or another aspect of self-concept. An emerging literature shows that when reminded to think of oneself as a social entity (part of a "we" group) or as a separate entity (a singular

"I") individuals do respond differently to social cues (Haberstroh, Oyserman, Schwarz, Kühnen, & Ji, 2002) and process and remember information differently (Kühnen & Oyserman, 2002; Oyserman & Lee, in press).

It seems reasonable that social identity and cultural perspectives be integrated to provide an identity-based sociocultural model of motivation, in which content of self-concept differs both chronically (based on differences in cultural milieu) and momentarily (based on momentary salience of social roles or group memberships). Whether due to chronic or momentary focus, when social roles or group memberships are salient, individuals should define themselves in terms of these roles and the traits and ways of being relevant to these. When individual difference is made salient, individuals should define themselves in terms of their traits and individual preferences. In both cases, motivation to act or refrain from acting will be identity based.

Unfortunately, the empirical base for such integration is as yet limited. An earlier review found that most cultural and cross-cultural research on content of self-concept was correlational and used as the dependent variable content coding from Kuhn and McPartland's (1954) Twenty Statements Task (see Oyserman, Coon, & Kemmelmeier, 2002, for a review). Although, as outlined below, quite a few priming studies were published in the past few years, these studies still rely on content coding of the Twenty Statements Task as the single dependent variable (see Oyserman & Lee, in press, for a review). While taken together, results of these studies do support the notion that content of salient or online self-concept shifts when one is primed to take into account social groups as opposed to individual differences, the Twenty Statements methodology does not lend itself to understanding the traits, characteristics, and future self-goals nested within social identities. To understand why this is so, the specific instructions and coding methods are outlined below. As will be seen, though at first seeming to be a reasonable method for highlighting content difference, the method does not live up to its promise.

Task instructions follow the form "In the twenty blanks below please make twenty different statements in response to the simple question (addressed to yourself), 'Who am I?' Answer as if you are giving the answers to yourself, not to somebody else. Write your answers in the order they occur to you. Don't worry about logic or importance. Go along fairly fast." These instructions are followed by 20 blank lines beginning with the words "I am." In some versions (e.g., Cousins, 1989), respondents are then told to go back over their responses and mark the five responses that are most important to them. Across all the studies using the Twenty Statements Task to study content of self-concept, respondents were primed with personal versus social self-focus, then generated responses to an "I am . . . " stem.

In a classic study, Cousins (1989) found that whereas Americans described themselves in terms of traits more than Japanese students when using the standard (decontextualized) version of the Twenty Statements Task, this tendency was reversed when the task was modified to ask for self-descriptions in context. Once contextualized,

that is, when instructions were to describe oneself "at home," "at school," and "with close friends," Japanese were more likely to use trait descriptors than Americans. Cousins shows that, instead of having decontextualized trait-based selves, Japanese participants had a set of contextualized trait-based selves. The idea that the traits and attributes that are part of a social identity (e.g., student) should become self-defining when that social identity is made salient is basic to social identity theory. However, this way of thinking about content of self-concept is not standard within a cultural psychology framework. Thus, this study was an important first step in making the link between cultural and social identity frameworks; cultural and social identity perspectives converge in predicting first that Japanese would be more likely to describe themselves in terms of the social identities (in this case as child, student, friend) and that once a social identity is made salient, relevant traits and attributes would come to mind. This interesting methodology has not been replicated by other cultural psychology researchers, with the exception of a conceptual replication is by Kanagawa, Cross, and Markus (2001), which unfortunately had inconclusive results.

Further cultural psychology research on content of self-concept used three different priming tasks. Trafimow and colleagues (1991) developed two of these priming tasks. In one priming task, participants read about a Sumerian warrior who needs to choose a general. The criteria he is described as using are the basis of the prime. In one condition, he makes the choice based on the general's traits (meant to be an individual self-prime). In the other condition, he makes the choice based on the general's ingroup connections to family and tribe (meant to be a collective self-prime). Participants read the paragraph and then respond to dependent variables. In the second priming task developed by Trafimow and colleagues, rather than read a passage, participants are asked to take a moment to think about either their similarities to or their differences from family and friends. The authors argue that thinking about difference should prime the private or personal self, while thinking about similarities should prime the collective, relational, or interdependent self. The third priming task involves a more subtle manipulation. Brewer and Gardner (1996) asked participants to read a paragraph and circle either first-person plural ("we") pronouns in a paragraph or third-person ("they") pronouns. This task was revised by Gardner, Gabriel, and Lee (1999). They created a paragraph-length story about a trip to the city in which pronouns to be circled were either first person singular pronouns (I, me, mine) or first person plural pronouns (we, our, us). Circling first-person singular pronouns was the independent self-prime whereas circling first-person plural pronouns was the interdependent self-prime.

Across studies using these primes to assess content of self-concept, the Twenty Statements Task was almost always used as the dependent variable. To determine whether an interdependent (social) self was evoked, content coding focused on collective or group-level self-descriptors (e.g., I am a woman or I am a student). The count or proportion of collective self-descriptors was pre-

dicted to increase following interdependent self-prime. To determine whether an independent (personal) self was evoked, content coding focused on personal trait self-descriptors (e.g., I am determined or I am smart). The count or proportion of trait-focused self-descriptors was predicted to increase following independent self-prime.

Across all studies, the modal response focused on traits (e.g., "I am smart"), coded as "private" self, with many fewer responses focused on group memberships (e.g., "I am a boy" or "I am a student"), coded as "collective" self. Less commonly coded for (or reported on) are responses focused on interpersonal aspects of self-concept (e.g., "I am shy"), those traits that explicitly require the presence of others. Some authors have suggested that these results imply that personal aspects of identity are always more motivationally powerful than social aspects of identity (Gaertner, Sedikides, Vevea, & Iuzzini, 2002). However, it is not possible to infer from the large preponderance of trait descriptors that content of self-concept is mostly focused on personal identities because, as demonstrated in Cousins's (1989) study and as argued by social identity theory, having primed social identities should increase the salience of traits relevant to the social identity. Because researchers in the described studies that follow focus on significant increase in collective identities but do not report on any attempt to code for social identity-relevant traits, it is not possible to tell if the traits brought to mind are appropriately thought of as part of personal or social identity aspects of self-concept.

Thus, six studies using the Sumerian warrior task show an increase in collective self-descriptors in the Twenty Statements Task when the collective versus the individual prime was used (between-subjects design). Participants were U.S. college students (Trafimow et al., 1991), Native American adults (Trafimow & Smith, 1998), European American college students (Ybarra & Trafimow, 1998), a nonspecified U.S. sample (Mandel, 2003), Hong Kong Chinese high school students and adults (Trafimow, Silverman, Fan, & Law, 1997), and European American and Asian American students (Gardner, Gabriel, & Dean, 2004).

In the latter study, "we" priming resulted in a significantly larger increase in responses focused on collective identity for Asian Americans compared to European Americans (Gardner et al., 2004). This may reflect the chronic propensity of people socialized in Asian cultures to conceive of the self in terms of social identities. Language used in each case was English, which may be an important contextual feature of the prime; Trafimow and colleagues (1997) found no effect of priming when materials were presented in Chinese. When responding in Chinese, about three-quarters of the responses were categorized as private and about one-quarter of the responses were categorized as social or collective. Of course, it is possible that responses in Chinese showed more subtle effects—traits may have been those relevant to the social identity brought to mind by the prime— though information about the identity the warrior prime brought to mind was not obtained.

Similarly, three studies using the similarities to/differences from family and friends prime show increase in collective self descriptors in the Twenty Statements Task when the collective versus the individual prime was used (between-subjects design). Participants are European American and Chinese college students in the United States (Trafimow et al., 1991), Hong Kong Chinese high school and college students (Trafimow et al., 1997), and female U.S. college students (Vohs & Heatherton, 2001). Language used in each case was English. As with the Sumerian prime, when Chinese was used in the Hong Kong-based sample, no effects of priming were found. However, effects do not appear to be limited to English. A fourth study conducted in German found significant results using as the dependent variable three collective self-items from the Singelis (1994) self-construal scale. Collective responses increased among German college students in the similarities to friends and family condition (Kühnen, Hannover, & Schubert, 2001). Because this study used a different dependent variable, it is not entirely clear whether effects would have been found with the Twenty Statements Task, again raising the question of whether current methods are adequate to detect effects.

The final set of studies shows shift in self-concept content following the pronoun-circling prime. In studies with European American participants the "we" prime increased interpersonal and collective self-descriptions compared with a "they" prime (Brewer & Gardner, 1996) as well as compared with the "I" prime (Gardner et al., 1999, 2004; Gardner, Gabriel, & Hochschild, 2002). Gardner and colleagues (2004) also showed effects with Asian American participants, effects that were significantly larger than in their European American sample; effects were of the same size whether the Sumerian warrior or the pronoun-circling tasks were used.

Thus, across prime type, "we" priming did shift content of self-concept toward social identities, suggesting that momentary contextual effects influence working or online self-concept. Thus, evidence supports a sociocultural approach to when social identity is cued or made salient. Evidence is limited by the fact that almost all the research involves American samples, and perhaps more importantly, by the method of combining all trait responses into a "private" self code. Studies were not set up to examine the content of primed social identities. It is possible that the traits described are those relevant to the primed social identity. This is a main feature of the social identity approach and is in fact congruent with a cultural psychology approach that would posit that the traits chosen to define the self are those that are culturally valued. Individuals who endorse individualistic cultural values are indeed more likely to describe themselves in terms of individualistically oriented traits, while individuals who endorse collectivistic values are more likely to describe themselves in terms of collectivistically oriented traits (Oyserman, 1993). Given that both perspectives would posit that individuals are motivated to take on the traits and characteristics valued by ingroups, the lack of research that could examine this assumption is particularly puzzling.

While cultural and social identity approaches both suggest that content of self-concept is cued by relevant contextual cues, neither approach provides explicit models of the process by which self-concept influences self-regulation. This process-level framing of self-concept has been articulated within a self-schema approach. As reviewed in the next section, a self-schema approach is highly compatible with a social identity framework and could be integrated within a cultural perspective, together creating an identity-based motivation process model.

IDENTITY-BASED MOTIVATION: A SELF-SCHEMA APPROACH

Social identity approaches assume a hierarchical organization to self-concept, suggesting that content of self-concept is organized within a series of hierarchically organized identities that may or may not feel connected with one another (Roccas & Brewer, 2002). The notion of hierarchical structure is not emphasized in the self-schema model; rather, this approach focuses attention on self-concept process and function (Markus, 1977; Markus & Wurf, 1987). Within a self-schema approach, self-concept is assumed to be made up of cognitive schemas about the self that mediate perception and regulate affect, motivation, and behavior, lending meaning and organization to thoughts, feelings, and actions and motivating action by providing incentives, standards, plans, strategies, and scripts for behavior (Oyserman & Markus, 1993). Rather than focus on hierarchical organization, this approach emphasizes temporal flow. Self-concept content includes an articulation of how one was in the past, is in the present, and might possibly be in the future. These temporal selves include both content relevant to social categories and social roles and content relevant to individual attributes.

The schema approach has already been adapted to a cultural frame (Oyserman & Markus, 1993). How the self is described, which content is included in self-concept, and the incentives, strategies, and scripts adopted to motivation and regulate the self are all likely to be culturally framed (Oyserman & Markus, 1993). That is, individuals are not schematic for all of the characteristics, traits, skills, and abilities that are true or observable about them (Markus, 1977). Instead, self-schemas reflect meaningful domains, those domains that are valued or marked as important in one's social context (Oyserman & Markus, 1993). Thus not all self-relevant content and knowledge becomes integrated into a self-schema, some images or conceptions about the self are tentative, fleeting, peripheral, or not well integrated, while others are more highly elaborated and more chronically accessible. It is these latter selves that function as enduring meaning-making interpretive structures, fostering coherence and forming the core of self-concept.

These salient identities (Burke, 2003; Stryker & Burke, 2000) or self-schemas (Markus, 1977) are packages of self-knowledge reflecting what an individual cares and thinks about and spends time and energy on, dimensions

along which individuals hold clear and distinct perceptions about themselves. They are domain-specific organized cognitive structures that provide generalizations about one's past and present and claims about one's possible future characteristics, actions, and skills (Montepare & Clements, 2001; Oyserman & Markus, 1993). As cognitive structures, self-schemas direct attention to self-relevant information and so influence what is perceived in the environment (Markus & Sentis, 1982). They direct memory and so influence what is remembered and what cues are recalled (Markus, 1977).

Information is assimilated into existing schemas where possible. Individuals process schema-relevant information more quickly and more efficiently than schema-irrelevant information (Markus, Smith, & Moreland, 1985) and are likely to misremember information in ways that reflect their own schemas (Markus, Crane, Bernstein, & Siladi, 1982; Markus & Wurf, 1987). Information irrelevant to self-schemas is likely to be disregarded (Markus, Hamill, & Sentis, 1987); ambiguous information is likely to be framed in ways relevant to self-schemas (Catrambone & Markus, 1987). When a domain becomes self-schematic, it becomes important to maintain a particular view of the self within this domain. Individuals are more likely to challenge, disbelieve, or try to refute negative or disconfirming schema-relevant rather than schema-irrelevant feedback (Markus, 1977). Negative or disconfirming feedback that is schema irrelevant is unlikely to result in mobilization of effort and resources to combat it.

Self-schema research has typically focused on self-concept at the level of personal or individual identities and has been criticized for being explicitly individually focused (Onorato & Turner, 2002). However, while the initial research focused on an individual trait (e.g., defining the self as "independent"), the self-schema conceptualization itself is not limited to personal identities (Oyserman et al., 2003). A self-schema approach has been used to examine the impact of self-concept defined in terms of social category memberships, such as being a man or a woman, being heavyweight and being a member of one's age category. This research shows that not all men and women have gender self-schemas (Markus et al., 1982), not all heavyweight people have "fat" self-schemas (Markus et al., 1987), and not everyone is schematic for his or her age (Montepare & Clements, 2001). Across each of these domains, those who are schematic are more likely to organize information in terms of these schemas and are better able than aschematic individuals to defend the self from negative schema-relevant feedback. Because they are likely to be chronically salient, social identity self-schemas, like personal identity self-schemas, are likely to influence ongoing meaning making, motivation, and persistence.

Following this logic, not all social roles and social categories to which one belongs will become schemas. For example, with regard to minority race and ethnicity a number of authors have argued that one's membership in a minority racial or ethnic group are likely to shift from being simply facts about the self to being important social identities only if life experiences make them central (e.g.,

Cross & Fhagen-Smith, 1996). When one's minority race–ethnicity is salient and contextually valued or marked, it is likely that self-schemas focused on this race or ethnicity will develop. Even when race–ethnicity is culturally marked, not everyone will have a racial or ethnic self-schema, just as not everyone has a gender self-schema (Oyserman, Brickman, & Rhodes, in press; Oyserman et al., 2003). Those who are aschematic will make sense of who they are without spontaneously thinking about race–ethnicity. Those who have a racial–ethnic schema are likely to make sense of themselves and their possibilities through the lens of this schema when it is made momentarily or chronically salient by social contexts. Like other self-schemas, racial–ethnic self-schemas (RES) are stable processing structures that guide the perception, encoding, and retrieval of information relevant to one's racial–ethnic group membership and the connection between membership in this ingroup and membership in larger society (Oyserman et al., 2003).

Following the self-schema model, race–ethnicity aschematic individuals will be more vulnerable to negative feedback based on race–ethnicity, including stereotypes or situational factors emphasizing their otherness, because they lack a cognitive structure to automatically process and fend off the negative self-relevant implications of racially tinged information. Those who define themselves in terms of their racial–ethnic ingroup are RES schematic and will make sense of their circumstance and focus their self-regulatory effort in terms of the content of the schema. Given that many groups would prefer to self-define in terms of generally valued traits and goals such as academic success, to the extent that RES does not explicitly contain links to these goals, schema-based processing carries the risk of disengaging effort from these goals (Oyserman et al., 1995, 2003; Rhodes, Oyserman, & Brickman, 2006).

Oyserman and her colleagues (2003) found evidence that racial–ethnic self-schemas function like other self-schemas in that they focused self-regulatory effort; when primed, RES that contained focus on school as ingroup defining improved academic persistence. When students' RES "bridged" ingroup and larger society by explicitly taking both into account, students were more academically engaged and fared better in school than when their RES focused on the ingroup only or they were RES aschematic. Controlling for prior grades, over the course of the school year, grades of low-income African American and Latino middle school students did not exhibit decline when their racial–ethnic identity schema contained both ingroup and larger society but did decline when they were aschematic for race–ethnicity and when their racial–ethnic identity schema was focused only on the ingroup.

Two subsequent studies of low-income middle and high school–age African American and Latino students also showed that content of RES influenced self-regulation (Rhodes et al., 2006). An experimental manipulation demonstrates the causal process: When content of RES was experimentally primed by having students write about what it means to be a member of their racial–ethnic group either before or after doing a math task, ef-

fects of bridging RES were found. That is, when RES was primed and content of RES focused on both positive connection to ingroup and bridge to larger society, math persistence improved (Oyserman et al., 2003). Thinking about motivation as identity-based clarifies the underlying process. When ingroup identity is contextually cued, individuals are motivated to engage in ingroup-relevant behaviors. If the ingroup is seen as linked with larger society, then larger societal goals like school attainment are cued. The impact of cuing larger societal goals and strategies should be positive—indeed increased persistance was found across various groups, including American Indians, African Americans, and Arab Israelis.

The notion that content of racial–ethnic identity self-schemas influences important behaviors was further explored in a series of studies examining the extent that health promotion behaviors are or are not included in social identity (Oyserman, Fryberg, & Yoder, 2006). Oyserman and her colleagues (2006) posed two questions:

1. Does racial–ethnic social identity include health promotion (such as exercising regularly) or unhealthy lifestyle behaviors (such as smoking or eating candy)?
2. How does content of racial–ethnic social identity influence cognitions and perceptions about health?

Minority college students rated health behaviors as White and middle-class things to do and were more likely to rate unhealthy than healthy behaviors as racial–ethnic ingroup things to do.

Not only did unhealthy behavior appear to be part of RES, but this content, when made salient, seems to have motivational consequences. When primed with race–ethnicity and social class, college and middle school students who are African American, Mexican American, and American Indian are significantly more likely to endorse a fatalistic perspective about health than in the control condition when social class and race–ethnicity are not made salient. A follow-up study with middle school students showed that priming RES also makes health information less cognitively accessible for low-income minority middle school students as compared to control condition. Follow-up studies with American Indian college students and reservation adults demonstrate that the undermining affects of making RES salient occur only when unhealthy behavior is incorporated into RES. Thus, an integration of social identity and self-schema approaches is fruitful in beginning to understand how social identities influence motivation and self-regulation.

Both sociocultural identity and self-schema approaches assume that content of self-concept is socially derived and demonstrate that when social contexts bring social groups or relational ways to thinking about the self to mind, social identities and social self-schemas are primed. However, neither social identity nor self-schema approaches provide an explicit process model of how social contextual information is incorporated into self-concept. To begin to build an identity-based motivation process model that articulates how social contextual information is incorporated into self-concept, it is necessary to turn to social cognition approaches.

Identity-Based Motivation: Integrating a Social Self-Schema Approach with Social Cognition Frameworks

A social cognition framework is a useful starting point in making predictions about the influence of contextual factors on salient content of self-concept and the influence of salient self-concepts on self-regulation and behaviors. In particular, the inclusion–exclusion (Schwarz & Bless, 1992, in press) or assimilation–contrast (Blanton, 2001; Schwarz, Bless, Wänke, & Winkielman, 2003) model provides insight into when social information is likely to be assimilated into one's judgment of who one is and what one might become and when this social information is likely to be used as a standard, excluded from self-concept, such that one's own successes or failures are judged relative to the standard. Because we live in a social world, social comparisons are ubiquitous. Answering the "who am I" and "how do I fit in" questions necessarily involves others, as role models, as yardsticks, or as parts of how we define ourselves.

The social comparison literature has classically proposed that individuals contrast themselves with others, feeling good when another is doing comparatively worse (a downward social comparison) and bad when another is doing comparatively better (an upward social comparison) (see Blanton, 2001; Collins, 1996, for reviews). Much research has focused on the use of downward social comparisons to improve self-evaluation (Gilbert, Giesler, & Morris, 1995; Pelham & Wachsmuth, 1995; Wills, 1981). These comparisons to a worse-off target provide a pleasing reminder of one's own superiority, especially if one cannot easily generate plausible parallels between the other's fate and one's own (Brewer & Weber, 1994). Downward comparisons are equally effective when the comparison target is an individual or a group; indeed, social identity theorists argue that downward outgroup comparisons contribute positively to social identity and that a primary function of social identities is to provide the basis for favorable self-evaluation (Tajfel, 1981).

Although downward comparisons are clearly effective, this strategy can be risky if downward comparison is seen as gloating (on a personal level) or when done on an intergroup level as blatant prejudice—whether racism, sexism, or classism or other negative group-based comparison. How can social comparisons avoid this particular problem yet still produce the desired positive self-evaluative boost? One possibility is upward comparisons, which can promote positive evaluation to the extent that the other is seen as a role model, or the other's success is viewed as similar enough to one's own to directly provide a boost (Collins, 1996). However, this strategy is risky because it clearly highlights the gaps between the other's positive characteristics and one's own less positive characteristics and it risks suggesting that these differences are unbridgeable, which may result in dampened self-

evaluation (Mussweiler, Rüter, & Epstude, 2004; Taylor & Lobel, 1989).

Another possibility is to avoid social comparison altogether and to simply assimilate the target's positive attributes into one's own self-evaluation, to "bask" in the reflected glory (BIRG) of the other (Cialdini et al., 1976). Assimilating the target to the self feels good without risk of threatening social comparison. Moreover, because assimilating the other's success does not denigrate the other, basking in reflected glory (BIRGing) is likely to be both safer than upward social comparisons and a more socially acceptable way to enhance self-worth than downward social comparisons. To BIRG, it is necessary to create a sense of closeness and to reduce boundaries that would otherwise trigger self-other contrasts (Arnett, German, & Hunt, 2003; Pelham & Wachsmuth, 1995). BIRGing is especially likely to produce gains in positive self-regard when the target's positive attributes are in self-irrelevant domains (Chen et al., 2004; Hirt, Zillman, Erickson, & Kennedy, 1992). Because the domain itself is not central to self-definition, the other's success can simply be included in the self and does not provide an upward comparison standard. By creating a symbolic link between the self and the target, one can feel good when the target succeeds. Thus, nonathletes can BIRG athletes; nonartists can BIRG musicians, artists, and the stylistically cutting edge; and the non-ecologically minded can BIRG those who preserve the natural environment.

In the initial demonstration of this effect, Cialdini and his colleagues (1976) showed that students were more likely to wear school-themed clothing and refer to their university as "we" rather than "they" on weekends in which the college football team won the game. In this way, students symbolically took on the positive attributes of the winning team. Follow-up research has focused on the impact of BIRGing of successful or prestigious individuals, organizations, and groups (e.g., a successful sports team— Bernhardt, Dabbs, Fielden, & Lutter, 1998; Boen, Vanbeselaere, & Feys, 2002; a high-ranked university—Bhattacharya, Rao, & Glynn, 1995; a winner in political elections—Boen et al., 2002; or a successful marketer—Arnett et al., 2003).

However, successful targets are not always assimilated into self-views. Lockwood and Kunda (1997) provide a useful example. After reading materials about a "superstar" student, participants were asked to rate their own current and possible future success. When rating current success, students used the information about the superstar as a standard, relative to which their own current success looked more modest than without the standard. When rating possible future success, however, students incorporated the superstar as a possible self and rated their own future possibilities more highly after being exposed to the superstar. That is, in the former case, the superstar was a standard against which one's own performance was contrasted, while in the latter case the superstar informed one's judgment about what was possible and so became incorporated or assimilated into the target of judgment—one's future chances. A number of

follow-up studies have asked whether the propensity to assimilate versus contrast information about the other into one's judgments about oneself is influenced by factors other than the whether the judgment is focused on the present versus one's possible future. In particular, researchers have asked whether the tendency to incorporate or assimilate information about another into one's self-judgment is carried by a chronic or primed tendency toward interdependence.

Research from a number of studies suggests that the tendency to assimilate or incorporate social information into one's self-judgment as opposed to using this information as a yardstick to assess one's relative standing is indeed influenced by interdependence. Kemmelmeier and Oyserman (2001) showed that both Palestinian Israeli women and European American women are more likely to assimilate a downward target into their self-judgment than are men. They replicate this work with Palestinian Israeli, German, and Turkish students asked to generate an upward social comparisons, again showing that women are more likely to assimilate their self-judgment to that of a same gender comparison who is performing better than they are, whereas men are more likely to show contrast effects (Kemmelmeier, Oyserman, & Brosh, 2005). Unfortunately, this work does not provide a direct assessment of the prediction that assimilation is driven by tendency to interdependence, relying instead on research documenting that women are chronically higher in interdependence than men (Cross & Madson, 1997).

Fortunately, this issue has also been addressed directly in the experimental literature utilizing the pronoun-circling prime developed by Brewer and Gardner (1996; Gardner et al., 1999). In a series of studies with Dutch college students, Stapel and Koomen (2001) show that "I" priming makes salient contrast with other; "we" prime makes salient assimilation with other. The pattern of assimilation with "we" priming is also shown in studies with German participants using the pronoun-circling task (Kuhnen & Haberstroh, 2004) and by writing down independence (interdependence)-relevant words in scrambled sentences (Kuhnen & Hannover, 2000).

Stapel and Koomen (2001) note that there is a self-serving asymmetric pattern to these contrast effects: Contrast effects are larger when comparing self to low-performing standard or when the other's positive results are in an unimportant domain. However, when instead of using the "I" priming task, participants were primed by unscrambling sentences with the words *compare, distinguish, differ*, and *opposition*, the asymmetry disappeared and respondents contrasted themselves to standards even when this resulted in negative self-definitions (Stapel & Koomen, 2001).

Although not addressed by the authors, this latter finding is important because it suggests that when focused on the self as different from or in opposition to others—as may occur in either intergroup or interindividual contexts—individuals are likely to use other's performance not as a model for one's own possibilities but as a contrasting standard, against which one's own perfor-

mance looks relatively worse. This negative assessment of one's current state and future possibilities may trigger disengagement from the goal, either because one no longer sees the goal as possible or one no longer feels competent to engage in relevant goal pursuit activities (Bandura, 2001; Wrosch, Scheier, & Carver, 2003). Indeed, Gardner and colleagues (2002) find that when primed to think in terms of social category memberships using the Trafimow and colleagues (1991) Sumerian warrior prime (in which participant read about a warrior who chooses a general due to his family and tribal ingroup connections), participants rated their friends as likely to succeed on a self-relevant task, whereas participants primed to think about individuals as having separate traits and characteristics (after reading about a Sumerian warrior who chose a general due to his skills) are significantly less likely to do so.

Although preliminary, taken together these studies suggest first that assimilation of information about another is more likely for individuals chronically (e.g., women) or situationally (e.g., after priming tasks) interdependently oriented and less likely for individuals chronically (e.g., men) or situationally (e.g., after priming tasks) independently oriented. Conversely, these studies suggest that using information about another as a contrasting standard from which to evaluate the self is more likely for individuals chronically or situationally independently oriented and less likely for individuals chronically or situationally interdependently oriented. When cued, ingroup belonging should evoke both motivation to be like the ingroup and information about ingroup characteristics. Ingroup belonging can be cued chronically or by specific situational information. Cues can be subtle and the process should proceed automatically once cued.

As a more general frame, the inclusion–exclusion model (Schwarz & Bless, in press) proposes that social information is included in the self-judgment unless the information is judged incompatible with the self. Social information that cannot be included in the self is used as a standard of comparison. Social information is more likely to be judged incompatible with self-concept when it is extreme relative to current self-content and when the social information is explicitly or implicitly presented as separate from the self. The inclusion–exclusion model articulates the circumstances in which information about another will be included in the self, so that the other's successes and failures become part of oneself, and when this information will be excluded from the self, so that the other's successes and failures will become a standard of comparison.

Because social information that is irrelevant to the judgment task is unlikely to be used, not all social information will be included in or excluded from self-judgments. Some social information will be ignored. Relevance is subjective. Thus, for some, knowledge of American students' low ranking in international comparisons of math and science achievement creates a sense of urgency because international comparisons are relevant. Other countries are a standard against which "we" are doing badly. For others, the information is simply not relevant—other countries are not "us." Of course, what constitutes a relevant comparison is likely to be context dependent. A social cognition perspective makes clear that what social information is deemed relevant and how it is used is highly dependent on what makes sense in context. Meaning is made in the moment, it feels sensible, obvious, and natural in the moment, but slight shifts in context will shift meaning.

Early Formulation of the Self as a Motivational System

Once an image has been deemed relevant and either included in self-definition or formulated as a standard against which one should compare oneself, how does the self proceed? James (1890/1927) developed what can be considered a precursor of current self-motivation theories. He conceptualized the self as the metacognitive experience of being a self on the one hand and as containing cognitive and affective content (self-knowledge and self-feelings) on the other. He proposed that all things being equal, individuals would desire to be and become all possible selves simultaneously ("a Greek scholar, a bon vivant"), expanding to incorporate ever more self-goals to strive toward.

He argued that this tendency to incorporate ever more selves as possible future selves is limited or reigned in by a number of factors. First, some self-projects simply cannot be pursued simultaneously because the actions needed to pursue them are incompatible. The bon vivant wants to stay out late with friends; the scholar wants to return to his books—resulting in the need to choose which of two competing goals to focus energies on. Second, some self-projects turn out to be unattainable either because of failure of strategies to attain the future self or because of lack of ability. After years of lessons fail to create the desired child prodigy self, and yet more years fail to create even a gifted pianist self, at some point the self-goal will need either to be abandoned (e.g., "I played the piano when I was younger") or reshaped (e.g., "I play the piano just for fun").

When future imagined selves cannot be worked on, they are unlikely to engage much attention or affective response and so may wither away. Holding onto a blocked or failed self-goal or possible self has negative consequences for self-valuation. Continued engagement in blocked or failed self-projects is limited by one's ability to tolerate the negative feelings that failures to attain possible selves entail. Rather than continue lessons imagining that one will become a gifted pianist, one may over time revise the goal to instead imagine an "enjoying music" possible self or to abandon the goal altogether.

Within James's model there is an implied innate desire to self-regulate and to attain ever more self-goals. This desire is limited only by the need to succeed in some proportion of one's efforts. In James's model, self-regulation is associated with self-esteem, operationalized as proportionate success, the ratio of selves one is attempting to become to selves one is succeeding in attaining.

Thus, James's framework articulates a model of self-regulation that is focused on incorporating all the selves a person can imagine becoming. The system is assumed

to have finite energy, so that some self-regulation tasks are incompatible with each other. Self-regulation is also assumed to have emotional consequences; it feels good to succeed at self-regulation and it feels bad to fail. These components are present in current social cognition frameworks of self-regulation as well.

WHAT IS SELF-REGULATION?

While self-regulation is a universal capacity that develops along with other cognitive and socioemotional capacities, it is at the same time a deeply personal and self-defining capacity. Not only is motivation identity based, but success at pursuing a goal feels good because it reinforces the identity in which it is based. Self-regulation is the self-in-action—successful self-regulation feels good; failed self-regulation feels bad. Without the capacity for self-regulation, goal attainment would be impossible. Self-regulation entails the channeling of energy, effort, and motivation toward a goal, the strategies relevant to goal attainment, and the goal itself. Thus when racial identity is cued, one is primed to pursue relevant goals. If goal pursuit is successful, it feels good in part because successful goal pursuit affirms membership in the social identity group.

Self-regulation or self-control is the coordination of neural, cognitive, affective, and behavioral processes to moderate reactivity, excitability, and arousal (Rothbart & Rueda, 2005). Self-regulation requires both behavioral inhibition and behavioral activation (Avila, 2001). Self-regulation allows planned, sustained, and sequenced action in service of desired end states to occur (Mischel & Ayduk, 2004). It involves controlling, channeling, or mastering the self to produce sought after results whether these results are attaining a better mood, more satisfactory grades, being liked, fitting in, or gaining power (Brandstätter, Lengfelder, & Gollwitzer, 2001).

Self-regulatory capacity can be described as a motivational resource that can be turned on to pursue one's goals. It can also be described as inhibition of a dominant response—sleeping in, hanging out, saying whatever comes to mind, eating snacks, and replacing the dominant response with another response—getting up at the sound of the early alarm, doing homework, maintaining civility during an unpleasant social event, eating healthy. Thinking of self-regulation as inhibition implies that we self-regulate because we have to, not because we want to (Baumeister, 2002; Baumeister & Heatherton, 1996).

Yet competent self-control is rewarding (White, 1959, 1960). Because it is essential for goal striving, self-regulation is a necessary component of happiness; to self-actualize (Maslow, 1970), to attain a state of "flow" when one is positively focused on fulfilling life tasks (Csikszentmihalyi, 1996), one must be able to focus attention on one's goals. Because it is critical for goal attainment, self-regulation is basic to human happiness, self-worth, and social regard (Bandura, 2001). Efforts to self-regulate are not necessarily effective; one may or may not succeed in movement toward one's goals and self-regulatory efforts may backfire and make things worse

(Thayer, 2001). Individuals differ in their propensity for and success with self-regulation (Baumeister & Vohs, 2003; Bogg & Roberts, 2004).

Self-regulation may be in service of finding out what one can do, convincing others of one's worth, changing or improving one's self, obtaining resources for one's self, or fitting into one's social context. Self-regulation always involves focusing attention and resources on responses relevant to a focal goal to the relative neglect of other goals. At any particular moment in time, focusing attention on one self-relevant goal necessarily means reduced attention to other self-relevant goals. Focusing one's attention on one goal (e.g., the goal of completing homework) means not focusing on other self-relevant goals such as being athletic, popular, or a cooperative group member.

Self-Regulation Is a Socially Scaffolded Developmental Process

Whatever capacity for self-regulation one has will be brought to bear when identity is turned on. There is a normal developmental increase in self-regulatory capacity, a shift toward increased ability to control reactions to stress and to maintain focused attention and increased ability to interpret one's own and other's mental states in order to successfully predict the likely responses to one's self-regulatory efforts. Early effortful control involves ability to focus and shift attention while maintaining perceptual sensitivity, inhibitory control, and low-intensity pleasure (Rothbart & Rueda, 2005). Although infants differ (Bronson, 2000), there is a normative developmental process of improved control over reactivity, excitability, and arousal (Rothbart & Rueda, 2005). Ability to self-regulate is related to early response to novel stimuli, termed "reactivity," "excitability," or "arousability," but also develops through maturation and experience with the social world. Early caregiving involves attunement to individual differences in reactivity and setting up appropriate experiences that scaffold infants' efforts at self-regulation, providing infants with a sense that the context can be controlled (Brooks-Gunn & Markman, 2005; Serbin & Karp, 2004).

Such maternal scaffolding is predictive of successful self-regulation at age 16 months (Conner & Cross, 2003). Effortful control, as assessed by gaze, is observable at 9 months (Bronson, 2000) and predicts effortful control in toddlers at age 18 months (Bronson, 2000). Effortful control that is discernible in toddlers (18 months of age) (Rodriguez, Ayduk, & Aber, 2005; Rodriguez, Mischel, & Shoda, 1989), becomes stable across lab tasks by age 2.5 (Rothbart & Rueda, 2005). Early (preschool) ability to delay gratification predicts adolescent academic and social skill (Ayduk et al., 2000).

At later ages, scaffolding by parents and other adults entails focusing children's attention on effort (Dweck & London, 2004). The actions that allow for self-regulation—not eating that extra bowl of potato chips, not having that third brownie—may not sound hedonically satisfying yet developmental research suggests

that self-regulation is intrinsically pleasurable (Bronson, 2000). Feeling that one is controlling contingencies is likely to produce a positive affective response whether the feeling is that one is making positive things more likely to happen (self-will) or that one is making negative things less likely to happen (self-control).

As children learn about and experience the world, they begin to develop theories about contingencies. The capacity to self-regulate develops from this early base of "if I–then" relationships—that kicks and hand thrusts produce movement of a rattle or mobile, that crying engages soothing caregivers' attentions. Because self-regulation is so essential to humanness, caregiving is likely to universally foster this emerging self-regulatory capacity. Early self-regulation is scaffolded by caregivers who set up environments to facilitate it. Children experience "if I–then" contingencies within social contexts that set up which goals are worth pursuing and what strategies are worth using. These "if I–then" contingencies alone do not direct motivation but rather are cued a part of identity-based motivation. Thus, when identity as "girl" is cued, motivation to act like a girl is cued, if girls behave well in class, pay attention to the teacher, and take notes, these identity-based behavioral beliefs will be translated via a series of "if I–then" contingencies to behavioral sequences to become more like a "girl."

Self-Goals and Self-Regulation

Self-goals are temporally proximal or distal images of oneself in the future. They can be images of the selves one ideally wants to become or feels one ought to become (Higgins, 1996), the possible selves one expects to become, hopes to become, or is afraid one may become but wishes to avoid becoming (e.g., Markus & Nurius, 1986; Oyserman et al., 2004). The gap between one's current self and these future selves is assumed to motivate efforts to reduce the gap or discrepancy between current selves and positive future selves and increase or enlarge the gap between oneself and negative possible selves (Carver, Sutton, & Scheier, 2000).

Self-regulation in pursuit of a self-relevant goal may fail in spite of ongoing investment of effort due to utilization of inappropriate, ineffective, inefficient, or even iatrogenic strategies. What we do to try to attain our goals can produce much heat but little light or even make things worse (like dieting strategies that involve such rigid monitoring that they eventually cannot be kept up, resulting in eventual weight gain instead of loss). Self-regulation may fail not because the outcome is not valued but because the effort required to attain the goal is underestimated or because the strategies brought to bear are not effective.

Even when the outcome is valued and strategies are effective, self-regulation may fail if attainment is directed to another goal. Clearly, not all goals can be pursued with equal vigor. Individuals are likely to have multiple goals that might draw their attention and resources. Goals are likely to differ along a variety of dimensions—some are short term and concrete, others longer term and more abstract. Pursuing some more proximal goals may increase chances of attaining more distal goals—the goal of college is more likely when the goal of good high school grades is pursued, the goal of good high school grades is more likely when the goal of good middle school grades is pursued. Indeed, when the future feels far away, self-regulation may require linkage of distal goals to more proximal ones or goals will not be pursued at all.

But goals are not necessarily compatible with one another, so that focus on one goal may necessarily mean reducing likelihood of attaining another goal. For example, the goal of buying a first home may not be compatible with the goal of being home with one's children if buying a home requires saving money and saving money requires working longer hours. Assuming that individuals have multiple goals, some in the present and some in the future, some congruent and others incongruent, not all goals can be simultaneously pursued. Self-regulation to attain one goal must mean at least temporary abandonment of another goal. A key question then is which of competing goals will be chosen for self-regulatory attention.

As outlined in Figure 18.1, personal and social identities include all these future images. These selves carry motivational characteristics, providing reasons to act and to refrain from acting in any particular situation, specific behaviors to engage in as well as persistence and desistence scripts (how much and how long to keep trying and when to pull back effort). In this way, social identities scaffold one's goals. To the extent that goals trigger action, when these social identities are brought to mind, they should trigger goal-focused behavior. Because even skilled behaviors, once acquired, are grouped together as behavioral sequences or scripts that can be performed without conscious awareness (for reviews, see Kruglanski et al., 2002; Wegner & Bargh, 1998), identities do not necessarily need to be consciously triggered for relevant behaviors to occur. That is, an identity carries with it behavioral tendencies, scripts for action, that are cued when the identity is cued.

While all cybernetic-control or feedback models (e.g., Carver & Scheier, 1990) assume motivation to work toward becoming like positive goals and to avoid becoming like negative or antigoals, these models do not suggest particular linkages between goals and how they are worked toward or the strategies likely to be chosen. This further specification of process depending on self-goal is provided by self-regulatory models that distinguish between behavioral activation and behavioral inhibition or approach and avoid systems (Avila, 2001; Carver, Sutton, & Scheier, 2000; Carver & White, 1994; Elliot & Covington, 2001; Gray, 1990; Higgins, 1996, 1997). The behavioral activation system responds to signals of reward, nonpunishment, and escape from punishment, while the behavioral inhibition system responds to signals of punishment, nonreward, and novelty (Carver et al., 2000; Gray, 1982).

Higgins's (1997, 1998) self-regulatory focus model builds on these distinctions between behavioral activation and behavioral inhibition and further articulates two systems, one focused on attaining successes and avoiding
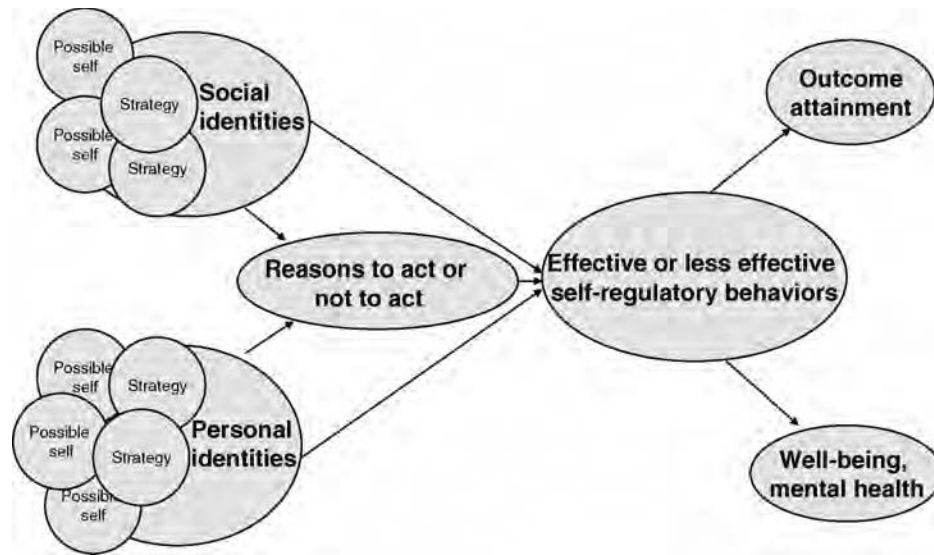
FIGURE 18.1. Identity-based motivation process model.

nonsuccesses (promotion focus) and the other focused on attaining nonfailure and avoiding failures (prevention focus). Different self-goals are likely to be regulated by differing self-regulatory systems; when seeking to attain or avoid failing to attain goals related to becoming like one's ideal selves, self-regulation focuses on promotion, a concern with attaining advancement, growth, and accomplishment. A parallel process occurs when one is seeking to attain or avoid failing to attain goals related to becoming like ought selves. In this case, self-regulation focuses on prevention, a concern with ensuring safety, being responsible, and meeting obligations (Higgins, 1997).

Primed or chronic promotion focus is associated with eagerness, risk taking, sensitivity to the presence or absence of gains, and motivation to ensure acceptance of relevant new behaviors and avoid incorrect rejection of relevant new behaviors. Conversely, primed or chronic prevention focus is associated with minimizing risk, sensitivity to the presence or absence of losses, and motivation to ensure correct rejections and avoid incorrectly accepting irrelevant new behaviors as relevant (Camacho, Higgins, & Luger, 2003; Crowe & Higgins, 1997; Higgins, 1997; Higgins, Idson, Freitas, Spiegel, & Molden, 2003; Liberman, Idson, Camacho, & Higgins, 1999; Liberman, Molden, Idson, & Higgins, 2001). This research suggests that individuals differ in their chronic styles and prefer using self-regulatory strategies that match their self-goals over mismatching ones (Higgins & Silberman, 1998). We prefer to work toward positive goals using eager approach strategies and to prevent negative goals using vigilant, caution-oriented strategies. Termed "value from fit," the underlying notion is that working toward a self-goal feels better when carried out with means that match the ends—eager pursuit of promotion goals feels better than vigilant pursuit of these goals; vigilant pursuit of prevention goals feels better than eager pursuit of these goals.

An identity-based motivation paradigm, a sociocultural identity model that includes self-regulation, is important because movement toward any of the basic self-goals—knowing, improving, bolstering, enhancing, or maintaining one's self all require self-regulation. While these self-goals are often assumed to be aspects of one's personal identity, following the logic of social identity and self-categorization theories (Foddy & Kashima, 2002; Onorato & Turner, 2002; Turner & Oakes, 1989) there is no reason why self-regulation cannot be part of social identity. Social identities include information about the self as a member of one or more social collectives (Abrams, 1994) as well as socially contextualized ways of being (Fiske, 1991; Fiske, Kitayama, Markus, & Nisbett, 1998; Oyserman, Coon, & Kemmelmeier, 2002). Therefore, self-goals and self-regulatory processes could equally be part of the one's social identity or interdependent self-concept. In this sense, motivation is identity based and personal goals are likely to be scaffolded by relevant social identities and the goals and strategies cued by these social identities.

Social and Personal Goals Require Self-Regulatory Focus

Which kinds of self-goals are the focus of self-regulation? To date, research on self-regulation has either omitted reference to whether something is or is or is not a self-goal or focused almost exclusively on personal goals (Gollwitzer, Fujita, & Oettingen, 2004; Gollwitzer, & Kirchhof, 1998). Thus, researchers have either asked about self-regulatory goals such as dieting and exercising (Bagozi & Kimmel, 1995; Herman & Polivy, 2004), or asked about pursuit of research tasks such as pressing buttons in response to the appearance of a letter or number (Brandstätter et al., 2001). These latter studies assume that the research task becomes a self-goal. Similarly, when children are studied, self-regulation often

focuses on children's ability (or willingness) to follow instructions of adults (typically the instructions of the child's mother or of the researcher). Children are asked to wait and not to eat a treat or to work on a boring task and not play with toys (Mischel, Shoda, & Rodriguez, 1989). Indeed, in children self-regulation is often called effortful control and delay of gratification to highlight the appetitive nature of the self-regulation being studied. Thus studies of "self-regulation" often omit actual self-goals. Yet motivation and thus "self-regulation" must be identity based to be meaningful.

When a self-goal is taken into account, goals are described in terms of personal identity. Yet social identity (Abrams, 1994; Abrams & Brown, 1989), self-categorization (Foddy & Kashima, 2002; Onorato & Turner, 2002), and cultural psychology (Markus & Kitayama, 1991; Oyserman, Kemmelmeier, & Coon, 2002) models all highlight the importance of examining the self as consisting of personal traits, abilities, and goals and the traits, abilities, and goals one takes on as part of inclusion in social units—such as the family—and social categories—such as gender, race–ethnicity, and other cultural groups. More generally, the goals that are the focus of self-regulation may be conceptualized as individual or group level. That is, the image one is controlling one's behavior in pursuit of can be a possible self embedded in a social identity. Thus, doing homework may be part of a "smart" or "successful" personal possible self, but it might also be part of self-regulation to attain "smart African American" social identity.

For example, doing homework may be seen as part of what "we" do. The "we" or social identity at the root of self-regulation may vary. In the case of doing homework, the social identity may be that of team member, family member, racial–ethnic, social class, or gender group. Doing homework may be part of meeting the minimal grade-point average (GPA) requirement to participate in sports, part of meeting the GPA requirements needed to be a "scholar athlete," part of being a good son or daughter, part of being a girl, or part of one's social class or racial–ethnic identity. To the extent that homework is linked with engagement in a social identity, desire to engage this identity will increase self-regulation. The same goal—high GPA—can thus be cued by a variety of social identities. In each case, pursuit of the goal will feel genuinely self-defining because motivation is identity based.

If the identity loses luster (e.g., becoming a good son feels less central to identity during adolescent identity development), then so too will the self-regulatory effort put into attaining goals relevant to the social identity. While self-regulation is central to our understanding of what it means to be a self, self-regulation is not commonly associated with social aspects of self-concept; social or collective identities and questions about differences in self-regulatory style or focus are only beginning to be examined in the fields of cultural and cross-cultural psychology.

In spite of its relative neglect in current research and theorizing, the notion that self-regulation is importantly directed by social identity has roots in Cooley's (1902) and Mead's (1934) frameworks. These early conceptualizations directed attention to the importance of social and interpersonal context in self-focused emotion and self-regulation. Indeed, it seems intuitively obvious that self-regulation is linked with the social aspects of identity and not simply with the personal aspects of identity. Much of self-regulation involves a combination of inhibition of socially inappropriate responses and centralization of socially appropriate responses or goals. Self-conscious emotions (i.e., shame and guilt) are likely to play an important part in motivating self-regulation to behave in socially appropriate ways (Baldwin & Baccus, 2004; Tracy & Robins, 2004).

What constitutes being socially appropriate, of course, depends on social identity—the person one is as characterized by fit with ingroup others. We imagine how others would respond, we feel pride or shame at ourselves as a result of these imagined responses, and these self-generated emotions focus self-regulatory effort to become the kind of person of whom relevant others are proud, not ashamed. In this sense, social identities provide ongoing context, clarifying both what would be prideful for people like me and what would be shameful for people like me. Because humans are wary of social approbation, we are mindful not to behave in ways that would cause shame. Because what would cause shame depends on what is valued, what is devalued, and what is irrelevant to the groups we belong to and have incorporated into our sense of self, social identities are important. Even the ways in which self-regulation is carried out are likely to be importantly shaped by social context and the social content of self-concept.

Possible Selves and Self-Regulation

Self-regulation is central to attaining one's self-relevant goals—one's possible selves (Markus & Nurius, 1986; Oyserman et al., 2003), one's wished for ideal selves, or obligated ought selves (Higgins, 1997). Possible selves have been shown to influence outcomes requiring self-regulation such as academic striving (Oyserman et al., 1995) and health-related behavior (Hooker & Kaus, 1994). Importantly, even when possible selves embedded in different social identities focus on the same issues, differences in social identities seem to inform content of strategies to attain these possible selves. For example, Oyserman and her colleagues (1995) found that whereas both African American and White first-generation college students had possible selves focused on academic attainment, the strategies these students described differed. African American students focused on what they could do to avoid becoming like their negative, feared academic failure possible selves, while White students described strategies to work toward positive academic possible selves.

Social identities that feel centrally defining and important are more likely to influence strategies. In the case of academic possible selves, compared with minority students, White students report that the social identity of "college student" is a more central and important identity and are more likely to believe that having this identity will facilitate attaining their important possible selves

(Cameron, 1999). However, relatively little attention has been paid to the ways that social identities influence self-regulation.

Even though self-regulation is typically described in terms of focus on attaining personal possible selves, because the goals one strives to achieve are likely to be the goals valued in one's social context, self-regulation is also central to social esteem and attainment of social possible selves (Bandura, 2001). Indeed, to the extent that all members of society are responsible for carrying out plans and fulfilling obligations, self-regulation is central to the social construction of humanness and social identity. While goals are typically described as part of the individualistically focused self, one's goals clearly are also embedded in social identities as well.

The efficacy of a possible self framework for improving self-regulation has been documented in research focused on school-focused possible selves and health-focused possible selves. Oyserman, Terry, and Bybee (2002) randomly assigned eighth graders to receive a seven-session after-school program of small-group activities that focused on possible selves. Youth in the intervention group attained better grades, had better in-class behavior, and skipped class less often. This initial test of a possible selves–based intervention was conceptually replicated as an in-school randomized clinical trial. Oyserman, Bybee, and Terry (2006) randomly assigned the cohort of eighth graders in three Detroit middle schools to receive the intervention as the first part of their elective sequence, during the first 11 weeks of school, with control youth receiving their regular elective. Follow-up data were collected at the end of the eighth-grade school year and again in the following year for a 2-year follow-up. The possible self-focused intervention improved both the self-control and the self-will aspects of self-regulation. In terms, of self-control, school records showed fewer unexcused absences (youth refrained from skipping class) and teacher report showed less engagement in disruptive behavior (hitting, threatening the teacher). Increase in self-will was reflected in teacher-reported increase in active engagement with learning (asking questions after class, coming to class prepared); youth also reported more time spent in homework preparation (based on a weekly diary method). Self-regulation had positive affective consequences—self-control (e.g., not skipping) predicted fewer depressive symptoms—and positive consequences for goal attainment—self-will (e.g., more time doing homework)—predicted better grades.

Effects were mediated by the impact of the intervention on possible selves. Youth in the intervention group had more balanced (positive and negative) school-oriented possible selves and were more concerned about avoiding off-track possible selves, such as becoming pregnant or involved with drugs, than youth in the control group. Self-regulation was also targeted; balanced school-oriented possible selves predicted more engagement with school and time spent in homework, not less skipping or less disruptive behavior. Feared off-track possible selves predicted less skipping school and less disruption but not more time spent in homework or engagement

with school. While these possible selves might be assumed to be part of these teen's personal identities, Oyserman and colleagues argued that they had become part of the teen's RES, showing that school-focused possible selves were positively associated with racial–ethnic identity in intervention youth but orthogonal to this social identity for control group youth (Oyserman, Bybee, & Terry, 2006).

Possible selves have also been used in a number of health and exercise-focused interventions. In one, possible selves of adults (averaging 68 years of age) were predictive of effective use of an exercise program (Whaley & Shrider, 2005). Another, briefer, possible selves–based intervention involved college students. Those asked to envision what they would be like in 10 or 20 years if they did not exercise regularly were more likely to report increased exercise a week later in a follow-up ostensibly unrelated phone interview (Ouellette, Hessling, & Gibbons, 2005). A third brief intervention had college students write for 20 minutes each day for 4 consecutive days about an important trauma, one's best (most positive) possible self, both, or neither (King, 2001). Five months later, those who wrote about a possible self, a trauma, or both a possible self and a trauma had better health outcomes; effects of writing about a possible self were as good or better than those for writing about a trauma. King (2001) suggests that either task evoked self-regulatory behavior. While these studies do not contextualize possible selves as part of either personal or social identities, our research on connection between racial–ethnic and social class-based social identities (Oyserman, Fryberg, & Yoder, 2006) suggests that these possible selves are likely to have been embedded in relevant social identities.

SELF-REGULATION OF SOCIAL IDENTITIES IS A CULTURALLY EMBEDDED PROCESS

Being able to control oneself is likely a universally developed skill. What is likely to be culturally determined are the goals toward which one self-controls and how one goes about controlling oneself as well as the circumstances that cue self-control. Clearly, pursuing traditional values requires self-regulation, so does pursuing postmodern values (see Inglehart & Baker, 2000). Whether living in a culture focused on individualism (vs. collectivism), secularism (vs. tradition), or self-expression (vs. survival) (Inglehart & Oyserman, 2004), the ability to control one's actions and will oneself into action is necessary. Self-regulation should be just as necessary whether focused on achieving idiosyncratic personal possible selves and goals or on consensually accepted social or relational possible selves and goals. Indeed, self-control is part of Schwartz's (Schwartz & Bilsky, 1987) restrictive-conformity value domain. Because self-regulation more generally is required to attain other universally valued goals, it is likely to be universally part of socialization. Childrearing values across countries do include socializing for self-regulation—variously termed "responsibility," "obedience," "determination," "perseverance," "thrift,"

and "good manners." Cultures differ in which of these particular aspects of self-regulation they most centrally endorse (e.g., Inglehart & Baker, 2000), but not on whether some form of self-regulation is a desired outcome of socialization. Like other core social values, universality of self-regulation does not mean uniformity in style (e.g., Schwartz & Bilsky, 1990); cultures are likely to vary systematically on the form that self-regulation takes—focusing on self-control and self-will in varying proportions and across different domains.

Developmentally, there is evidence of interplay between that which is culturally rewarded and self-regulation. Rodriguez and colleagues (2005) document cross-cultural difference in the relationship between effortful self-control and positive affectivity (extraversion/surgency) versus negative affectivity by 6 or 7 years of age. They compare U.S. and Chinese samples. Specifically, in Chinese (People's Republic of China) children of these ages, effortful control is negatively associated with extraversion/surgency—operationalized as activity, smiling and laughing, high-intensity pleasure, impulsivity, lack of shyness, and positive anticipation. Yet extraversion/surgency is orthogonal to effortful control among U.S. children of these ages. Conversely, in the United States, effortful control is negatively associated with negative affectivity—operationalized as being fearful, angry, sad, difficult to sooth, and high in discomfort—yet negative affectivity is orthogonal to effortful control among Chinese children. These findings suggest between-culture differences in the behaviors viewed as worthy of control (negative affect in the United States, outgoing behavior in China). These early differences in factors related to effortful control are congruent with differences in cultural values found among college students (e.g., Oyserman, Coon, & Kemmelmeier, 2002).

These results suggest that self-regulatory style and capacity are not only personal but also contextually and culturally shaped. Contexts make salient appropriate future horizons for self-regulation, varying from more proximal (e.g., getting through the afternoon without insulting Aunt Millie) to more distal (e.g., finishing high school) and even lifelong (e.g., being successful). Cultures endorse and therefore make salient some ways of self-regulating over others (e.g., Is it best to "shoot for the stars" and "say what you think" or does "haste make waste" and "fools rush in where angels fear treading"?). Within the context of universal socialization for self-regulation, the style with which goals are pursued may differ—does one take aim at attaining the goal, focusing on success with little concern for possible negative consequences of failure, or take care in attaining the goal, focusing on possible repercussions and negative implications along the way (Higgins, 1997, 2000; O'Brien & Oyserman, 2006)? Indeed, initial work in this area suggests that cross-culturally, prevention focus (typically studied as loss-framed focus) is more common in Eastern than in Western contexts (Aaker & Lee, 2001; Briley & Wyer 2002; Lee, Aaker, & Gardner, 2000).

Differences in focus of self-regulatory style do not imply differences in valuation of self-regulation. For example, within American culture, self-regulation is a valued trait; failures are assumed to be due to insufficient effort or insufficient exertion of will. Self-help manuals can be seen as the cultural artifact embodying this belief. They are produced in an ever updated abundance and bought in great numbers by individuals who believe that the potential for change is limited only by one's willingness to self-improve. Americans believe in the perfectible possible self—one could be one's thinner, better-toned, more patient, more ecologically friendly, better parent and more religiously observant self, if one just tried. At the same time, Americans also believe that talent is a fixed entity, not something that can be learned or attained through effort (Dweck, 2002). Things one cannot do well are assumed to be "not me" arenas, the proclivity to accept that which is as that which is inevitable melds with the belief that the willful yet untalented can only go so far. As reflected in the "for dummies" manuals, there are many things that the willful yet untalented can learn even if mastery is reserved for the talented.

American cultural frame embodies a Protestant focus on free will. The ability to control one's self is a basic assumption regulating not only personal goal setting but also one's relation to others and obligations within the social system. Yet belief in will power is not solely a Protestant cultural artifact—Catholicism also carries with it a focus on will via endurance of conditions that cannot be changed (Tropman, 2002). Moreover, self-regulation is clearly not simply an American or a Western cultural style. Non-Western cultures also centralize the ability to endure, to cheerfully do one's duty in the face of odds. This formulation of will is deeply part of other distinctly different cultural frames such as Hinduism (Weber, 1958) and Confucianism (Finegan, 1952).

Will in these non-Protestantism-infused contexts may focus less on the self as controlling the environment as on controlling oneself to fit the needs of the context or one's station within a larger context. Self-regulation thus can involve learning to control oneself to meet the demands of the context just as much as it can involve using one's resources to pursue personal goals. In this way, self-regulation can be just as central to humanness when the goal is shaping the self to the exigencies of the context as when the goal is pursuit of personally defined goals. Self-regulation can assume that effort; improves all goal pursuit or that effort only goes so far; in either case, the nature of motivation is identity based.

CONCLUDING COMMENTS

While the exact nature of what is universal and what is culture specific in the relationship between social identity and self-regulation has yet to be fully researched, it seems reasonable to assume that the process model described in this chapter is broadly applicable. Cultural and social factors are likely to influence the centrality of possible selves explicitly embedded in social identities as well as the appropriate style chosen to pursue these possible selves. While there is some reason to assume that self-regulatory style, and therefore the strategies one uses to avoid failure and attaining

success, is culturally linked, it is also clear that strategies are more concretely linked to specific social identities, what people like "us" do.

I have developed an identity-based motivation process model linking social identity with self-regulation by integrating a number of relevant lines of research. Within this model, culture matters for self-concept because it influences both content and process of self-concept. It influences what is of value, what matters, and therefore how one is likely to define the self but also which means toward goal attainment are endorsed, which are merely accepted, and which are denigrated. All cultures value self-regulation—controlling the self and molding the self to become more like valued possible self-goals. However, cultures differ in which self-regulation processes are likely to be primed and whether self-regulation is framed more in terms of fitting into a social role or creating a unique self.

When studied in terms of the individualism–collectivism axes, culture influences chronic salience of social identities as well as how one is likely to self-regulate—by eagerly pursuing goals in ways likely to maximize chances of success or cautiously moving forward in ways unlikely to produce errors and regret. Social identities contain traits and characteristic ways of being that are relevant to the social group defined by the identity. Therefore, they influence the possible self-goals and strategies to attain them of individuals who define themselves in terms of these social identities. Individuals are likely to incorporate social information as part of their identity unless the social information is framed as separate from the self. This is likely when social identities are primed and the information is tagged as relevant to an outgroup that cannot be assimilated into ingroup identity and when a bridge between the outgroup and ingroup identities has not previously been created.

Current social cognition theories focus on self-regulation as a contextually cued cognitive or "hot" cognitive process, infused with affect. The cognitive processes underlying self-regulation are likely to be universal, triggered by self-goals formulated as an "I" or "we" identity and carried out with strategies that are "I" or "we" identity congruent. Thus, self-regulation can involve controlling the self via inaction—not engaging in currently hedonically pleasurable activities (not sleeping in, refraining from smoking, not eating certain foods). This form of self-regulation or self-control makes sense when inaction is in pursuit of longer-term goals (being successful, being healthy, being a good member of one's religious group) that require not engaging in undermining actions along the way. Self-regulation can also involve willing the self into action or sustaining action—engaging in action that may or may not be pleasurable (preparing for class, studying, setting an alarm). This form of self-regulation or self-will makes sense when action is in pursuit of longer-term goals (e.g., learning, getting good grades, becoming successful, or fulfilling social role obligations) that require constant vigilance and action. Whether providing reasons for action or for inaction, social identities are central to the self-regulatory process, motivation is identity based.

REFERENCES

Aaker, J., & Lee, A. (2001). "I" seek pleasures and "we" avoid pains. *Journal of Consumer Research*, *28*, 33–49.

Abrams, D. (1994). Social self-regulation. *Personality and Social Psychology Bulletin [Special issue: The self and the collective]*, *20*, 473–483.

Abrams, D. (1999). Social identity and social cognition. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 197–229). Malden, MA: Blackwell.

Abrams, D., & Brown, R. (1989). Self-consciousness and social identity: Self-regulation as a group member. *Social Psychology Quarterly*, *52*, 311–318.

Arnett, D., German, S., & Hunt, S. (2003). The identity salience model of relationship marketing success. *Journal of Marketing*, *67*, 89–105.

Avila, C. (2001). Distinguishing BIS-mediated and BAS-mediated disinhibition mechanisms: A comparison of disinhibition models of Gray (1981, 1987) and of Patterson and Newman (1993). *Journal of Personality and Social Psychology*, *80*, 311–324.

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P., & Rodriguez, M. (2000). Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity. *Journal of Personality and Social Psychology*, *79*, 776–792.

Bagozzi, R., & Kimmel, S. (1995). A comparison of leading theories for the prediction of goal-directed behaviours. *British Journal of Social Psychology*, *34*, 437–461.

Baldwin, M., & Baccus, J. (2004). Maintaining a focus on the social goals underlying self-conscious emotions. *Psychological Inquiry*, *15*, 139–144.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26.

Baumeister, R. (2002). Ego depletion and self-control failure: An energy model of the self's executive function. *Self and Identity*, *1*, 129–136.

Baumeister, R., & Heatherton, T. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.

Baumeister, R., & Vohs, K. (2003). Self-regulation and the executive function of the self. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 197–217). New York: Guilford Press.

Bernhardt, P. C., Dabbs, J. M., Fielden, J. A., & Lutter, C. D. (1998). Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiology and Behavior*, *65*, 59–62.

Bhattacharya, C., Rao, H., & Glynn, M. (1995). Understanding the bond of identification: An investigation of its correlates among art museum members. *Journal of Marketing*, *46*, 46–57.

Blanton, H. (2001). Evaluating the self in the context of another: Assimilation and contrast effects in social comparison. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 75–87). Mahwah, NJ: Erlbaum.

Blanton, H., Christie, C., & Dye, M. (2002). Social identity versus reference frame comparisons: The moderating role of stereotype endorsement. *Journal of Experimental Social Psychology*, *38*, 253–267.

Boen, F., Vanbeselaere, N., & Feys, J. (2002). Behavioral consequences of fluctuating group success: An internet study of soccer-team fans. *Journal of Social Psychology*, *142*, 769–781.

Boen, F., Vanbeselaere, N., Pandelaere, M., Dewitte, S., Duriez, B., Snauwaert, B., et al. (2002). Politics and basking-in-reflected-glory: A field study in Flanders. *Basic and Applied Social Psychology*, *24*, 205–214.

Bogg, T., & Roberts, B. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, *130*, 887–919.

Brandstätter, V., Lengfelder, A., & Gollwitzer, P. (2001). Implementation intentions and efficient action initiation. *Journal of Personality and Social Psychology*, *81*, 946–960.

Branscombe, N. R., & Ellemers, N. (1998). Coping with group-based discrimination: Individualistic versus group-level strategies. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 243–266). San Diego, CA: Academic Press.

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, *77*, 135–149.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, *71*, 83–93.

Brewer, M., & Weber, J. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, *66*, 268–275.

Briley, D., & Wyer, R. S., Jr. (2002). The effect of group membership salience on avoidance of negative outcomes: Implications for social and consumer decisions. *Journal of Consumer Research*, *29*, 400–415.

Brooks-Gunn, J., & Markman, L., (2005). The contribution of parenting to ethnic and racial gaps in school readiness. *Future of Children*, *15*, 139–168.

Bronson, M. B. (2000). *Self-regulation in early childhood: Nature and nurture*. New York: Guilford Press.

Brown, J. (1998). *The self.* Boston: McGraw-Hill.

Burke, P. J. (2003). Relationships among multiple identities. In P. J. Burke, T. J. Owens, R. T. Serpe, & P. A. Thoits (Eds.), *Advances in identity theory and research* (pp. 195–214). New York: Kluwer Academic/Plenum Press.

Camacho, C., Higgins, E. T., & Luger, L. (2003). Moral value transfer from regulatory fit: What feels right *is* right and what feels wrong *is* wrong. *Journal of Personality and Social Psychology*, *84*, 498–510.

Cameron, J. (1999). Social identity and the pursuit of possible selves: Implications for the psychological well-being of university students. *Group Dynamics: Theory, Research, and Practice*, *3*, 179–189.

Carver, C., & Scheier, M. (1990). Principles of self-regulation: action and emotion. In E. T. Higgins & R. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 3–52). New York: Guilford Press.

Carver, C., Sutton, S., & Scheier, M. (2000). Action, emotion, and personality: Emerging conceptual integration. *Personality and Social Psychology Bulletin*, *26*, 741–751.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*, 319–333.

Catrambone, R., & Markus, H. (1987). The role of self-schemas in going beyond the information given. *Social Cognition*, *5*, 349–68.

Chen, S., Chen, K., & Shaw, L. (2004). Self-verification motives at the collective level of self-definition. *Journal of Personality and Social Psychology*, *86*, 77–94.

Cialdini, R., Borden, R., Thorne, A., Walker, M., Freeman, S., & Sloan, R. (1976). Basking in reflected glory. *Journal of Personality and Social Psychology*, *45*, 1232–1242.

Collins, R. (1996). For better or for worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*, 51–69.

Conner, D., & Cross, D. (2003). Longitudinal analysis of the presence, efficacy and stability of maternal scaffolding during informal problem-solving interactions. *British Journal of Developmental Psychology*, *21*, 315–334.

Cooley, C. (1902). *Human nature and the social order*. New York: Scribner.

Cousins, S. (1989). Culture and self-perception in Japan and the United States. *Journal of Personality and Social Psychology*, *56*, 124–131.

Crocker, J., & Major, B. (1989). The self-protective properties of stigma. *Psychological Review*, *96*, 608–630.

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, *60*, 218–228.

Cross, S. E., Gore, J. S., & Morris, M. L. (2003). The relation-interdependent self-construal, self-concept consistency, and well-being. *Journal of Personality and Social Psychology*, *85*, 933–944.

Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin*, *122*, 5–37.

Cross, S. E., Morris, M. L., & Gore, J. S. (2002). Thinking about oneself and others: The relational-interdependent self-construal and social cognition. *Journal of Personality and Social Psychology*, *82*, 399–418.

Cross, W., & Fhagen-Smith, P. (1996). Nigrescence and ego-identity development: accounting for differential black identity patterns. In P. Pedeson, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures* (4th ed., pp. 108–123). Thousand Oaks, CA: Sage.

Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes*, *69*, 117–132.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. New York: HarperCollins.

Dweck, C. (2002). The development of ability conceptions. In A. Wigfield & J. Eccles (Eds.), *Development of achievement motivation* (pp. 57–88). San Diego, CA: Academic Press.

Dweck, C., & London, B. (2004). The role of mental representation in social development. *Merrill-Palmer Quarterly*, *50*, 428–444.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*, 73–92.

Finegan, J. (1952). *The archeology of world religions: The background of primitivism, Zoroastrianism, Hinduism, Jainism, Buddhism, Confucianism, Taoism, Shinto, Islam, and Sikhism*. Princeton, NJ: Princeton University Press.

Fiske, A. (1991). *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. New York: Free Press.

Fiske, A., Kitayama, S., Markus, H., & Nisbett, R. (1998). The cultural matrix of social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.) *The handbook of social psychology* (Vol. 2, pp. 915–981). New York: McGraw-Hill.

Foddy, M., & Kashima, Y. (2002). Self and identity: What is the conception of the person assumed in the current literature? In Y. Kashima, M. Foddy, & M. Platow (Eds.), *Self and identity: Personal, social, and symbolic* (pp. 3–25). Mahwah, NJ: Erlbaum.

Gaertner, L., Sedikides, C., Vevea, J. & Iuzzini, J. (2002). The "I," the "we," and the "when": A meta-analysis of motivational primacy in self-definition. *Journal of Personality and Social Psychology*, *83*, 574–591.

Gardner, W., Gabriel, S., & Dean, K. (2004). The individual as "melting pot." *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *22*, 181–201.

Gardner, W., Gabriel, S., & Hochschild, L. (2002). When you and I are "we," you are not threatening: The role of self-expansion in social comparison. *Journal of Personality and Social Psychology*, *82*, 239–251.

Gardner, W., Gabriel, S., & Lee, A. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, *10*, 321–326.

Gilbert, D., Giesler, R., & Morris, K. (2005).When comparisons arise. *Journal of Personality and Social Psychology*, *69*, 227–236.

Gollwitzer, P. M., Fujita, K., & Oettingen, G. (2004). Planning and the implementation of goals. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 211–228). New York: Guilford Press.

Gollwitzer, P., & Kirchhof, O. (1998). The willful pursuit of identity. In J. Heckhausen & C. Dweck (Eds.), *Motivation and self-regulation across the lifespan* (pp. 389–423). New York: Cambridge University Press.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septohippocampal system*. New York: Oxford University Press.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion*, *4*, 269–288.

Haberstroh, S., Oyserman, D., Schwarz, N., Kuhnen, U., & Ji, L. (2002). Is the interdependent self more sensitive to question context than the independent self? *Journal of Experimental Social Psychology*, *38*, 323–329.

Herman, C. P., & Polivy, J. (2004). The self-regulation of eating: Theoretical and practical problems. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 492–508). New York: Guilford Press.

Higgins, E. T. (1996). The "self digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*, 1062–1083.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*, 1140–1153.

Higgins, E. T., & Silberman, I. (1998). Development of regulatory focus: promotion and prevention as ways of living. In J. Heckhausen & C. Dweck (Eds.), *Motivation and self-regulation across the lifespan* (pp. 78–113). New York: Cambridge University Press.

Hirt, E. R., Zillman, D., Erickson, G. A., & Kennedy, C. (1992). Costs and benefits of allegiance: Changes in fans' self-ascribed competencies after team victory versus defeat. *Journal of Personality and Social Psychology*, *63*, 724–738.

Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage.

Hogg, M. A. (2003). Social identity. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 462–479). New York: Guilford Press.

Hooker, K., & Kaus, C. R. (1994). Health-related possible selves in young and middle adulthood. *Psychology and Aging*, *9*, 126–133.

Inglehart, R., & Baker, W. (2000). Modernization, cultural change and persistence of traditional values. *American Sociological Review*, *65*, 19–51.

Inglehart, R., & Oyserman, D. (2004). Individualism, autonomy, self-expression and human development. In H. Vinken, J. Soeters, & P. Ester (Eds.), *Comparing cultures, dimensions of culture in a comparative perspective* (pp. 74–96). Leiden, The Netherlands: Brill.

James, W. (1927). *The principles of psychology*. New York: Dover. (Original work published 1890)

Kâgitçibasi, Ç. (1994). A critical appraisal of individualism and collectivism: Toward a new formulation. In U. Kim, H. C. Triandis, Ç. Kâgitçibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method, and applications* (pp. 52–65). Thousand Oaks, CA: Sage.

Kâgitçibasi, Ç. (2002). Patterns of autonomy and interdependence. In R. Liljestrom & E. Özdalga (Eds.), *Autonomy and dependence in the family* (pp. 19–38). Istanbul, Turkey: Swedish Research Institute.

Kanagawa, C., Cross, S., & Markus, H. (2001). "Who am I?": The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, *27*, 90–103.

Kemmelmeier, M., & Oyserman, D. (2001). Gendered influence of downward social comparison on current and possible selves. *Journal of Social Issues*, *57*, 129–148.

Kemmelmeier, M., Oyserman, D., & Brosh, H. (2005). *Gender and upward social comparison*. Unpublished manuscript, University of Michigan.

King, L. (2001). The health benefits of writing about life goals. *Personality and Social Psychology Bulletin*, *27*, 798–807.

Kruglanski, A., Shah, J., Fishbach, A., Friedman, R., Chun, W., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 331–378). New York: Academic Press.

Kuhn, M., & McPartland, T. (1954). An empirical investigation of self-attitudes. *American Sociological Review*, *19*, 68–76.

Kühnen, U., & Haberstroh, S. (2004). Self-construal activation and focus of comparison as determinants of assimilation and contrast in social comparisons. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *22*, 289–310.

Kühnen, U., & Hannover, B. (2000). Assimilation and contrast in social comparisons as a consequence of self-construal activation. *European Journal of Social Psychology*, *30*, 799–811.

Kühnen, U., & Oyserman, D. (2002). Thinking about the self influences thinking in general: Cognitive consequences of salient self-concept. *Journal of Experimental Social Psychology*, *38*, 492–499.

Lee, A., Aaker, J., & Gardner, W. (2000). The pleasures and pains of distinct self-construals: The role of interdependence in regulatory focus. *Journal of Personality and Social Psychology*, *78*, 1122–1134.

Liberman, N., Molden, D., Idson, L., & Higgins, E. T. (2001). Promotion and prevention focus on alternative hypotheses: Implications for attributional functions. *Journal of Personality and Social Psychology*, *80*, 5–18.

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, *73*, 91–103.

Mandel, N. (2003). Shifting selves and decision making: The effects of self-construal priming on consumer risk-taking. *Journal of Consumer Research*, *30*, 30–40.

Markus, H. (1977). Self-schemas and social processing information about the self. *Journal of Personality and Psychology*, *35*, 63–78.

Markus, H., Crane, M., Bernstein, S., & Siladi, M. (1982). Self-schemas and gender. *Journal of Personality and Social Psychology*, *42*, 38–50.

Markus, H., Hamill, R., & Sentis, K. (1987). Thinking fat: Self-schemas for body weight. *Journal of Applied Social Psychology*, *17*, 50–71.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *20*, 568–579.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Markus, H., & Oyserman, D. (1989). Gender and thought: The role of the self-concept. In M. Crawford & M. Gentry (Eds.), *Gender and thought: Psychological perspectives* (pp. 100–127). New York: Springer-Verlag.

Markus, H., & Sentis, K. (1982). The self in social information processing. In J. Suls (Ed.), *Social psychological perspectives on the self* (Vol. 1, pp. 41–70). Hillsdale, NJ: Erlbaum.

Markus, H., Smith, J., & Moreland, R. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology*, *49*, 1494–1512.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337.

Maslow, A. (1970). *Motivation and personality*. New York: Harper & Row.

Matsumoto, D. (1999). Culture and self: An empirical assessment of Markus and Kitayama's theory of independent and interdependent self-construals. *Asian Journal of Social Psychology*, *2*, 289–310.

Mead, G. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Mischel, W., & Ayduk, O. (2004). Willpower in a cognitive-affective processing system: The dynamics of delay of gratification. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 99–129). New York: Guilford Press.

Mischel, W., Shoda, Y., & Rodriguez, M. (1989). Delay of gratification in children. *Science*, *244*, 933–938.

Montepare, J., & Clements, A. (2001). "Age schemas": Guides to processing information about the self. *Journal of Adult Development*, *8*, 99–108.

Mussweiler, T., & Bodenhausen, G. V. (2002). I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology*, *82*, 19–32.

Mussweiler, T., Rüter, K., & Epstude, K. (2004).The ups and downs of social comparison: Mechanisms of assimilation and contrast. *Journal of Personality and Social Psychology*, *87*, 832–844.

O'Brien, B., & Oyserman, D. (2006). *Promoting law and order or preventing injustice? How self-regulatory focus relates to responses to crime*. Unpublished manuscript, University of Michigan.

Onorato, R., & Turner, J. (2002). Challenging the primacy of the personal self: The case for depersonalized self-conception. In Y. Kashima, M. Foddy, & M. Platow (Eds.), *Self and identity: Personal, social, and symbolic* (pp. 145–178). Mahwah, NJ: Erlbaum.

Osborne, J. W. (1995). Academics, self-esteem and race: A look at the underlying assumptions of the disidentification hypothesis. *Personality and Social Psychology Bulletin*, *21*, 449–455.

Ouellette, J., Hessling, R., & Gibbons, F. (2005). Using images to increase exercise behavior prototypes versus possible selves. *Personality and Social Psychology Bulletin*, *31*, 610–620.

Oyserman, D. (1993). The lens of personhood: Viewing the self, others, and conflict in a multicultural society. *Journal of Personality and Social Psychology*, *65*, 993–1009.

Oyserman, D., Brickman, D., & Rhodes, M. (in press). Racial-ethnic identity: Content and consequences for African American and Latino youth. In A. Fuligni (Ed.), *Social categories, identities and educational participation*. New York: Russell Sage.

Oyserman, D., Bybee, D., & Terry K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology*, *91*, 188–204.

Oyserman, D., Bybee, D., Terry, K., & Hart-Johnson, T. (2004). Possible selves as roadmaps. *Journal of Research in Personality*, *38*, 130–149.

Oyserman, D., Coon, H., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3–72.

Oyserman, D., Fryberg, S., & Yoder, N. (2006). *Identity-based motivation: McDonald's and I are one*. Unpublished manuscript, University of Michigan.

Oyserman, D., Gant, L., & Ager, J. (1995). A socially contextualized model of African American identity: Possible selves and school persistence. *Journal of Personality and Social Psychology*, *69*, 1216–1232.

Oyserman, D., Kemmelmeier, M., & Coon, H. (2002). Cultural psychology: A new look. *Psychological Bulletin*, *128*, 110–117.

Oyserman, D., Kemmelmeier, M., Fryberg, S., Brosh, H., & Hart-Johnson, T. (2003). Racial–ethnic self-schemas. *Social Psychology Quarterly*, *66*, 333–347.

Oyserman, D., & Lee, W. S. (in press). Priming culture: Culture as situated cognition. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology*. New York: Guilford Press.

Oyserman, D., & Markus, H. (1993). The sociocultural self. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 4, pp. 187–220). Hillsdale, NJ: Erlbaum.

Oyserman, D., Terry, K., & Bybee, D. (2002). A possible selves intervention to enhance school involvement. *Journal of Adolescence*, *24*, 313–326.

Pelham, B., & Wachsmuth, J. (1995). The waxing and waning of the social self. *Journal of Personality and Social Psychology*, *69*, 825–838.

Reed, A. (2004). Activating the self-importance of consumer selves: Exploring identity salience effects on judgments. *Journal of Consumer Research*, *31*, 286–295.

Rhodes, M., Oyserman, D., & Brickman, D. (2006). *Racial-ethnic self-schemas well-being and academic outcomes*. Unpublished manuscript, University of Michigan.

Roccas, S., & Brewer, M. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6*, 88–106.

Rodriguez, M., Ayduk, O., & Aber, J. (2005). A contextual approach to the development of self-regulatory competencies: The role of maternal unresponsivity and toddlers' negative affect in stressful situations. *Social Development*, *14*, 136–157.

Rodriguez, M., Mishel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social Psychology*, *57*, 358–367.

Rothbart, M., & Rueda, R. (2005). The development of effortful control. In U. Mayr, E. Awh, & S. Keele (Eds.), *Developing individuality in the human brain: A tribute to Michael Posner* (pp. 167–188). Washington, DC: American Psychological Association.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*, 194–201.

Schwartz, S., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, *53*, 550–562.

Schwartz, S., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values. *Journal of Personality and Social Psychology*, *58*, 878–891.

Schwarz, N., & Bless, H. (in press). Mental construal processes: The inclusion/exclusion model. In D. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology*. New York: Psychology Press.

Schwarz, N., & Bless, H. (1992). Construction reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, NJ: Erlbaum.

Schwarz, N. Bless, H., Wänke, M., & Winkielman, P. (2003). Accessibility revisited. In G. Baudenhausen & A. Lambert (Eds.), *Foundations of social cognition* (pp. 51–77). Mahwah, NJ: Erlbaum.

Seeley, E., & Gardner, W. (2003). The "selfless" and self-regulation: the role of chronic other-orientation in averting self-regulatory depletion. *Self and Identity*, *2*, 103–117.

Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, *87*, 38–56.

Serbin, L., & Karp, J. (2004). The intergenerational transfer of psychosocial risk: Mediators of vulnerability and resilience. *Annual Review of Psychology*, *55*, 333–363.

Sherman, S. J., Judd, C. M., & Park, B. (1989). Social cognition. *Annual Review of Psychology*, *40*, 281–326.

Singelis, T. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*, 580–591.

Stapel, D. A., & Koomen, W. (2001). I, we, and the effects of others on me: How self-construal level moderates social comparison effects. *Journal of Personality and Social Psychology*, *80*, 766–781.

Stryker, S., & Burke, P. (2000). The past, present, and future of identity theory. *Social Psychology Quarterly*, *63*, 284–297.

Tajfel, H. (1972). Experiments in a vacuum. In J. Israel & H. Tajfel (Eds.), *The context of social psychology: A critical assessment* (pp. 69–119). London: Academic Press.

Tajfel, H. (1981). *Human groups and social identity*. Cambridge, UK: Cambridge University Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (2nd ed., pp. 33–48). Chicago: Nelson-Hall.

Taylor, S., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*, 569–575.

Thayer, R. (2001). *Calm energy: How people regulate mood with food and exercise*. New York: Oxford University Press.

Tracy, J., & Robins, R. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, *15*, 103–125.

Trafimow, D., Silverman, E., Fan, R., & Law, J. (1997). The effects of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology*, *28*, 107–123.

Trafimow, D., & Smith, M. (1998). An extension of the "two-baskets" theory to Native Americans. *European Journal of Social Psychology*, *28*, 1015–1019.

Trafimow, D., Triandis, H., & Goto, S. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology*, *60*, 649–655.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts *Psychological Review*, *96*, 506–520.

Triandis, H. C. (1995). *Individualism and collectivism.* Boulder, CO: Westview Press.

Tropman, J. (2002). *The Catholic ethic and the spirit of community*. Washington, DC: Georgetown University Press.

Turner, J., & Oakes, P. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, *25*, 237–252.

Turner, J. C., & Oakes, P. J. (1989). Self-categorization theory and social influence. In P. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 233–275). Hillsdale, NJ: Erlbaum.

Vohs, K., & Heatherton, T. (2001). Self-esteem and threats to self: Implications for self-construals and interpersonal perception. *Journal of Personality and Social Psychology*, *81*, 1103–1118.

Weber, M. (1958). *The religion of India; the sociology of Hinduism and Buddhism* (H. Gerth & D. Martindale, Trans. & Eds.). Glencoe, IL: Free Press.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 446–496). New York: McGraw-Hill.

Whaley, D., & Shrider, A. (2005). The process of adult exercise adherence: Self-perceptions and competence. *Sport Psychologist*, *19*, 148–163.

White, R. (1959). Motivation reconsidered: The concept of competence. *Psychological Bulletin*, *104*, 36–52.

White, R. (1960). Competence and psychosexual stages of development. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 8). Lincoln: University of Nebraska Press.

Wills, T. A. (1981). Downward social comparison principles in social psychology. *Psychological Bulletin*, *90*, 245–271.

Wilson, T., & Dunn, E. (2004). Self-knowledge: its limits, value and potential for improvement. *Annual Review of Psychology*, *55*, 493–518.

Wrosch, C., Scheier, M., & Carver, C. (2003). The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial. *Self and Identity*, *2*, 1–20.

Ybarra, O., & Trafimow, D. (1998). How priming the private self or collective self affects the relative weights of attitudes and subjective norms. *Personality and Social Psychology Bulletin, 24*, 362–370.

CHAPTER 19

Value

E. TORY HIGGINS

To understand why people do what they do, why they feel what they feel, psychologists need to know what matters to them, what they want and don't want. A central concept concerned with "mattering" and "wanting" is *value*. Indeed, Allport (1961) suggested that value priorities were the "dominating force in life" (p. 543). Scholars have been studying what it means to value something and where value comes from for centuries. For example, an entire branch of classical and modern philosophy is concerned with understanding the nature and source of one kind of value—moral or ethical value. It is surprising, then, that social psychologists who care deeply about why people do what they do and feel what they feel have paid relatively little attention to the concept of value per se. If one checks the subject index of social psychology textbooks and handbooks, one discovers only a few references to "value." Those references that do appear mostly involve discussions of people's shared beliefs about general desirable ends (e.g., social values such as "freedom" or "security").

This is not to say that social psychologists are unconcerned with what makes something valuable to someone and how the value of something influences his or her behavior and feelings. Indeed, they are very concerned. However, these issues are addressed in terms of other concepts, such as "commitment," "goal," "norm," and, especially, "attitude." For example, a central component of attitudes is evaluation, which certainly relates to the concept of value. However, the attitude literature has been more concerned with the implications of there being an evaluative dimension of attitudes, such as what it

means for predicting behavior, than examining where the "goodness" and "badness" itself come from. Psychological work related to the concepts of "commitment," "goal," "norm," and "attitude" is also clearly relevant to understanding the nature and sources of value (e.g., work on attitude formation; see Zanna & Rempel, 1988), but, with some significant exceptions (e.g., Brickman, 1987), it is rarely concerned directly with what value is and where it comes from. A need remains to study value as a general principle in its own right, which is the purpose of this chapter.

I begin by briefly considering how "value" has been formally defined, which provides a preliminary answer to the question, "What *is* value?" Next, I review in more detail sociological and psychological answers to the question, "*Where does value come from?*" The major purpose of this chapter is to review the answers to this question that have been offered in the literature. Five major viewpoints on where value comes from are reviewed: *value from need satisfaction*, *value from shared beliefs about what is desirable*, *value from actual self-relation to end states*, *value from evaluative inference*, and *value from experience*. The section on value from experience is the most extensive because this viewpoint has played a dominant role within philosophy and psychology generally and within social psychology in particular. I also believe that value from experience is critical to understanding the psychological nature of value. Five different value experiences are distinguished: *hedonic experiences*, *moral or ethical experiences*, *regulatory fit experiences*, *understanding experiences*, and *agentic experiences*. At the end of the chapter, I return to the more

general question, "What is value?" I propose an answer that is inspired by the insights and perspectives provided by previous answers to where value comes from, especially those that have emphasized experience as a source of value.

WHAT *IS* VALUE?

What is meant by something having value to someone? According to the *Oxford English Dictionary* (1971, p. 3587), to value something is to estimate or appraise in respect of value. But what is the "value" being estimated or appraised in respect of? The same dictionary provides the following definitions of "value":

1. That amount of some commodity, medium of exchange, etc., which is considered to be an equivalent for something else; a fair or adequate equivalent or return.
2. The material or monetary worth of a thing; the amount at which it may be estimated in terms of some medium of exchange or other standard of a similar nature.
3. Possessed of (a specified) material or monetary worth.
4. The equivalent (in material worth) of a specified sum or amount.
5. Worth or worthiness (of persons) in respect of rank or personal qualities.
6. The relative status of a thing, or the estimate in which it is held, according to its real or supposed worth, usefulness, or importance.

Compare these definitions of "value" to those provided by Webster's dictionary, which is another standard dictionary (*Webster's Ninth New Collegiate Dictionary*, 1989, p. 1303):

1. A fair return or equivalent in goods, services, or money.
2. The monetary worth of something; marketable price.
3. Relative worth, usefulness, or importance; degree of excellence.
4. Something intrinsically desirable.

There is considerable agreement between these dictionaries in their definitions of value. It is noteworthy that in both cases the primary definitions of value refer to equivalence to something in exchange, such as some material, commodity, or service, but especially equivalence in money—*the value of something as its monetary worth or marketable price*. What this type of definition accomplishes is to provide an operational definition of "value." It is the kind of definition that would please behaviorists, or traditional economists for that matter. Indeed, it is useful as a measure that can be used across various contents and as a dependent measure in research on value. What it does not do is answer the question of what value is conceptually. What exactly is this "value" that people will exchange money for? Why are people willing to give up goods and services in exchange for this thing called "value"?

Another way of thinking about these primary definitions is that they concern the consequences of something having value—*what happens when something has value*? If

something has value, it can be exchanged for something else that is considered equivalent (i.e., a fair exchange). The other dictionary definitions, such as the usefulness, importance, rank, status, excellence, or intrinsic desirability of something, are not operational definitions of value like the exchange definitions, but they also define value in terms of consequences. If something has high value, it will be judged to be high in usefulness, or importance, or rank, or status, or excellence, or intrinsic desirability. The question remains, What is it about something that has value that makes it useful or important or desirable? Where does value come from? The next sections of this chapter review in some detail different answers to this question. For now it is sufficient to carry on with a working definition that is broad enough to capture the discussions of value in the literature—value as something about an object, action, activity, or event that makes it good or bad to some perceiver. With this working definition, let me turn now to the central issue in this chapter: *Where does value come from?*

VALUE FROM NEED SATISFACTION

At the turn of the 20th century, psychologists with theoretical perspectives ranging from behavioristic to Gestalt to psychodynamic proposed that value comes from need satisfaction. This viewpoint relates to dictionary definitions regarding the *usefulness* of something—its ability to satisfy physical needs or reduce drives or deficiencies, to increase an individual's survival in the world. This perspective on value as usefulness includes what an object or activity affords, such as a chair affording sitting (e.g., Gibson, 1979).

In the classic version of this value-from-need-satisfaction viewpoint, behavior is directed toward the removal of tissue deficits. Drives were manifest in behavior, had physiological correlates, and naturally gave rise to man's desires (see Woodworth, 1918). Value derives from homeostasis and physiological equilibrium (see Weiner, 1972). A striking illustration of this source of value is provided in Woodworth and Schlosberg's (1954) classic textbook. If one feeds an animal for a few days on a diet that is deficient in vitamin B, one creates a biological need for this vitamin. If one then offers the animal a choice between a meal that is rich in vitamin B versus one that lacks vitamin B, the animal will choose the vitamin B-rich meal.

The value-from-need-satisfaction viewpoint has had less influence within social psychology than the other viewpoints on where value comes from that are reviewed next. I discuss later why I would not include the influential consistency models (e.g., balance theory and cognitive dissonance theory) within this viewpoint, although one could arguably do so given the standard reference to "need for consistency." Clearly relevant to the need-satisfaction viewpoint is the research in the attitude literature that attempts to form attitudes through operant or classical conditioning (see Eagly & Chaiken, 1993). Unfortunately, the research on operant conditioning typically used a paradigm that confounded need satisfaction

with other sources of value. For example, the experimenter in most studies approved or disapproved of some class of response by the participant. This procedure not only satisfied a need, such as a need for approval or a need to be accepted, but also potentially created a shared belief about what is desirable or created a standard of excellence for self-regulation, which are alternative sources of value reviewed below. Still, after a careful review of the literature, Eagly and Chaiken (1993) concluded that the traditional conditioning account, which involves need satisfaction, is reasonable as an explanation for the findings in the attitude literature.

The value-from-need-satisfaction viewpoint in the attitude literature is not restricted to studies on conditioning. Evidence for attitudinal effects of mere exposure (i.e., that familiarity leads to liking or preferences) (Titchener, 1910; Zajonc, 1968) and the persuasive impact of fear appeals (Hovland, Janis, & Kelley, 1953; Rogers, 1975) can also be considered relating generally to value from need satisfaction in that anxiety reduction is likely to be involved. In addition, the literature on "message matching" (e.g., Clary, Snyder, Ridge, Miene, & Haugen, 1994) could be considered to reflect a need-satisfaction viewpoint. In their classic functional approach to attitudes, Smith, Bruner, and White (1956) and Katz (1960) proposed that attitudes have different motivational bases that serve different needs, such as the value-expressive function that satisfies the need to affirm desired identities and the social adjustment function that maintains (or disrupts) social relationships.

VALUE FROM SHARED BELIEFS
ABOUT WHAT IS DESIRABLE

An important feature of the value-from-need-satisfaction viewpoint is that value can occur at the biological system level without any need for reflection or beliefs. Moreover, value from this perspective is not restricted to humans. A very different answer to where value comes from has neither of these features—it is restricted to humans and involves beliefs. It relates to the common use of the term "value" that was mentioned earlier—value from people's shared beliefs about general desirable ends. This is typically the meaning of value when people talk about their personal values or the values of others. Although these values are personal in the sense of being internalized, they are acquired within a social context and thus are shared with others (see Williams, 1979). In a clear statement of this viewpoint, Rokeach (1980) describes values as "shared prescriptive or proscriptive beliefs about ideal modes of behavior and end states of existence" (p. 262). Similarly, Merton (1957) points out, "Every social group invariably couples its cultural objectives with regulations, rooted in the mores or institutions, of allowable procedures for moving toward these objectives" (p. 133). The cultural objectives are the "things worth striving for" (Merton, 1957, p. 133)— what has outcome value in the culture. The "allowable procedures" concern the acceptable ways to strive for the worthwhile

things—what has process value in the culture (see also Rokeach, 1979; Schwartz, 1992). This viewpoint, then, concerns *shared beliefs about both desired objectives or end states and desired procedures or means for attaining and maintaining them*. It includes norms about what goals are worth pursuing and what moral principles or standards of conduct one should live by. As captured in the concept of "procedural justice" (e.g., Thibaut & Walker, 1975; Tyler & Lind, 1992), for example, people in many societies value the fairness of decision procedures independent of decision outcomes.

Given the importance of cultural or socialized values in what makes humans a unique kind of animal, it is not surprising that these values have received special attention in the literature on the psychology of values (e.g., Seligman, Olson, & Zanna, 1996), including by such pioneers as Rokeach (e.g., 1973) and Schwartz (e.g., Schwartz & Bilsky, 1987). These shared beliefs about desired states of existence concern both the outcomes of goal pursuit, such as social justice, freedom, and social recognition, and the processes of goal pursuit, such as honesty, logic, and obedience. This viewpoint of value from shared beliefs about what is desirable relates to dictionary definitions regarding some *standard of excellence*. Lewin (1952) noted the special nature of such values. He said that such values influence behavior differently than do goals. For example, people do not attempt to "reach" the value of fairness even though it guides their behavior. He pointed out that such values function to define the valence of an action or activity in a given situation, to define which have positive valence and which have negative valence.

The viewpoint of values as shared beliefs about what is desirable provides a complement to the viewpoint on value from need satisfaction or affordance in its emphasis on social construction of values and potential variability across people as a function of socialization rather than nature. The "sharing" can be considered at more than one level of analysis—at both the broader community or societal level where the values apply to people in general or at the more interpersonal level where significant others apply values specifically to oneself. Values from shared beliefs at the broader societal level are usually treated as social values, whereas the interpersonal level with significant others is usually treated as personal values. The literature on personal values is somewhat different from the literature on social values and is considered in more detail in the next section.

VALUE FROM ACTUAL SELF-RELATION
TO END STATES

Classic cybernetic and control process models treat value in terms of the relation between a current state and some end state functioning as a standard or reference point. Approaching desired end states and avoiding undesired end states have positive value (e.g., Carver & Scheier, 1981, 1990; Miller, Galanter, & Pribram, 1960; Powers, 1973; Wiener, 1948). With respect to human motivation,

the desired and undesired end states that function as guides for self-regulation are typically acquired from interactions with significant others. A person's self-guides begin developmentally as representations of what that person's caretakers would like that person ideally to be (ideals) or believe that person should be (oughts). When that person develops his or her own standpoint, the ideals or oughts for that person from the significant others' standpoint can be shared by that person from his or her own standpoint. This creates a shared reality about desired end states—identified or internalized self-guides. With respect to the desired end states of ideals and oughts, there is considerable evidence that congruencies (or matches) have positive value and discrepancies (or mismatches) have negative value and (Higgins, 1987, 1991, 1998; James, 1890/1948; Moretti & Higgins, 1999; Rogers, 1961).

This viewpoint of value from actual self-relation to end states (desired and undesired) is also found in the literature on social comparison. Comparing yourself with another person who possesses less of a desirable (and self-relevant) attribute than you has positive value, whereas comparing yourself with another person who possesses more of a desirable attribute than you has negative value (e.g., Tesser, 1988). Value from comparing to others also occurs for end states represented by desired and undesired groups such as positive and negative reference groups (Hyman, 1942; Merton & Kitt, 1952). It has also been noted that positive value created by matching desired end states can be achieved through indirect means as well. For example, people who are made to feel insecure or uncertain about whether they have the kind of identity they want to possess, such as being a musician, will engage in activities or present themselves in ways that are traditionally associated with being that kind of person—"symbolic self-completion" (Wicklund & Gollwitzer, 1982). People will also form relationships with other people who possess an attribute they do not have but want to have, such as great beauty, in order to possess the attribute by association—"basking in reflected glory" (Cialdini et al., 1976; Tesser, 1988). People whose personal identity fails to meet their standards of excellence will act to enhance the value of the groups to which they belong in order to possess a social identity that does meet their standards (Tajfel & Turner, 1979).

It should be noted that the viewpoint of value from actual self-relation to end states treats value in two different ways. Like the previous viewpoint on value from shared beliefs about what is desirable, the person's ideal and ought standards of excellence that are shared with significant others are themselves desired end states that function as personal values. But in addition, the congruencies or discrepancies of the actual self from the ideal and ought self-guides themselves have positive or negative value. It is this second aspect of the proposal (i.e., value from actual self-relation to personal standards) that uniquely characterizes this viewpoint. It introduces the notion of value from monitoring one's success or failure in meeting standards or attaining goals—value from the answer to "how am I doing?" (Bandura, 1986; Boldero

& Francis, 2002; Carver & Scheier, 1990; Duval & Wicklund, 1972). When people speak of their life as having value (or not), it is typically in reference to this kind of value. Monitoring one's actual self-relation to end states is an important type of self-evaluative process. There are other processes of evaluating oneself that are also a source of value. They differ from the value created by actual self-relations to end states in their requiring an inference to identify the value. Such evaluative inferences as a source of value are considered next.

VALUE FROM EVALUATIVE INFERENCE

The value created from monitoring discrepancies or congruencies in relation to end states does not require an inference to create value. An actual-self congruency with ideals and oughts *is* good. An actual-self discrepancy with ideals and oughts *is* bad. But there are times when people do infer what the value of some activity or object is to them. Perhaps the most influential theory about such evaluative self-inferences in social psychology was proposed by Bem (1965, 1967) in his self-perception theory. A basic assumption of self-perception theory is that people function like behavioral scientists engaged in hypothesis testing (cf. Heider, 1958; Jones & Davis, 1965; Schachter & Singer, 1962). The critical proposal is that people make inferences about themselves in the same way that an uninterested outside observer would make inferences—based on observable evidence. They observe their own behavior and test hypotheses about its meaning or significance. When people engage in an activity, for example, one hypothesis is that they chose to do that activity because they value it. In Bem's terms, adapted from Skinner (1953, 1957), a person hypothesizes that the force to do the activity came from within and the behavior was emitted spontaneously (a tact). An alternative hypothesis is that the behavior was demanded from without and was elicited by outside forces or pressures (a mand). If there is strong support for the second hypothesis, such as a person was promised a large reward to do the behavior, then the person is less likely to infer that he or she values the activity. Classic studies on bolstering and undermining the value of an activity support the Bemian perspective on value from evaluative inference (see Kruglanski, 1975; Lepper, Greene, & Nisbett, 1973; Salancik & Conway, 1975).

The viewpoint of value from evaluative inference is broader than just the Bemian perspective. For Bem, the evidence must be publicly observable. But Andersen (1984) points out that the notion of evidence to test hypotheses and draw inferences about what one values need not be restricted to publicly observable evidence. Indeed, she reports that people are likely to use what they were thinking or feeling about some thing as evidence to infer its value to them more than their behaviors regarding it (cf. Schwarz & Clore, 1988). An especially broad form of evidence is whatever the current stimulus reminds one of from one's stored knowledge and past experiences. People use analogies, representa-

tiveness, and familiarity regarding their knowledge and experience of the world to assign meaning to a current object or event (e.g., Gilovich, 1981; Holyoak & Thagard, 1997; Tversky & Kahneman, 1974). Thus, for example, they could infer the value to them of some new activity from the value of a previously experienced activity to which the new activity is similar.

It is notable that the Bemian logic used to make an evaluative inference, like that originally proposed in Heider's (1958) theory of people's lay reasoning, is relatively simple—the likelihood of drawing an inference from your behavior about the value of something to you decreases as the evidence of external forces that could have produced the behavior increases. There are other more complex kinds of logical reasoning that can be used in the inferential process that creates value, such as Bayesian logic (e.g., Trope, 1986a) or the information gain logic that is proposed in Higgins and Trope's activity engagement theory (see Higgins, Trope, & Kwon, 1999). Of special interest historically is the reasoning postulated to underlie ethical value. One of the most influential models of inferential reasoning as a source of ethical value is that proposed by Kohlberg (1969, 1976) based on the earlier work of Piaget (1932/1965). According to this model, the moral value of an action depends on the reasons for doing or not doing it, and the reasons vary in their cognitive level of development. More generally, philosophers from Aristotle to Kant and Marx have proposed that determinations of what is morally good or bad should be based, and to some extent *are* based, on logical reflection using criteria of justice that are provided by religious and political authorities (cf. Haidt, 2001). Psychological research on the criteria people actually use to infer the ethical value of some action has found that people use multiple criteria and engage in trade-offs among them, such as trade-offs between equity and equality for distributive actions (Tyler & Smith, 1998).

Before closing this section on value from evaluative inferences, one other important factor in evaluative inferences must be mentioned—the role of context in providing factual frames of reference. People do not infer the value of something in a vacuum. They consider the value of something in relation to whatever factual standards are currently available (or accessible) to them, which can vary as a function of context (Higgins, Strauman, & Klein, 1986). The value of something can be assimilated toward or contrasted away from the context of current alternatives (Helson, 1964; Higgins & Stangor, 1988; Sherif & Hovland, 1961). It can also vary depending on which mental account the context suggests is appropriate for calculating its value (Thaler, 1999). It can increase or decrease in intensity depending on what the context suggests is normal (Kahneman & Miller, 1986) or could have happened instead (Kahneman & Tversky, 1982; Markus & Nurius, 1986; Roese, 1997). Standards of excellence that are desired end states have been discussed earlier. What is notable here is that some standards are factual states rather than desirable states, but they nonetheless influence value inferences. Moreover, whereas the standards of excellence that a person uses tend to be rela-

tively chronic, the factual standards a person uses often vary by context.

VALUE FROM EXPERIENCE

It is curious that value from experience is not explicitly mentioned in the dictionary definitions of value listed earlier. Nonetheless, the viewpoint of value from experience has a long history and has received considerable attention by psychologists. In philosophy, value from experience has a checkered history, especially with respect to ethical or moral value. Generally speaking, the dominant philosophical viewpoint has been more compatible with the value from evaluative inference viewpoint that was discussed in the previous section, but with the specific emphasis being more on using reason and reflection to create an objective basis for determining what is good or bad (Haidt, 2001; Williams, 1985). The tension between a reason versus an experiential account of ethical value is captured in Plato's (4th century B.C./1949) myth, described in *Timaeus*, of the gods beginning by creating human heads filled with reason. Because the heads alone could not move around in the world, the gods were forced to add bodies to the heads to permit movement. Emotions resided in the bodies. The moral of the story was that people were left with heads (i.e., reason), that must struggle with their passionate bodies to get them to behave ethically (and rationally). This moral was well accepted by most influential philosophers (e.g., Aristotle and Kant) and is embodied in Freud's (1923/1961) classic conflict between the reasonable ego and the passionate id. For centuries emotions have been connected with desires, desires with the potential for sin, and reason as the defense against sin (Haidt, 2001).

Plato's myth can be interpreted in another way, however. After all, the gods are not crazy and they understood that the heads alone could take no action. The problem of bridging the gap between knowing and doing has been well understood by mere mortal psychologists as well. There is the famous story of Edwin Guthrie, the "practical behaviorist," making a joke on this point at the expense of Edward Tolman, one of the founders of cognitive psychology. Tolman's influential theory of behavior emphasized hypotheses and expectations and Guthrie joked that Tolman's rats would never leave the start box to get to the goal because they would be "buried deep in thought." Reason alone does not produce movement. Something must provide the movement, must bridge the gap. The gods' solution of adding emotional bodies to the heads was an inspired solution. I propose, then, that Plato's myth has the useful moral that, without adding value from experience, value from reason cannot account for motivation and action. Value from experience is critical. I return to this point later after reviewing the different kinds of value from experience that have been proposed in the literature. At this point I should add that although somewhat in the minority, there have certainly been highly distinguished philosophers who have argued that experience is critical to ethical value

(e.g., Bentham, 1781/1988; Hume, 1777/1975; Smith, 1759/1997). Their arguments are considered next.

Hedonic Pleasure/Pain Experiences

The term "hedonic," which derives from the Greek term for "sweet," means relating to or characterized by pleasure (*Webster's Ninth New Collegiate Dictionary*, 1989, p. 561). When considering value experiences, hedonic experiences have received the most attention. From the time of the Greeks, hedonic experiences have been linked to the classic motivational principle that people approach pleasure and avoid pain. Our understanding of the subjective experience of pleasure and pain has become much more sophisticated (e.g., Kahneman & Tversky, 1979), and interest in hedonic experiences has never been greater (see Kahneman, Diener, & Schwarz, 1999). Perhaps the most influential early statement on the importance of hedonic experiences to value–both ethical and nonethical–was given by Jeremy Bentham (1781/1988) who stated: "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne" (p. 1). Bentham (1781/1988) defined the principle of *utility* as "that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question" (p. 2). Utility referred both to the property of producing pleasure and avoiding pain.

Kahneman (2000a) points out that the concept of utility has had two different meanings. One is like the primary dictionary definitions in its being an operational (behavioristic) definition–utility is inferred from observed choices. Kahneman refers to this concept of utility as decision utility. The second meaning reflects Bentham's perspective on utility as experiences of pleasure and pain and is called experienced utility by Kahneman. Kahneman argues that the best way to measure actual experienced utility is by moment-based methods where experienced utility of an episode is derived from real-time measures of the pleasure and pain that an individual experienced during that episode. But people also use their memory of pleasure and pain experiences to retrospectively evaluate past episodes, which Kahneman calls remembered utility. Redelmeier and Kahneman (1996) found, for the case of pain, that remembered utility was less influenced by the duration of the pain episode than by the "peak-end rule," which is the average of the most intense level of pain over the episode and the level of pain at the end of the episode. This indicates that remembered utility is not an accurate reflection of actual experienced utility, but people use what they remember to make choices. That is, decisions concerning the value of things (decision utility) can be determined more by retrospective value experiences based on memory (remembered utility) than by what pleasures or

pains were actually experienced at the time of the episode (experienced utility). Of course, remembered utility *is* the subjective hedonic experience at the time of the decision.

The role of peak experiences in hedonic value experiences may not be restricted to humans. Influential animal learning theorists argued that animals' emotional responses to stimuli played a critical role in their learning and performance (e.g., Miller, 1963; Mowrer, 1960; Spence, 1958). Spence (1958) suggested that animals differ characteristically in the magnitude of their emotional response to the same intensity of stressful stimulation and that the highly emotional animals would show a higher level of performance in aversive conditioning. In other words, higher emotionality would be related to a higher-level peak experience for the painful episode, which would produce a more negative remembered episode, which would enhance performance.

Hedonic experiences have received most attention in the voluminous literature on emotions and affect. Early on, Spinoza (1677/1986) proposed that all emotions can be reduced to some form of pleasure and pain. Since then it has been suggested that the primary function of emotional experiences is to signal or provide feedback about self-regulatory success or failure (e.g., Frijda, 1986; Mandler, 1984; Simon, 1967). Both appraisal and circumplex models propose a basic dimension that distinguishes between pleasant and painful emotions (e.g., Diener & Emmons, 1984; Feldman Barrett & Russell, 1998; Frijda, Kuipers, & ter Schure, 1989; Green, Goldman, & Salovey, 1993; Larsen & Diener, 1985; Ortony, Clore, & Collins, 1988; Roseman, 1984; Russell, 1980; Scherer, 1988; Smith & Ellsworth, 1985; see also Schlosberg, 1952; Wundt, 1896/1999).

As evident in the attitude and affect literatures (Eagly & Chaiken, 1993; Schwarz & Clore, 1996), the bulk of research in social psychology that is concerned with value experiences has emphasized basic hedonic experiences, such as basic distinctions between good and bad moods or between liking and disliking something. Influential theories and findings in decision science have also emphasized basic hedonic experiences, such as the pleasure of gains and the pain of losses (e.g., Kahneman & Tversky, 1979) or the pleasure of hope and the pain of fear (e.g., Lopes, 1987). This is not to say that there have been no theories and research in these literatures that have distinguished among types of pleasure and among types of pain. Self-discrepancy theory (Higgins, 1987), for example, distinguishes between the pleasures of ideal versus ought success or congruency (cheerfulness vs. quiescence, respectively) and between the pains of ideal versus ought failure or discrepancy (dejection vs. agitation, respectively). More generally, the emotion and affect literature distinguishes among many different types of pleasure and different types of pain based on the activation level of the experience, the target of the emotion (e.g., reaction to events versus agents), and other factors (e.g., Ortony et al., 1988; Roseman, 1984; Russell, 1980). But beyond identifying different types of hedonic experiences, it is important to identify types of value experi-

ences other than hedonic experiences. The remainder of this section reviews such types of experience.

Moral or Ethical Experiences

As mentioned earlier, for Bentham (1781/1988) the standard of right and wrong, what determined what one ought to do, the basis for approving or disapproving an action, was the pleasures and pains associated with the action. His influential system of ethics known as *utilitarianism* considered the amount of pain and pleasure an action produced for all interested parties in a community—the greatest happiness for the greatest number. Two things should be noted about Bentham's perspective. First, it can be used to distinguish ethical from nonethical experiences in that nonethical can concern only an agent's happiness whereas ethical must concern the happiness of all interested parties. Second, all the experiences being considered are hedonic experiences.

For other early moral philosophers moral experiences were broader than just hedonic experiences. Adam Smith (1759/1957), for example, described "the first perceptions of right and wrong" (p. 182). He stated that the first perceptions were *not* derived from reason but from immediate sense and feeling (see also Hume, 1777/1975). Although Smith did not agree with philosophers such as Francis Hutcheson that there was a specific "moral sense" that had its own peculiar nature, distinct from all other faculties or sentiments, he did believe that there was a specific process underlying moral sentiment—sympathy and empathy with what others were feeling. This included feelings of approval when our sentiment regarding the conduct of a third person coincides with the sentiment of another person toward the same conduct and feelings of disapproval when the sentiments are in opposition. Moral approval, according to Smith (1759/1957) does include a utilitarian component, but it also includes sympathizing with the motives of the agent and empathizing with the gratitude of those benefiting from what the agent did. He explicitly distinguished his perspective from the perspective that

> places virtue in utility, and accounts for the pleasure with which the spectator surveys the utility of any quality from sympathy with the happiness of those who are affected by it. This sympathy is different both from that by which we enter into the motives of the agent, and from that by which we go along with the gratitude of the persons who are benefited by his actions. It is the same principle with that by which we approve of a well contrived machine. But no machine can be the object of either of those last two mentioned sympathies. (pp. 192–193)

It is not clear how Smith's kind of moral feelings relates to what Haidt (2001) describes as moral intuition. According to Haidt (2001), a moral intuition involves the sudden appearance in consciousness of a moral judgment, including an affective valence. He does relate moral intuition to a process akin to aesthetic judgment that Hutcheson and Hume talked about—a moral sense in which one sees or hears about a social event and in-

stantly feels approval or disapproval. What is clear is that the moral experience from a moral intuition is *not* equivalent to a hedonic experience. It is more than feeling pleasure and pain. It is feeling approval or disapproval. As stated by a modern moral philosopher, Peacocke (2004): "What, additionally, is required for an emotion to be a moral emotion is that the positive or negative emotion, such as the delight or the indignation, be delight or indignation at a state of affairs that is (or is thought by the person having the emotion to be) morally good or morally bad respectively" (p. 254). That is, hedonic pleasure or pain is not sufficient. Moral experience includes a feeling of approval or disapproval. This position is also consistent with the dictionary defines "ethical" as "involving or expressing moral approval or disapproval" (*Webster's Ninth New Collegiate Dictionary*, 1989, p. 427).

Finally, it should be noted that popular culture for centuries has distinguished hedonic from ethical concerns in such maxims as "It is not enough to do good; one must do it in the right way," "What counts is not whether you win or lose, but how you play the game, "The ends don't justify the means," and "Never good through evil." It is notable that the ethical tends to be associated with the process of goal pursuit rather than the outcomes of goal pursuit. However, goal pursuit outcomes also vary in how ethical they are, and goal pursuit processes also vary in how hedonic they are. Thus, the critical distinction is between hedonic and ethical concerns, and these different concerns produce distinct value experiences.

Fit Experiences

When people "feel right" or "feel wrong" about something based on feelings of approval or disapproval, respectively, it can be properly considered a moral response. There is recent evidence, however, that there is another self-regulatory variable that can make people feel right or wrong about something—*regulatory fit*. People experience regulatory fit when their goal orientation is sustained (vs. disrupted) by the manner in which they pursue the goal (Higgins, 2000). For example, some students working to attain an "A" in a course are oriented toward the "A" as an accomplishment or aspiration (a promotion focus) whereas others are oriented toward the "A" as a responsibility or as security (a prevention focus). Some students read material beyond the assigned readings as a way to attain the "A" (an eager strategy) whereas others are careful to fulfill all course requirements (a vigilant strategy). An eager strategy fits a promotion focus better than a prevention focus, whereas the reverse is true for a vigilant strategy (see Higgins, 1997).

Regulatory fit experiences differ fundamentally from hedonic and ethical experiences because they arise from a different aspect of self-regulation. Fit experiences occur during the process of goal pursuit and are independent of the outcomes or consequences of the goal pursuit (both ethical and nonethical). This is not to say that hedonic and ethical experiences cannot arise as well during the process of goal pursuit, but they do so in relation to other goals that are or are not attained. For hedonic

experiences the other goals could be to have fun during the goal pursuit process ("Getting there is half the fun") or to expend as little effort as possible. For ethical experiences the other goals could be to pursue the goal in a way that would meet with others' approval or gratitude. In contrast to experiences from goal attainment, fit experiences arise from the relation between the manner of goal pursuit and the actor's current orientation toward the goal pursuit. Given one's current orientation to the goal pursuit, it is an experience of the suitability of one's manner of pursuing the goal. Unlike hedonic and ethical experiences, fit is not an experience of attaining some other goal. Perhaps this is why justifying one's choice of goal pursuit strategy by referring to hedonic or ethical experiences seems reasonable (e.g., "it was fun to do it that way" or "it was the socially proper way to do it") but attempting to justify it by referring to fit experiences does not seem reasonable ("it was the way to do it that happened to fit my motivational state at the time").

People experience regulatory fit when they pursue a goal in a manner that suits their current orientation toward the goal, and fit makes them "feel right" about what they are doing. Regulatory fit theory proposes that the "feel right" fit experience (Higgins, 2000; Higgins, Idson, Freitas, Spiegel, & Molden, 2003) is an experience that can influence subsequent evaluations of objects and actions. Regulatory fit increases the strength of people's engagement with what they are doing, thereby intensifying their reaction to what they are doing. When there is regulatory fit, people "feel right" about their positive reaction to something and "feel right" about their negative reaction to something (Cesario, Grant, & Higgins, 2004). Transfer of the regulatory fit experience can influence both utility and ethical evaluations.

In one set of studies, Higgins and colleagues (2003) examined the effect of regulatory fit on utility evaluations. Participants' chronic promotion and prevention orientations were measured prior to the experimental session. When the participants arrived for the experiment they were told that, over and above their usual payment for participating, they could choose between a coffee mug and a pen as a gift. (Pretesting indicated that the mug was clearly preferred.) The strategic manner of making the decision was manipulated through framing. Half of the participants were told to think about what they would gain by choosing the mug or the pen (an eager strategy), and the other half were told to think about what they would lose by not choosing the mug or the pen (a vigilant strategy). As expected, almost all participants chose the coffee mug. These participants were then asked either to assess the price of the chosen mug or to offer a price to buy it. Participants in the fit conditions (promotion/eager; prevention/vigilant) gave a much higher price for the mug than participants in the nonfit conditions (40–60% higher). Importantly, this fit effect on the money offered to buy the mug was independent of the participants' reports of their pleasure/pain feelings after making their decision. The fit effect was also independent of their perception of the efficiency (ease) and effectiveness (instrumentality) of the means they used to make their choice. Finally, another study demonstrated that regulatory fit created in one situation (from listing strategies that did or did not fit participants' personal goals) can influence later evaluations of a totally separate object (the "goodnaturedness" of pictured dogs).

Another set of studies by Camacho, Higgins, and Luger (2003) examined the effect of regulatory fit on ethical evaluations. One study, for example, had participants think back to a time in their lives when they had a conflict with an authority figure who resolved the conflict. The participants were asked to recall different kinds of resolution. For example, some participants recalled a resolution where the other person encouraged them to succeed (the pleasant/eager condition), whereas other participants remembered a resolution where the other person safeguarded them against anything that might go wrong (the pleasant/vigilant condition). Independent of whether the manner of resolution was itself pleasant or painful, and independent of their own pleasant or painful mood while making their judgments, participants judged the resolution to be more "right" if they were predominant promotion and the conflict was resolved in an eager manner or they were predominant prevention and the conflict was resolved in a vigilant manner. In another study by Camacho and colleagues (2003), a proposed policy for a new after-school program was judged to be more "morally right" when the manner of carrying out the program sounded eager than when it sounded vigilant for participants with a predominant promotion focus, but the opposite was true for participants with a predominant prevention focus.

As described earlier, there is evidence from studies by both Higgins and colleagues (2003) and Camacho and colleagues (2003) that "feeling right" from regulatory fit is more than just feeling pleasure or pain. Additional evidence of the independence of these value experiences was found by Cesario and colleagues (2004) in a study on persuasion. Similar to one of the Camacho and colleagues studies described earlier, Cesario and colleagues measured participants' chronic regulatory focus and manipulated whether a persuasive message advocating a new policy used either eager or vigilant means framing. The participants' hedonic mood after receiving the message was also measured. The study found that higher pleasant mood predicted greater message effectiveness in changing opinions in the advocated direction. At the same time, fit had an independent effect of increasing message effectiveness. Thus, regulatory fit and hedonic pleasure/pain mood each had independent effects on persuasion. In addition, the manipulation of regulatory fit did not affect participants' hedonic mood.

Because regulatory fit in this study by Cesario and colleagues (2004) involves the relation between participants' orientation and message framing, it is possible that some kind of "message matching" was involved. This could mean that value from need satisfaction might have contributed to the effects. Cesario and colleagues controlled for this possibility in another study, which also included a measure of hedonic mood. The study tested whether participants would be more persuaded by an

identical message if they had, versus had not, experienced regulatory fit prior to even receiving the message. Using an "unrelated studies" paradigm, participants in the first phase of the experimental session were asked to list either two promotion goals (i.e., listed two of their personal hopes or aspirations) or two prevention goals (i.e., listed two of their duties or obligations). Then they listed for each of their goals either eager means of pursuit (i.e., strategies they could use to make sure everything goes right) or vigilant means (strategies they could use to avoid anything that could go wrong). Promotion goals/eager means and prevention goals/vigilant means were the regulatory fit conditions. Following the goal/strategy listing, all participants received the identical persuasive message. Once again, the manipulation of regulatory fit did not affect participants' hedonic mood, and both pleasant mood and regulatory fit had independent effects on increasing the effectiveness of the message.

Together, these studies demonstrate that the regulatory fit, "feel right" experience is a distinct type of experience that can influence other types of value experience. It should be noted that the transfer is *not* a main effect of fit increasing the positive value of some action or object, unlike what one might expect from pleasant hedonic experiences for example. Rather, fit makes people *"feel right" about their response to an object or event*. This increases the strength of that response, whether it is positive or negative. Another study by Cesario and colleagues (2004) demonstrated this fact. They used a persuasive message of moderate persuasive strength such that participants varied in their positive or negative reactions to the message as measured by a standard thought-listing technique. They found that the fit effect on increasing attitude change occurred only for participants who had positive thoughts about the message. For participants who had negative thoughts about the message, fit had the opposite effect—it *decreased* attitude change. Thus, the "feel right" value experience from regulatory fit is a distinct experience of intensifying one's reaction to what one is doing, whatever that reaction happens to be.

As mentioned earlier, the regulatory fit experience differs from hedonic and ethical experiences in being an experience within the process of self-regulation that is independent of the ultimate, instrumental value of the goal pursuit. Regulatory fit is an experience produced by engaging in the goal pursuit activity itself that influences value. There are two other general types of value-related experience produced during activity engagement—understanding experiences and agency experiences. To understand the world and to act on it effectively can be thought of as two basic human motives (Pittman & Higgins, 2006; Stevens & Fiske, 1995; White, 1959), and the mechanisms by which people try to understand the world and act on it effectively have received considerable attention by psychologists generally and social psychologists especially (Anderson, Krull, & Weiner, 1996; Gollwitzer & Moskowitz, 1996; Heider, 1958; Snyder & Cantor, 1998). What is of interest in this chapter is the nature of understanding and agency as experiences that have value. These two general types of experience are considered next.

Understanding Experiences

People want to know how things work and they go about trying to figure things out and cope with the unknown (Kruglanski, 1989; Sorrentino & Roney, 2000). In social psychology, people's experiences from success or failure in making sense of the world have been examined most extensively by the various *consistency models* (Abelson et al., 1968), such as Heider's (1958) balance theory, McGuire's (1968) syllogistic model, Abelson and Rosenberg's (1995) psychologic model, Osgood and Tannenbaum's (1955) congruity principle, and, most influential of all, Festinger's (1957) cognitive dissonance theory.

Although multiple motives can play a role in consistency models, the central motive is to make sense of the world and one's relations to it (Fiske, 2004; Kruglanski, 1989). Consistencies and inconsistencies of various kinds produce different understanding experiences. Heider (1958), for example, describes an imbalanced state as leaving people with a feeling of disturbance, feeling that the factors in the situation "do not add up" (p. 180). The feeling of disturbance is only relieved when there is a change within the situation to achieve balance. Importantly, Heider (1958) also notes that the tension produced by imbalance can have a pleasing effect and balance can be unpleasantly boring. Thus, it is not the hedonic nature of balance or imbalance that is critical to the experience. Rather, it is the feeling that things do or do not add up—an understanding experience. Even when imbalance is hedonically pleasing, feeling that things "do not add up" is a value experience that motivates changing the situation, moving away from the current psychological situation to a new state.

According to Festinger (1957), " the human organism tries to establish internal harmony, consistency, or congruity among his opinions, attitudes, knowledge, and values" (p. 260). When people fail to do so, they experience dissonance, which gives rise to pressures to reduce that dissonance. Importantly, Festinger states: "In short, I am proposing that dissonance, that is, the existence of nonfitting relations among cognitions, is a motivating factor in its own right" (p. 3). What this suggests is that the motivational significance of dissonance does not derive from hedonic experiences alone but from it being a distinct type of value experience relating to nonfitting relations among cognitions. Indeed, part of the fascination with cognitive dissonance theory and other consistency models was their counterintuitive predictions that cognitive consistency as one kind of understanding experience had sufficient value that people would choose to engage in an unpleasant activity in order to achieve it, such as choosing to receive painful electric shocks (Comer & Laird, 1975) or agreeing to a request to install a large ugly sign in front of one's house (Freedman & Fraser, 1966).

The understanding experience of cognitive consistency is not the only kind of understanding experience that has value to people. There is also the value that derives from reducing uncertainty and drawing conclusions regarding some area of knowledge. Moreover, there is

evidence that these understanding experiences also differ from hedonic experiences. When it comes to the value of understanding, it is not the case that what matters is simply the pleasure or pain that results from the knowledge that is the product of the understanding. Sorrentino and Roney (2000) give the example of two mathematicians who solve a complex problem about the purpose of the universe and prove that it equals zero. The knowledge that is the product of their new understanding is itself hedonically painful but the understanding experience itself is the feeling of resolving uncertainty about the purpose of the universe. This feeling, like the feeling of discovery more generally, is independent of the hedonic consequences of the knowledge produced by the understanding. Similarly, some people value a feeling of certainty even when it concerns unpleasant knowledge about themselves (Sorrentino & Roney, 2000; Swann, Wenzlaff, Krull, & Pelham, 1992). Kruglanski's work on epistemic motivation (Kruglanski, 1989; Kruglanski & Webster, 1996) distinguishes between the need for specific closure that is concerned with whether the contents of the final knowledge are pleasant or unpleasant versus the need for nonspecific closure that is concerned only with arriving at a conclusion, having the experience of confidently knowing something, regardless of whether that something is pleasant or unpleasant.

As mentioned earlier, most emotion theorists, from Spinoza (1677/1986) to modern scholars, consider all emotions to be, in one form or another, hedonic experiences of pleasure and pain. It is interesting in this regard that Ortony and colleagues (1988) explicitly reject "surprise" as an emotion because it "can arise in the absence of a valenced reaction" (p. 32). They consider it to be, instead, a variable that can modulate the intensity of other emotions. The emphasis here is on surprise as a kind of understanding experience, where something happens that is different from what one thought one knew. Consistent with this position, surprise can be accompanied by either a pleasant or a painful experience. Its distinctive nature derives not from hedonic valence but from its relation to understanding. This is true of feelings like "wonder" and "interest" as well. Although these experiences are more often associated with pleasure, they are certainly not always pleasant. What distinguishes them is their relation to understanding—paying attention to something new to one's experience, to something that can increase one's knowledge. Even when understanding experiences are more clearly pleasant like "curiosity" or unpleasant like "doubt" or "uncertainty," their distinct character is their relation to understanding. Indeed, despite having a different valence, "curiosity" and "doubt" are both included in the definition of "wonder" (*Webster's Ninth New Collegiate Dictionary*, 1989, p. 1357).

Agentic Experiences

People not only want to understand how things work. They also want to make things work, to have a sense of control over what happens in the world (Bandura, 1977, 1982; Deci, 1975; Pittman & Higgins, 2006; Stevens &

Fiske, 1995; Weary, Gleicher, & Marsh, 1993). Two different kinds of agentic experiences have been described in the psychological literature: (1) experiencing oneself as being the agent of what happened versus something or someone else as the agent—feelings of personal control, autonomy or self-determination; and (2) experiencing the agency of bringing to bear one's resources to make something happen—feelings of effort, effectiveness or mastery. These two agentic experiences are related to the dictionary definitions (*Webster's Ninth New Collegiate Dictionary*, 1989) of "agent" as something that produces an effect, the efficient cause (p. 64) and "agency" as the capacity, condition, or state of acting or of exerting power (p. 63).

It has also been recognized in the literature that value from agentic experiences is *not* the same as value from either need satisfaction or hedonic value. Woodworth (1940) decades ago distinguished agentic value from need satisfaction:

> To some thinkers on these matters it appears self-evident that dealing with the environment occurs only in the service of the organic needs for food, etc. They say that the muscles and sense organs have evolved simply as tools for the better securing of food and other organic necessities, and for reproducing the race. Only the organic needs, on this view, are entitled to rank as primary drives; all activity dealing with the environment is secondary. The facts of evolution do not compel us to adopt this view, for motility and responsiveness to the environment are present way down to the bottom of the scale of animal life. There is no more reason for saying that the muscles exist for the purpose of obtaining food than for saying that food is needed to supply energy for the muscles. . . . (p. 374)

In an influential paper, White (1959) rejected the need satisfaction perspective on value experiences and emphasized instead the experience of dealing effectively in interactions with the environment, including evidence of humans and other animals being interested in exploring and manipulating their environment (see also Woodworth & Schlosberg, 1954). Recently, Kahneman (1999) distinguished agentic value from hedonic value when discussing the agentic "flow" experience described by Csikszentmihalyi (1990): "In particular, the measure of instant utility must allow for states of 'flow' in which one is so involved in an experience or activity that hedonic value fades into the background of experience" (p. 6).

Let me begin with the first kind of agentic experience from experiencing oneself, versus someone or something else, as being the causal agent. Perceiving an action in terms of actor as agent versus nonactor as agent is a major distinction in the person perception literature (Heider, 1958). Emotional experiences themselves have been distinguished in terms of whether the focus is on self as agent or other as agent (e.g., Ortony et al., 1988; Weiner, 1985). It is not surprising, then, that self versus other as agent underlies an important value experience. In an early statement of the value of experiencing personal agency, deCharmes (1968) stated: "Man strives to be a causal agent, to be the primary locus of causation

for, or the origin of, his behavior; he strives for personal causation" (p. 269). Deci (1975) also proposed that people engage in activities in order to feel a sense of self-determination. Deci and Ryan (1985) report that when people feel that they are in control of what they are doing, rather than being pressured by others or circumstances to do something, they enjoy more what they are doing, are more interested in doing it again, and generally experience a more positive emotional tone.

Value from the experience of personal agency is distinct from simple hedonic pleasure. This is illustrated by the fact that interest in an activity can be undermined by adding a pleasant reward for doing it if that reward changes the agentic experience of the activity from self-determined/endogenous to other-determined/exogenous (Deci, 1975; Krulganski, 1975). It should be noted as well that experiencing others as having influence on one's activity engagement will not always undermine the value of that activity. It will depend on whether the other's influence is experienced as "not me" versus "we." For example, undermining will not occur if young children or individuals in a collectivist or interdependent culture or communal relationship (Clark & Mills, 1979; Markus & Kitayama, 1991; Triandis, 1989) experience the influence of others (peers, parents, or trusted others more generally) as part of their own self-regulation. There is evidence consistent with this conclusion (Costanzo & Dix, 1983; Iyengar & Lepper, 1999). Thus, what is critical is not whether one experiences others as having an influence on one's activity engagement but whether that influence undermines one's own experience of personal agency.

Let me turn now to the second kind of agentic experience mentioned above in which people experience the agency of bringing to bear their resources to make something happen and to overcoming resistance through their efforts. The importance of this agentic experience has also been recognized for a long time, as evident in the following statement from James (1890/1948):

> Of course we measure ourselves by many standards. Our strength and our intelligence, our wealth and even our good luck, are things which warm our heart and make us feel ourselves a match for life. But deeper than all such things, and able to suffice unto itself without them, is the sense of the amount of effort which we can put forth . . . the effort seems to belong to an altogether different realm, as if it were the substantive thing which we *are* . . . (p. 458)

One important form of interacting with the environment occurs when people offer *resistance* or opposition to interfering forces, pressures, and circumstances with adverse significance for them. Woodworth (1940), for example, stated that people and other animals exert considerable resistance to environmental forces on them in order to maintain a degree of independence. They resist wind that is trying to blow them over and gravity that is trying to make them fall. They have an active give-and-take relation with the environment and value in the form of interests "springs from the individual's ability to deal effectively with some phase of the environment" (p. 396).

Value creation from resistance is central to social psychological research on reactance (Brehm, 1966; Brehm & Brehm, 1981; Wicklund, 1974). When people feel that their freedom to make their own choice on some matter is being taken away by outside pressure, they resist that pressure and attempt to restore their freedom. This resistance can create value (e.g., Brehm, Stires, Sensenig, & Shaban, 1966). Placing a barrier to attaining a desired object can also create resistance to losing freedom of access that increases the value of the object. Wright (1937), for example, reports an intriguing series of studies in which access to attaining food was made more difficult by adding a barrier. In one study, the experimenter arranged desserts on a serving table at various distances from where the waitresses stood to pick them up to serve. As long as the distance was not too great, the desserts further away were selected more than those that were closer. Another study found that kindergarten children chose a more distant candy with a wire sieve over it instead of one closer with no sieve.

In addition to value creation from the experience of resistance as opposition to external interfering forces, there is also value creation from the experience of overcoming personal resistance to doing something. When there are known adverse consequences of engaging in an activity, there is a natural resistance to doing the activity. When people overcome this internal resistance by freely choosing to engage in the activity, knowing its adverse consequences, they experience *commitment* to and value in the activity (Brickman, 1987). Great figures in psychology, including Sigmund Freud, Kurt Lewin, and Jean Piaget, have recognized that overcoming internal resisting forces is a special kind of agentic experience that relates to psychological commitment and "will" (Brickman, 1987; Deci, 1980).

There is evidence in the psychological literature that when people and other animals face adverse or difficult circumstances when engaging in some activity and continue the activity despite knowing those circumstances, this agentic experience of overcoming personal resistance increases the value of that activity. In a study by Lewis (1964), for example, rats had to displace either a heavy weight (high effort) or a minimal weight (low effort) in order to get to Rice Krispies as a food reward. Then all the rats were placed in a straight maze and allowed to run freely to the goal area (i.e., no weight to displace) where they were rewarded each time. The measure of the attractiveness of eating Rice Krispies was running speed to attain them, and the "high effort" rats ran faster (see also Lewis, 1965). Similarly, Carder and Berkowitz (1970) found that rats preferred pressing a lever to attain pellets than getting the same pellets freely available in a dish near the lever. Indeed, sometimes a rat would push the food-filled dish away from the food magazine to push the lever for a pellet (see also Jensen, 1963).

In the previous section I discussed how cognitive dissonance theory, like other cognitive consistency models, concerns value from an understanding experience. It should be noted, however, that this does not mean that cognitive dissonance only involves an understanding experience. Indeed, the cognitive dissonance literature has

clearly identified agentic experiences as being involved as well. Both the personal control kind of agentic experience (agent) and the overcoming personal resistance kind of agentic experience (agency) have been mentioned as being critical to creating the dissonant state. Some dissonance experts have emphasized the importance for dissonance induction that the actors' experience "free choice" or autonomy or personal responsibility for their decisions (e.g., Brehm & Cohen, 1962; Cooper & Fazio, 1984; Zimbardo, 1969). Others have emphasized the importance for dissonance induction that, in making their decisions, the actors experience personal resistance (which they overcome) from the negative features of the decision (e.g., Brickman, 1987; Wicklund & Brehm, 1976). Brickman (1987), for example, states that the effort involved in the dissonance studies creates value for the activity by increasing commitment to it. The effort occurs because the participants freely choose to do some activity that has negative consequences. The negative consequences produce resistance to do the activity. The effort involved when freely choosing to overcome this resistance increases commitment to the activity and increases its value.

As a final note, the case of cognitive dissonance illustrates how a phenomenon can involve both understanding and agency experiences. This is not surprising, perhaps, given that "thinking is for doing" (Fiske, 1992; James 1890/1948). The example of cognitive dissonance is complicated by the fact that for Festinger (1957) it was the understanding experience that was critical whereas for Brickman (1987) it was the agency experience that was critical. More directly in line with "thinking is for doing" is Pittman's theory of control deprivation (Pittman & D'Agostino, 1989; Pittman & Pittman, 1980). This theory proposes that experiencing an inability to control is disturbing because it calls into question one's understanding of how the world works, thereby undermining feelings of effectance or agency. Pittman and his colleagues find that following control deprivation, the value of accurate understanding increases so that people process subsequent information more thoroughly. The theory and research on control deprivation nicely illustrates how agency and understanding experiences can combine. Conflict between self-guides provides another example of how agency and understanding experiences combine. Van Hook and Higgins (1988) found that when individuals have discrepancies or conflict between their ideal and ought self-guides, they report feeling confused, unsure of themselves, and distractible. A discrepancy between two self-guides, such as what one ideally wants to be and what one's mother believes one ought to be, is both a problem of understanding who one is supposed to be and a problem of planning effective action.

Value from Experience versus Need Satisfaction

Several of the value experiences discussed in this section could be characterized as experiences associated with satisfying some need, such as pleasant experiences from fulfilling a need for consistency, a need for autonomy or agency, or a need for regulatory fit. Certainly there are experiences associated with need satisfaction, and the experiences I have described can relate to some kind of need satisfaction. However, the emphasis of the value from experience viewpoint is different from the value-from-need-satisfaction viewpoint. For the value-from-experience viewpoint, it is the experience itself that matters motivationally—it is the experience itself that has value implications. In contrast, for the value-from-need-satisfaction viewpoint the experience is just a consequence of satisfying the need—it is the need satisfaction itself that is critical for value. The difference between these positions is especially evident when the value experience has effects independent of whether any need is actually satisfied.

Consider the classic animal studies on value in relation to need satisfaction. Do animals value something solely as a function of whether it satisfies their biological needs or will they choose something simply because it provides hedonic pleasure experiences? There is clear evidence that animals will choose on the basis of value experiences independent of any biological need being satisfied (Eisenberger, 1972). There are early studies, for example, showing that sweet water with saccharine that has no physiological benefit is preferred to regular water, and that animals prefer sweet food to a physiologically better food, such as a food that is more beneficial for a vitamin deficiency they have (Woodsworth & Schlosberg, 1954). There are also studies showing that rats will work to press a bar that activates the pleasure area in the brain but does not satisfy any biological need (Olds & Milner, 1954). In the original and follow-up studies, metallic electrodes were implanted in certain regions of the lateral hypothalamus and some rats would push the lever up to 5,000 times an hour even to the point of collapsing. Some mother rats even abandoned their newborn pups in order to press the lever thousands of times per hour. In a T-maze, with both arms baited with mash, the rat would stop at a point halfway down the right arm to self-stimulate, never going to the food at all. Finally, there is evidence from classical conditioning studies of animals learning to value the conditioned cues themselves, even as objects of consumption. For example, autoshaped pigeons will make eating pecks at light cues that predict edible rewards and will make drinking pecks at the same light cues when they predict liquid reward even though the light cues themselves do not satisfy hunger or thirst (Berridge & Robinson, 2003).

WHAT IS VALUE?

The foregoing review of the psychological literature on proposed sources of value—where value comes from—sets the stage for answering the more general question of what is value psychologically. In brief, I agree with those who believe that value is best understood as a psychological experience. It is not simply a belief, inference, or judgment. I believe that inferences, judgments, and beliefs contribute to value, but they do so through their impact on experience. Consider the classic refrain that people often do not appreciate something (i.e., don't value

it) until they don't have it anymore—"Sometimes you don't know what you want until you've lost it." A related notion is that people take things for granted (i.e., don't value things that are objectively important to their well-being, if not critical to their survival). A classic example, especially for younger people, is being in good physical health. Young people know that being healthy is a good thing, but that knowledge alone is not enough to give health the value it objectively deserves—" Youth is wasted on the young." The general point of such refrains is that because the need is generally being satisfied, or the desired end state maintained, without the need to pay much attention to it or make it a priority, the *experience* of attraction to the value target is weak. When suddenly the need is not being satisfied or the desired end state is not being maintained, *then* people experience their attraction to it. Only *then* do they value it highly. This is true not only for health but also for resources such as electricity, water, and so on that are normally readily available (at least in wealthier nations). Something that satisfies a need, maintains a desired end state, and meets a standard is not valuable for that reason alone. It can still be taken for granted and not be appreciated. *Experiencing* the attraction to it is essential.

What exactly is the nature of this value experience? Inspired by Lewin's (1951) discussion of valence, I have proposed in a recent paper (Higgins, 2006) that value is a *motivational* experience. For Lewin, value related to *force*, which has direction and strength. For Lewin, the forces on a person's life space were analogous to natural physical forces on objects rather than something that a person experiences. Nonetheless, following Lewin's lead, I believe that value is a *force experience* that has direction and strength. Experiencing something as having positive value corresponds to experiencing attraction toward it (e.g., trying to move in the direction toward it), and experiencing something as having negative value corresponds to experiencing repulsion from it (e.g., trying to move in a direction away from it). Value experiences vary in strength. The experience of attraction toward something can be relatively weak or strong (low or high positive value), and the experience of repulsion from something can be relatively weak or strong (low or high negative value).

To understand the psychology of value, one must begin with the classic hedonic experiences of experiencing pleasure or pain from something (cf. Kahneman et al., 1999). This experience is critical to the resultant force experience because it provides *direction*. Many of the sources of value described earlier can be understood as contributing to hedonic experience. Pleasure can be experienced when a need is satisfied, a goal is attained, or a standard is met (although, as mentioned earlier, this is not always the case). One can distinguish between different contents of value experience, such as economic value (e.g., material resources), moral value (e.g., ethical character), political value (e.g., individual power), social value (e.g., social support networks), and so on, but they can all be understood as contributing to hedonic experience—making one's life more pleasurable. It is also notable that the "cognitive" sources of value described earlier can in-

fluence the experience of hedonic direction. For example, the shared beliefs about what is desirable and what is undesirable—both social values and personal ideals and oughts—directly concern what has positive value and what has negative value. The evidence used to make evaluative inferences also provides information about the positive or negative value of something.

In my recent paper (Higgins, 2006), I proposed that the value experience of the force of attraction or repulsion has two sources. One source is the hedonic pleasure/pain experience of the target. As discussed earlier, it has been recognized for centuries that pleasure has associated with it an approach motivation and that pain has associated with it an avoidance motivation—the classic hedonic principle. However, I propose that there is a second source of the value experience that does not involve the hedonic experience of pleasure or pain per se but rather involves the *experience of the motivational force* to make something happen (experienced as a force of attraction) or make something not happen (experienced as a force of repulsion). Although these two force experiences—hedonic experience and motivational force experience—often are experienced wholistically, they are distinct from one another. Some activity that provides little hedonic pleasure, for example, may have a strong motivational force associated with it because it is the proper thing to do or matches shared beliefs about appropriate procedures of goal pursuit—I don't "enjoy" doing this but I feel "compelled" to do it. What is critical here is the notion that value is not just an experience of pleasure or pain. Instead, value is essentially a directional force experience—the force of attraction toward or repulsion away from something. Because value is an experience of directional force and not just an experience of pleasure or pain, there can be contributions to value that are independent of hedonic experience.

What are the sources of the motivational force experience? Importantly, the pleasure/pain hedonic experience is itself a critical determinant of the motivational force experience, but it is *not* the only determinant. I have proposed (Higgins, 2006) that there is also a *nondirectional* source of motivational force experience—*strength of engagement*. The state of being engaged with some object or activity is to have sustained attention, to be involved, occupied, and interested in it. Strong engagement is to be absorbed or engrossed with something. Independent of their experience of a target's properties, for example, individuals can experience their opposition to interfering forces or experience overcoming their own personal resistance in relation to that target, and these experiences can increase the strength of engagement with that target. In contrast to experienced properties of the value target, strength of engagement has strength but not valence direction in itself. Strength of engagement concerns how much people are absorbed in what they are doing regardless of the strength of the pleasure/pain experience of what they are doing.

Historically, strength of engagement relates to the notion that people can be interested in something independent of its hedonic valence and that this has value implications. Perry (1926), for example, said that an object is

valuable when interest in taken in it. Ziff (1960) said that "good" means answering to certain interests. Mandler (1984) noted that what makes us attend to things also invests them with value, and events that are interesting may or may not be positively valued. Berlyne (1973) also distinguished "interesting" from pleasing or pleasant. Strength of engagement alone does not make something attractive or repulsive. Strength of engagement does not have direction. Instead, strength of engagement contributes to the experience of motivational force, to the strength of the experienced force of attraction or repulsion. Strength of engagement contributes to *how* positively or *how* negatively something is experienced.

There are various kinds of evidence that hedonic experience and motivational force from strength of engagement can independently contribute to value. Earlier I discussed how regulatory fit can produce a particular kind of experiential feeling–making people "feel right" or "feel wrong" about something. But this distinct feeling experience is not the only effect of regulatory fit. Regulatory fit also increases strength of engagement in what one is doing (Higgins, 2000), and this contributes to motivational force independently of hedonic experience, thereby influencing value. Some recent regulatory fit studies by Idson, Liberman, and Higgins (2004) provide evidence for these independent effects. Idson and colleagues proposed that imagining successfully making a more desirable choice has higher regulatory fit for people in a promotion focus (because success maintains the eagerness that sustains their concern with advancement) than for people in a prevention focus (because success reduces the vigilance that sustains their concern with safety). Therefore, people in a promotion focus, more than people in a prevention focus, should be more engaged and experience a stronger motivational force of attraction when they imagine successfully making a more desirable choice. In contrast, imagining failing to make a desirable choice has higher regulatory fit for people in a prevention focus (because failure maintains the vigilance that sustains their focus) than for people in a promotion focus (because failure reduces the eagerness that sustains their focus). Therefore, people in a prevention focus, more than people in a promotion focus, should be more engaged and experience a stronger motivational force of repulsion when they imagine failing to make a more desirable choice (i.e., imagine making a less desirable choice).

To test these predictions, Idson and colleagues (2004) modified a well-known example from Thaler (1980). All participants were instructed to imagine that they were in a bookstore buying a book for their classes. The orientation toward the buying decision was framed in two different ways–a promotion framing and a prevention framing. Participants in the promotion, "gain/nongain" framing condition were then told: "The book's price is $65. As you wait in line to pay for it, you realize that the store offers a $5 discount for paying in cash. Of course you would like to pay $60 for the book. You have both cash and a credit card and have to choose between them." After reading the scenario, half of the participants answered a "gain" question (i.e., "How would it feel

paying in cash and getting the $5 discount?"), and the other half answered a "nongain" question (i.e., "How would it feel using your credit card and giving up the $5 discount?"). Participants in the prevention,"nonloss/ loss" framing conditions were told instead: "The book's price is $60. As you wait in line to pay for it, you realize that the store charges a $5 penalty for paying in credit." Half of these participants answered a "nonloss" question (i.e., "How would it feel paying in cash and avoiding the $5 penalty?"), and half answered a "loss" question (i.e., "How would it feel using your credit card and paying the $5 penalty?"). All participants answered these questions about the value of the choice they imagined making on a scale that went from "feeling very bad" to "feeling very good."

As one would expect, the participants felt good when they imagined successfully making the more desirable cash choice and paying just $60 for the book, and they felt bad when they imagined failing to make the more desirable cash choice (i.e., using their credit card) and paying $65 for the book. This is the classic outcome valence effect. As predicted, however, there were also significant effects within the positive outcome condition and within the negative outcome condition. The positive response to the choice success/positive outcome was more positive in the promotion focus condition (fit) than in the prevention focus condition (nonfit). The negative response to the choice failure/negative outcome was more negative in the prevention focus condition (fit) than in the promotion focus condition (nonfit).

In a subsequent study, rather than inducing regulatory focus by framing the bookstore scenario in terms of gain/nongain from a discount (promotion) or nonloss/ loss from a penalty (prevention), regulatory focus was *experimentally primed* separately from the scenario itself. An unrelated-studies paradigm was used where participants began the session by writing about either their personal hopes and aspirations (promotion ideal priming) or their personal sense of duty and obligation (prevention ought priming). All the participants were given a scenario of planning to buy a book for a course and waiting a few days to go to the bookstore to buy it. Half of the participants in each priming condition were given either a positive outcome scenario (i.e., "Imagine that you go to the bookstore in a few days, and when you arrive, you find out that there are still copies of the book for sale. You won't have to go to other bookstores to look for it.") or a negative outcome scenario (i.e., "Imagine that you go there in a few days, but when you arrive, you find out that all the copies of the book are gone. You will have to go to other bookstores to look for it.").

The results for this study in which regulatory focus was manipulated through priming were the same as those for the previous study in which regulatory focus was manipulated through framing. Once again, these findings demonstrate that regulatory fit does not add positivity to outcomes. Rather, it intensifies whatever happens to be the reaction to something. Positive reactions to desirable events are more positive when there is fit than nonfit, and negative reactions to undesirable events are more negative when there is fit than nonfit. In addition to pro-

viding support for this part of the model, the Idson and colleagues (2004) studies also examined whether pleasure/pain hedonic experience and motivational force make independent contributions to the value of something. In addition to measuring how good or bad participants felt about the imagined decision outcome, separate measures of pleasure/pain intensity and strength of motivational force were taken. The framing study and the priming study used slightly different measures to provide convergent validity. The priming study, for example, measured *pleasure/pain intensity* by asking the participants how pleasant would it be for them to find the book at the local bookstore (in the positive outcome condition) or how painful would it be for them to not find it (in the negative outcome condition) and *strength of motivational force* by asking them how motivated would they be to make the positive outcome happen (in the positive outcome condition) or how motivated would they be to make the negative outcome not happen (in the negative outcome condition).

Both studies found that pleasure/pain intensity and strength of motivational force each made significant independent contributions to the perceived value of the decision outcome (i.e., its goodness/badness). The contribution of positive versus negative outcomes to the overall goodness/badness of the decision outcome was predicted by the pleasure/pain ratings but not by the motivational force ratings. The contribution of regulatory fit to the overall goodness/badness of the decision outcome (i.e., fit being more good for a positive outcome and more bad for a negative outcome) was predicted by the motivational force ratings but not by the pleasure/pain ratings. Both studies also found that the regulatory fit manipulation did not influence pleasure/pain ratings but did influence motivational force ratings, supporting the path from regulatory fit to strength of engagement to strength of motivational force to intensity of value experience. In sum, the results of these studies suggest that both pleasure/pain hedonic experience and motivational force contribute to the value of choice outcomes (i.e., their goodness or badness), and that regulatory fit is a source of motivational force through increasing strength of engagement.

In addition to regulatory fit, other sources of where value comes from that were reviewed earlier in the chapter can also be considered as contributors to the force experience of strength of engagement. For example, the agentic experiences of opposition to interfering forces and overcoming personal resistance increase strength of engagement and motivational force that can increase the value of what one is doing (see Higgins, 2006). The ethical experience of doing what one should or meeting shared beliefs about what is proper could also increase strength of engagement and motivational force. The pursuit of need satisfaction when a need is stronger would both increase strength of engagement and motivational force during the pursuit itself and increase the hedonic experience when the need was or was not satisfied.

What about value as actual self-relations to end states? Some self-monitoring or self-evaluation models treat the value experience as something separate from the actual self-congruency or discrepancy to a desired end state. Akin to Schachter and Singer's (1962) model of emotions, there would be a congruency or discrepancy and then some interpretation of it that would produce the value experience. From this perspective, the value experiences follow from inferential interpretations. An alternative perspective is that the value experiences are directly embodied in the congruencies or discrepancies themselves without there being any need for an additional inferential interpretation. For example, I have suggested in an earlier paper (Higgins, 1996) that an actual self-congruency to a personal ideal is a type of psychological situation, a type of presence of a positive, that *is* cheerfulness (happy, joyful), and an actual self-congruency to a personal ought is a type of psychological situation, a type of absence of a negative, that *is* quiescence (calm, relaxed). Similarly, an actual self-discrepancy to a personal ideal is a type of psychological situation, a type of absence of a positive, that *is* dejection (sad, disappointed), and an actual self discrepancy to a personal ought is a type of psychological situation, a type of presence of a negative, that *is* agitation (tense, nervous). From this perspective, people have direct value experiences when systems regulating the actual self to meet desired states succeed or fail. For example, when the promotion-ideal system fails people feel sad just like they feel thirsty when the water intake system fails. I believe that if the value experiences were not present, a self-regulatory success or failure would not feel like a success or failure and its critical feedback and incentive role would be lost.

The value-as-evaluative-inference viewpoint would seem to be quite different from the value-as-experience viewpoint. However, they might be more compatible than they seem. When people construct or form an attitude about some potential choice alternative, for example, it is their predictions about the potential, future value *experiences* from that choice (e.g., future pleasure or happiness experiences) that could be critical to the value they assign to that choice (Kahneman, 2000b; Wilson & Gilbert, 2003). Kahneman and Snell (1990) defined the predicted utility of an outcome as the individual's beliefs about its experienced utility at some future time, where the experienced utility of an outcome concerns the hedonic experience of that outcome.

What about evaluative inferences from past experiences such as those described in studies on self-perception theory and activity engagement theory? As reviewed earlier, these studies have been interpreted in purely cognitive inferential terms, but it is possible to add an experiential component. Indeed, the added experiential component could be critical. From the perspective of self-determination theory (Deci, 1971; Deci & Ryan, 1985), for example, the undermining of the value of the activity that occurs when individuals are promised an award for doing the activity happens because the external force of the promised award diminishes the agentic value experience (i.e., the experience of autonomy or self-determination; see also deCharmes, 1968). Similarly, one might argue that the diminished activity value that occurs when individuals switch back and forth between two liked activities that has been found in activ-

ity engagement research (Higgins et al., 1999) happens not only because of the inferences that are made but also because the actors experience themselves choosing to disengage from an activity they expected to like. In contrast to a Bemian perspective that emphasizes self-perception like an outside observer with no access to private experiences, this perspective emphasizes the experience itself. As mentioned earlier, for example, whether the value of an activity or choice to a person is undermined by another person influencing that person depends on whether the influence is experienced as external pressure or losing control, and this can vary by the developmental age of the actor and his or her cultural background (e.g., Iyengar & Lepper, 1999).

A key direction for future research is to consider how the different sources of where value comes from described in the literature might influence not only hedonic experiences but also motivational force experiences. In fact, the critical role of experience in value may have been underestimated in the previous literature because of the tendency to consider only the role of hedonic experiences in value creation rather than the more general role of motivational force experiences. Value creation that was not clearly attributable to hedonic experience has tended to be attributed to nonexperiential sources (e.g., inferential judgments) rather than considering nonhedonic, motivational experiences as underlying the value creation. More attention needs to be paid to the nonhedonic, motivational force experiences that create value. Another critical direction for future research is to consider how the different sources of hedonic and motivational force experiences interact with one another, and, especially, how one source of value can be confused with another. Indeed, the possibilities for value confusion, and what the benefits and costs of such confusion might be, are perhaps the most intriguing and significant issues that need to be explored.

REFERENCES

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. H. (Eds.). (1968). *Theories of cognitive consistency: A sourcebook.* Chicago: Rand McNally.

Abelson, R. P., & Rosenberg, M. J. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, *3*, 1–13.

Allport, G. W. (1961). *Pattern and growth in personality.* New York: Holt, Rinehart & Winston.

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology*, *46*, 294–307.

Anderson, C. A., Krull, D. S., & Weiner, B. (1996). Explanations: Processes and consequences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 271–296). New York: Guilford Press.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bandura, A. (1982). The self and mechanisms of agency. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 3–39). Hillsdale, NJ: Erlbaum.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*, *1*, 199–218.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*, 183–200.

Bentham, J. (1988). *The principles of morals and legislation.* Amherst, NY: Prometheus Books. (Original work published 1781)

Berlyne, D. E. (1973). The vicissitudes of aplopathematic and thelematoscopic pneumatology (or The hydrography of hedonism). In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, and preference.* New York: Academic Press.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, *26*, 507–513.

Boldero, J., & Francis, J. (2002). Goals, standards, and the self: Reference values serving different functions. *Personality and Social Psychology Review*, *6*, 232–241.

Brehm, J. W. (1966). *A theory of psychological reactance.* New York: Academic Press.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance.* New York: Wiley.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*, 109–131.

Brehm, J. W., Stires, L. K., Sensenig, J., & Shaban, J. (1966). The attractiveness of an eliminated choice alternative. *Journal of Experimental Social Psychology*, *2*, 301–313.

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control.* New York: Academic Press.

Brickman, P. (1987). *Commitment, conflict, and caring.* Englewood Cliffs, NJ: Prentice-Hall.

Camacho, C. J., Higgins, E. T., & Luger, L. (2003). Moral value transfer from regulatory fit: "What feels right *is* right" and "what feels wrong *is* wrong." *Journal of Personality and Social Psychology*, *84*, 498–510.

Carder, B., & Berkowitz, K. (1970). Rats' preference for earned in comparison with free food. *Science*, *167*, 1273–1274.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior.* New York: Springer-Verlag.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right." *Journal of Personality and Social Psychology*, *86*, 388–404.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, *34*, 366–375.

Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, *37*, 12–24.

Clary, E. G., Snyder, M., Ridge, R. D., Miene, P. K., & Haugen, J. A. (1994). Matching messages to motives in persuasion: A functional approach to promoting volunteerism. *Journal of Applied Social Psychology*, *24*, 1129–1149.

Comer, R., & Laird, J. D. (1975). Choosing to suffer as a consequence of expecting to suffer: Why do people do it? *Journal of Personality and Social Psychology*, *32*, 92–101.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229–265). New York: Academic Press.

Costanzo, P. R., & Dix, T. H. (1983). Beyond the information processed: Socialization in the development of attributional processes. In E. T. Higgins, D. N. Ruble, & W. W. Hartup (Eds.), *Social cognition and social development: A sociocultural perspective* (pp. 63–81). New York: Cambridge University Press.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

deCharmes, R. (1968). *Personal causation.* New York: Academic Press.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*, 105–115.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.

Deci, E. L. (1980). *The psychology of self-determination*. Lexington, MA: Heath.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, *47*, 1105–1117.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. New York: Harcourt Brace Jovanovich.

Eisenberger, R. (1972). Explanation of rewards that do not reduce tissue needs. *Psychological Bulletin*, 77, 319–339.

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*, 967–984.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laser photo. *Journal of Personality and Social Psychology*, *63*, 877–889.

Fiske, S. T. (2004). *Social beings: Core motives in social psychology*. New York: Wiley.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*, 196–202

Freud, S. (1961). The ego and the id. In J. Strachey (Ed. & Trans.), *Standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 3–66). London: Hogarth Press. (Original work published 1923)

Frijda, N. H. (1986). *The emotions*. New York: Cambridge University Press.

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, *57*, 212–228.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilovich, T. (1981). Seeing the past in the present: The effect of associations to familiar events on judgments and decisions. *Journal of Personality and Social Psychology*, *40*, 797–808.

Gollwitzer, P. M., & Moskowitz, G. B. (1996). Goal effects on action and cognition. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399). New York: Guilford Press.

Green, D. P., Goldman, S. L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, *64*, 1029–1041.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.

Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1991). Development of self-regulatory and self-evaluative processes: Costs, benefits, and tradeoffs. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development: The Minnesota symposia on child psychology* (Vol. 23, pp. 125–165). Hillsdale, NJ: Erlbaum.

Higgins, E. T. (1996). Emotional experiences: The pains and pleasures of distinct regulatory systems. In R. D. Kavanaugh, B. Zimmerberg, & S. Fein (Eds.), *Emotion: Interdisciplinary perspectives* (pp. 203–241). Mahwah, NJ: Erlbaum.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Higgins, E. T. (2006). Value from hedonic experience and engagement. *Psychological Review*, *113*, 439–460.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*, 1140–1153.

Higgins, E. T., & Stangor, C. (1988). A "change-of-standard" perspective on the relations among context, judgment, and memory. *Journal of Personality and Social Psychology*, *54*, 181–192.

Higgins, E. T., Strauman, T., & Klein, R. (1986). Standards and the process of self-evaluation: Multiple affects from multiple stages. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 1. Foundations of social behavior* (pp. 23–63). New York: Guilford Press.

Higgins, E. T., Trope, Y., & Kwon, J. (1999). Augmentation and undermining from combining activities: The role of choice in activity engagement theory. *Journal of Experimental Social Psychology*, *35*, 285–307.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist*, *52*, 35–44.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. New Haven, CT: Yale University Press.

Hume, D. (1975). *An equiry concerning the principles of morals* (J. B. Schneewind, Ed.). Cambridge, UK: Hackett. (Original work published 1777)

Hyman, H. H. (1942). The psychology of status. *Archives of Psychology*, 269.

Idson, L. C., Liberman, N., & Higgins, E. T. (2004). Imagining how you'd feel: The role of motivational experiences from regulatory fit. *Personality and Social Psychology Bulletin*, *30*, 926–937.

Iyengar, S. S., & Lepper, M. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology*, *76*, 349–366.

James, W. (1948). *Psychology*. New York: World. (Original work published 1890)

Jensen, G. D. (1963). Preference for bar pressing over "freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology*, *65*, 451–454.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York: Academic Press.

Kahneman, D. (1999). Objective happiness. In D. Kahnemen, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage.

Kahneman, D. (2000a). Experienced utility and objective happiness: A moment-based approach. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 673–692). New York: Cambridge University Press.

Kahneman, D. (2000b). New challenges to the rationality assumption. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 758–774). New York: Cambridge University Press.

Kahneman, D., Diener, E., & Schwarz, N. (1999). *Well-being: The foundations of hedonic psychology*. New York: Russell Sage.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.

Kahneman, D., & Snell, J. (1990). Predicting utility. In R. M. Hogarth (Ed.), *Insights in decision making* (pp. 295–310). Chicago: University of Chicago Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*, 163–204.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally.

Kohlberg, L. (1976). Moral stages and moralization. In T. Lickona (Ed.), *Moral development and behavior*. New York: Holt, Rinehart & Winston.

Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review*, *82*, 387–406.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263–283.

Larsen, R. J., & Diener, E. (1985). A mutitraitmultimethod examination of affect structure: Hedonic level and emotional intensity. *Personality and Individual Differences*, *6*, 631–636.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the overjustification hypothesis. *Journal of Personality and Social Psychology*, *28*, 129–137.

Lewin, K. (1951). *Field theory in social science*. New York: Harper.

Lewin, K. (1952). Constructs in field theory [1944]. In D. Cartwright (Ed.), *Field theory in social science: Selected theoretical papers by Kurt Lewin* (pp. 30–42). London: Tavistock.

Lewis, M. (1964). Some nondecremental effects of effort. *Journal of Comparative Physiological Psychology*, *57*, 367–372.

Lewis, M. (1965). Psychological effect of effort. *Psychological Bulletin*, *64*, 183–190.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, p. 255–295). New York: Academic Press.

Mandler, G. (1984). *Mind and body: The psychology of emotion and stress*. New York: Norton.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

Merton, R. K. (1957). *Social theory and social structure*. Glencoe, IL: Free Press.

Merton, R. K., & Kitt, A. S. (1952). Contributions to the theory of reference-group behavior. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (2nd ed., pp. 430–444). New York: Holt, Rinehart & Winston.

McGuire, W. J. (1968). Theory of the structure of human thought. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A source book* (pp. 140–162). Chicago: Rand McNally.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Miller, N. E. (1963). Some reflections on the law of effect produce a new alternative to drive reduction. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 11, pp. 65–112). Lincoln: Nebraska University Press.

Moretti, M. M., & Higgins, E. T. (1999). Internal representations of others in self-regulation: A new look at a classic issue. *Social Cognition*, *17*, 186–208.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.

Newcomb, T. M. (1953). An approach to the study of communicative acts. *Psychological Review*, *60*, 393–404.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*, 419–427.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. New York: Cambridge University Press.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, *62*, 42–55.

Oxford English Dictionary. (1971). *The compact edition* (Vols. I, II). Oxford, UK: Oxford University Press.

Peacocke, C. (2004). *The realm of reason*. Oxford, UK: Oxford University Press.

Perry, R. B. (1926). *General theory of value: Its meaning and basic principles construed in terms of interest*. Cambridge, MA: Harvard University Press.

Piaget, J. (1965). *The moral judgment of the child*. New York: Free Press (Original work published 1932)

Pittman, T. S., & D'Agostino, P. R. (1989). Motivation and cognition: Control deprivation and the nature of subsequent information processing. *Journal of Experimental Social Psychology*, *25*, 465–480.

Pittman, T. S., & Higgins, E. T. (2006). *Human motivation: Comprehending and acting instrumentally and expressively*. Unpublished manuscript, Colby College.

Pittman, T. S., & Pittman, N. L. (1980). Deprivation of control and the attribution process. *Journal of Personality and Social Psychology*, *39*, 377–389.

Plato. (1949). *Timaeus* (B. Jowett, Trans.). Indianapolis, IN: Bobbs-Merrill. (Original publication 4th century B.C.)

Powers, W. T. (1973) *Behavior: The control of perception*. Chicago: Aldine.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338–342.

Redelmeier, D., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *116*, 3–8.

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*, 133–148.

Rogers, C. R. (1961). *On becoming a person*. Boston: Houghton Mifflin.

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. *Journal of Psychology*, *91*, 93–114.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Rokeach, M. (1979). Change and stability in American value systems, 1968–1971. In M. Rokeach (Ed.), *Understanding human values: Individual and societal*. New York: Free Press.

Rokeach, M. (1980). Some unresolved issues in theories of beliefs, attitudes, and values. In H. E. Howe, Jr. & M. M. Page (Eds.), *1979 Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.

Roseman, I. J. (1984). Cognitive determinants of emotion: A structural theory. *Review of Personality and Social Psychology*, *5*, 11–36.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Salancik, G. R., & Conway, M. (1975). Attitude inferences from salient and relevant cognitive content about behavior. *Journal of Personality and Social Psychology*, *32*, 829–840.

Schachter, S., & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review*, *69*, 379–399.

Scherer, K. R. (1988). Criteria for emotion-antecedent appraisal: A review. In V. Hamilton, G. H. Bower, & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 89–126). Norwell, MA: Kluwer Academic.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, *44*, 229–237.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York: Academic Press.

Schwartz, S. H., & Bilsky, W. (1987). Toward a universal structure of human values. *Journal of Personality and Social Psychology*, *53*, 550–562.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? The informative function of affective states. In K. Fiedler & J. Forgas

(Eds.), *Affect, cognition and social behavior* (pp. 44–62). Toronto: Hogrefe.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford Press.

Seligman, C., Olson, J. M., & Zanna, M. P. (1996). *The psychology of values: The Ontario Symposium* (Vol. 8). Mahwah, NJ: Erlbaum.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication*. New Haven, CT: Yale University Press.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29–39.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton.

Smith, A. (1997). *The theory of moral sentiments*. Washington, DC: Regnery. (Original work published 1759)

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*, 813–838.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. New York: Wiley.

Snyder, M., & Cantor, N. (1998). Understanding personality and social behavior: A functionalist strategy. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 635–679). New York: McGraw-Hill.

Sorrentino, R. M., & Roney, C. J. R. (2000). *The uncertain mind: Individual differences in facing the unknown*. Philadelphia: Psychology Press.

Spence, K. W. (1958). A theory of emotionality based drive (D) and its relation to performance in simple learning situations. *American Psychologist*, *13*, 131–141.

Spinoza, B., de. (1986). *Ethics and on the correction of the understanding* (A. Boyle, Trans.). London: Dent. (Original work published 1677)

Stevens, L. E., & Fiske, S. T. (1995). Motivation and cognition in social life: A social survival perspective. *Social Cognition*, *13*, 189–214.

Swann, W. B., Jr., Wenzlaff, R. M., Krull, D. S., & Pelham, B. W. (1992). Allure of negative feedback: Self-verification strivings among depressed persons. *Journal of Abnormal Psychology*, *101*, 293–306.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). San Diego, CA: Academic Press.

Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, *1*, 39–60.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*, 183–206.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Erlbaum.

Titchener, E. B. (1910). *A text-book of psychology* (rev. ed.). New York: Macmillan.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *93*, 506–520.

Trope, Y. (1986a). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Trope, Y. (1986b). Self-enhancement and self-assessment in achievement behavior. In R. M. Sorrentino & E. T. Higgins

(Eds.), *Handbook of motivation and cognition: Vol. 1. Foundations of social behavior* (pp. 350–378). New York: Guilford Press.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–421.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *85*, 1124–1131.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–192). New York: Academic Press.

Tyler, T. R., & Smith, H. J. (1998). Social justice and social movements. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 595–629). New York: McGraw-Hill.

Van Hook, E., & Higgins, E. T. (1988). Self-related problems beyond the self-concept: The motivational consequences of discrepant self-guides. *Journal of Personality and Social Psychology*, *55*, 625–633.

Weary, G., Gleicher, F., & Marsh, L. (Eds.). (1993). *Control motivation and social cognition*. New York: Springer-Verlag.

Weber, M. (1967). Subjective meaning in the social situation. In G. B. Levitas (Ed.), *Culture and consciousness: Perspectives in the social sciences* (pp. 156–169). New York: Braziller.

Webster's ninth new collegiate dictionary. (1989). Springfield, MA: Merriam-Webster.

Weiner, B. (1972). *Theories of motivation: From mechanism to cognition*. Chicago: Rand McNally.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333.

Wicklund, R. A. (1974). *Freedom and reactance*. New York: Wiley.

Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Hillsdale, NJ: Erlbaum.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Erlbaum.

Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine*. Cambridge, MA: MIT Press.

Williams, B. (1985). *Ethics and the limits of philosophy*. Cambridge, MA: Harvard University Press.

Williams, R. M., Jr. (1979). Change and stability in values and value systems: A sociological perspective. In M. Rokeach (Ed.), *Understanding human values: Individual and societal*. New York: Free Press.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 345–411). New York: Elsevier.

Woodworth, R. S. (1918). *Dynamic psychology*. New York: Columbia University Press.

Woodworth, R. S. (1940). *Psychology* (4th ed.). New York: Holt.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (rev. ed.). New York: Holt, Rinehart & Winston.

Wright, H. F. (1937). *The influence of barriers upon strength of motivation*. Durham, NC: Duke University Press.

Wundt, W. (1999). Outlines of psychology. In R. H. Wozniak (Ed.), *Classics in psychology, 1896* (Vol. 35). Bristol, UK: Thoemmes Press. (Original work published 1896)

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1–27.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge, UK: Cambridge University Press.

Ziff, P. (1960). *Semantic analysis*. Ithaca, NY: Cornell University Press.

Zimbardo, P. (1969). *The cognitive control of motivation: The consequences of choice and dissonance*. Glenville, IL: Scott, Foresman.

Basic Human Needs

Thane S. Pittman
Kate R. Zeigler

It is vain to do with more what can be done with less.
—attributed to William of Occam (c. 1285–1349)

There is always an easy solution to every human problem—neat, plausible, and wrong.
—Mencken (1949, p. 443)

It has been a long time since a chapter devoted to the subject of basic human needs appeared in a major handbook in social psychology; indeed, there has never been one (Gilbert, Fiske, & Lindzey, 1998; Higgins & Kruglanski, 1996; Lindzey, 1954; Lindzey & Aronson, 1968, 1985; Murchison, 1935). A search of chapter titles in the *Annual Review of Psychology* also came up empty. The discovery of these facts gave us considerable reason to pause. But as interest in using a motivational perspective for the generation of hypotheses and the interpretation of findings has increased (cf. Higgins & Kruglanski, 2000; Pittman, 1998; Pittman & Heller, 1987), theorists have begun to return to the question, "What are the basic human needs?" It thus does seem to be an appropriate time to assess the ways in which ideas about basic human needs have been and are being used in social–psychological theories and research.

Rather than providing a thorough literature review of all research using constructs proposed to represent the operation of basic human needs, we instead review much more selectively the current state of theories about basic human needs, with a little historical context. We have chosen six theories for comparison. All these theories have been given extensive explication and review elsewhere, so we focus particularly on how these theories are structured and at what level of analysis they are designed to apply. We will find a little agreement but perhaps considerable food for thought.

HUMAN NATURE

What is human nature? Is there such a thing? Although it is not the way that social psychologists have approached this question, an informal consideration of other species in comparison suggests that there must be such a thing as human nature. We know that dogs are different from cats in many ways, and that neither is the same as a horse, a rat, a dolphin, or an orangutan. These mammals are very different in size and shape, but we also sense from their behavior that the differences among them go beyond those obvious physical characteristics. Humans share many basic similarities with all these animals but are also unlike any of them in many aspects of their physical construction and, we suspect, in their psychological processes. But what are those psychological differences, and to what extent are they simply differences in degree (e.g., in the extent of information-processing capability) versus qualitative differences (perhaps, for example, in ambition, or in the inclination and ability to construct symbolic meaning)? Such questions are not easy to answer, but it is also the case that most of our work in psychology has not been designed to address them.

To pursue this line of thinking a bit further, nevertheless, it is easy to distinguish a human from other animals based on physical appearance. Bipedal and relatively hairless would seem to do it. "Find the human" is not a common item on tests of intelligence. To make such dis-

tinctions based on behavior is also quite simple. A clue such as "makes automobiles" resolves all ambiguity as to which species we have in mind. Of course, not all humans make automobiles, but archaeologists do not need to expend a great deal of energy answering the question "made by which species?" upon discovery of clay tablets, tools, or pottery shards found in digs around the globe. Anyone reading this chapter indoors is surrounded by, sitting on, probably wearing, and using things fabricated by *homo sapiens*. Humans make things.

When we move into the realm of psychology, matters become more complex and considerably less clear. In considering what, psychologically, might constitute human nature, social psychologists have not taken up the method of cross-species comparisons illustrated in the aforementioned musings. In fact, in psychology more generally, instead of looking for what is unique about human nature, psychologists have tried repeatedly to come up with a few general principles that are meant to apply across all or virtually all species, essentially arguing that psychologically all species are governed by the same fundamental principles. This approach constitutes a scientifically sound strategy, in the spirit of Occam's Razor (i.e., given two equally good explanations, the more simple one is preferred), as an attempt to understand complexity by reducing it to a few simple laws that apply to many species. Familiar examples of this approach include the various serious attempts by behaviorists to explain everything in terms of basic principles of reinforcement (Hull, 1943; Pavlov, 1927; Skinner, 1938, 1981; Watson, 1930). These ideas were tested and refined extensively with experiments on rats and pigeons as well as humans, were extended into such realms as social learning (e.g., Miller & Dollard, 1941), attitude formation and change (e.g., Doob, 1947), and group behavior (Zajonc, 1965), and are still echoing in the traditional halls of economics. Here we have also found, repeatedly and particularly at social psychology's level of analysis, that as useful as such simple constructions have been, they are not adequate to the task of fully explaining, understanding, and predicting the behavior of humans (cf. Dweck, 1996; Festinger, 1957; Henchy & Glass, 1968; Kahneman & Tversky, 1979).

If social psychologists have not pursued a comparative or a multispecies approach, other than to show that simple cross-species analyses are not sufficient to understand human psychology, then how have they gone about the task of understanding human nature? Leaving aside the question of what makes humans distinctive or unique, and instead focusing on trying to understand human behavior in its own right, one approach (our subject) has been to specify the basic set of things that humans need. This is not the only way to proceed, and use of the need concept is not universally thought to be a good idea. But as part of a motivational approach to human behavior (also not embraced by all), trying to understand human nature means understanding the basic things that humans want and need. Of course, we learn to want all sorts of things—even to need them—but can we find a fundamental and universal set of basic needs that characterize human nature? What are the fundamental

dimensions of human desire, and how might understanding those basic human needs help us to understand human behavior?

WHAT ARE NEEDS, AND DO WE NEED THEM?

> Man is a wanting animal and rarely reaches a state of complete satisfaction except for a short time. As one desire is satisfied, another pops up to take its place. When this is satisfied, still another comes into the foreground, etc. It is a characteristic of the human being throughout his whole life that he is practically always desiring something.
> —MASLOW (1970, p. 24)

Motivational approaches focus on the generative aspect of human behavior, on the forward-moving, internally driven aspect of behavior that derives in part from internal impulses that include but also go beyond simple tissue deficits (see Higgins & Kruglanski, 2000; Pittman, 1998; Pittman & Heller, 1987, for reviews of motivational research in social psychology). Motivational analyses do not, however, necessarily include the use of any assumptions about fundamental needs, nor do they necessarily embrace the emphasis on motives as internal drivers. Instead of employing the concept of need, some have argued for a focus on *incentives* and *goals* as the way to understand motivation. For example, on the question of the nature of control motivation (Pittman, 1993; Pittman & D'Agostino, 1989; Weary, Gleicher, & Marsh, 1993), Bandura (1997) has argued against thinking of an inborn or acquired need for control as providing the *push* from internal motives and instead in favor of the *pull* from anticipated environmental incentives:

> Some theorists regard the striving for control as an expression of an inborn drive (Deci & Ryan, 1985; White, 1959). . . . In social cognitive theory, people exercise control for the benefits they gain by it. Some of these benefits may involve biological gratifications, but the striving for control is not a drive in its own right. (p. 16)

Similarly, this general expectancy × value incentive-based analysis, along with elaborations of the concept of negative feedback loops (Miller, Galanter, & Pribram, 1960), has been employed in Carver and Scheier's (1981, 1998, 2000) approach to self-regulation. Research using goals and self-regulation of goal pursuit also tends to be either purely or primarily cognitive in its theorizing and does not depend on assumptions about basic needs (cf. Gollwitzer & Moskowitz, 1996; Kruglanski, 1996).

The approach to motivation represented in the theories on which we will concentrate is one in which motives are seen as variable internal states that, when activated or aroused, energize and direct behavior. In these models, *deprivation* and *deviation* (from desired states) play a central role. In the early work on learning and conditioning, a common method for getting an animal prepared for a session would be to deprive the animal of food for some specified length of time. The practical effect of this prac-

tice was to create in the animal a real desire for food and a willingness to expend considerable energy to get it. Hungry rats will more reliably and persistently press a bar for food than will rats that are partially or completely satiated. It seemed clear, at least to many, that rats *need* food and are *motivated* to get it when deprived.

The recognition that there are physical nutrients such as food, water, and oxygen that are necessary for survival was used by theorists such as Hull (1943) to develop the concept of drive, a motivational force that provided the impetus toward behavior, an impetus that would wax and wane as a function of the extent of tissue deficits. As we have said, early attempts were made to explain all of animal (including human) behavior with a small set of learning and conditioning principles based on these tissue deficits. But the work of psychologists such as White (1959) on effectance, Hunt (1965) on optimal levels of stimulation, Berlyne (1960) and Harlow (1953) on curiosity, and Festinger and Carlsmith (1959) on the inverse relationship between incentives and attitude change began to make it clear that to understand the complexity of human behavior, it would be necessary to go beyond simple principles of learning and an exclusive reliance on building only on basic tissue deficit needs such as hunger and thirst.

How Should Needs Be Defined?

> It is characteristic of this deeper analysis that it will always lead ultimately to certain goals or needs behind which we cannot go; that is, to certain need-satisfactions that seem to be ends in themselves and seem not to need any further justification or demonstration. These needs have the particular quality in the average person of not being seen directly very often but of being more often a kind of conceptual derivation from the multiplicity of specific conscious desires. In other words then, the study of motivation must be in part the study of the ultimate human goals or desires or needs.
>
> —MASLOW (1970, p. 22)

A source of potential confusion lies in the varying meanings that are intended when theorists use the term "need." *Webster's Seventh New Collegiate Dictionary* (Gove, 1963) defines the noun "need" as a lack of something required, useful, or desired. Defining a need as something that is required leads to a set of implications that are quite different from those that would follow from defining a need as something that is useful or desired.

Some needs are clearly required, necessary in that without them the organism would, in a relatively short time, cease to be alive. Without food, a human would survive for a period of time conveniently counted in weeks; without water, days. In the frigid waters of the North Atlantic or the Artic Sea, mariners know that death from hypothermia following immersion would come in a matter of minutes. Without oxygen, human life only endures for seconds, perhaps a few minutes. These required forms of sustenance, the antidotes for otherwise fatal tissue deficits, provided the foundation for early studies of the phenomena of learning and conditioning (cf. Cofer & Appley, 1964).

Sexual relations, however, although needed in the sense of being required for the continued existence of a species and needed in the sense that they are at times strongly desired and clearly biologically based, are not necessary for individual survival in the same sense as are food, water, and air. Although sexual desire can be considered in terms of drive, and as something fundamentally desired, it is not needed by the individual in the same way as food and water are needed.

When social–psychological theorists talk about basic human needs, they are usually not talking about tissue deficits but such things as control, understanding, or self-esteem. It is worth thinking about what this use of the term "need" actually is intended to mean. One possibility is to consider that some basic needs are necessary for continued existence, but that the time scale for continued existence without satisfying those needs is more conveniently measured in years or decades rather than in minutes, hours, or days. We can take for an example the assumption that there is a need to belong (Baumeister & Leary, 1995). A need to belong is included in a number of the theories to be considered. At the beginning of life, survival without the assistance of others is not possible, because it is only with their help that an infant can obtain basic sustenance. At such young ages, social support is just as essential as food and water for survival. But beyond those early years, the data on survival rates as a function of the extent of social support (cf. Deeg, van Zonneveld, van der Maas, & Habbema, 1989) would still be consistent with this definition of basic need (i.e., something that is required for existence), but the time scale for damage due to deprivation would be considerably longer than it is at the beginning of life.

However, instead of using a definition of need as something required for existence, most current social–psychological theories of basic needs employ some version of the concept of *thriving*. For example, Deci and Ryan (2000) argue that truly basic needs are those that influence a person's well-being. Experiences that satisfy these basic needs are thought of as nutriments, and insufficient amounts of these nutriments result in a failure to thrive. Inadequate degrees of satisfaction of these basic needs may not lead to premature death but instead are revealed in the failure to achieve one's potential or to function as well as one might under more optimal conditions of need fulfillment. Need satisfaction should promote well-being and psychological thriving (Sheldon, Elliot, Kim, & Kasser, 2001). Those emphasizing survival-oriented considerations (e.g., Fiske, 2004; Pyszczynski, Greenberg, & Solomon, 1997) also seem to use the concepts of basic needs or motives as something that will leave the organism better off and probabilistically, if not necessarily individually, more likely to survive and thrive if satisfied.

Baumeister and Leary (1995), in their argument for considering belongingness to be a basic human need, suggested a set of criteria that could be used to evaluate whether a need is truly fundamental, distinguishable from needs that may be derived from other more fundamental ones. They argue that a *fundamental need* should:

1. Produce effects readily under all but adverse conditions.
2. Have affective qualities.
3. Direct cognitive processing.
4. Lead to ill effects when thwarted (e.g., poor health or adjustment).
5. Elicit goal-oriented behavior designed to satisfy it—subject to motivational patterns such as object substitutability and satiation.
6. Be universal in the sense of applying to all people.
7. Not be derivative of other motives.
8. Affect a broad variety of behaviors.
9. Have implications that go beyond immediate psychological functioning.

Although adoption of this or some similar set of criteria might seem sensible and broadly compatible with how the concepts of basic motives or needs are typically used in social–psychological theories, such tests are not typically employed by theorists to develop or evaluate a set of basic needs. Instead, the more common strategy is to pursue a hypotheticodeductive process of hypothesis generation and testing given the assumed basic need structure.

How Many Needs Do We Need?

In social and personality analyses, the notion of needs characteristic of the human organism had an early history, as did the recognition that specification of internal cognitive, emotional, and motivational processes would be required for an effective approach to understanding human psychology. But whether this approach also needed to include a specification of basic human needs was and continues to be a matter on which theorists disagree. Making lists of needs began to acquire a bad reputation with McDougall's (1926) concept of instincts, an approach that was soon rejected for using circular reasoning and leading to a plethora of assumed inborn tendencies. Murray's (1938) list of more than 20 needs was also seen by many as going down a path that would only lead to longer and longer lists, of dubious usefulness. It may be that the baby was thrown out with the bath, however, as the majority of experimentally inclined social and personality psychologists eschewed using the concept of basic needs altogether in favor of a focus on situational constraints and individual differences as the twin paths to understanding.

Most of those who made specific motivational assumptions (e.g., Festinger's, 1957, use of the consistency principle as a source of motivation) took no clear position on whether these motives were acquired or innate and also tended to focus on a single motivational dimension. This is still the dominant approach. Investigators can assume that something (e.g., self-esteem) is a strong motive or need and use their theoretical assumptions about how it works to generate hypotheses. They can do so without having to take a position on whether it is a fundamental need or an acquired motive, a derived motive based on some more fundamental need or on what other needs also might be fundamental. This approach leads to what

are sometimes described as minitheories as opposed to theories with a more comprehensive sweep.

Recently we have seen a reemergence of interest in specifying the set of basic needs, with an eye toward sticking to a short list that would avoid the criticisms that had been applied to the long and lengthening sets of needs of earlier theorists such as McDougall and Murray. Given the checkered past of long lists of needs, these theories confine their basic set to a number that can be counted on the fingers of one (human) hand. Before proceeding to our set of comparison theories, a brief review of Murray's approach provides a transition to the current theories of basic human needs that we will compare. Considering Murray's list may offer a useful perspective on the question of how many needs should be specified in the set of basic human needs.

Murray's List of Needs: A Nest of Boxes?

> . . . the list of drives one chooses to list depends entirely on the degree of specificity with which one chooses to analyze them. The true picture is not one of a great many sticks lying side by side, but rather of a nest of boxes in which one box contains three others, and in which each of these three contains ten others, and in which each of these ten contains fifty others, and so on.
>
> —MASLOW (1970, p. 25)

Murray's (1938) list was generated empirically and resulted in a large and well-known set of both manifest and latent needs. Several of these needs have been investigated intensively, particularly the needs for achievement (Atkinson, 1958; Atkinson & Raynor, 1978; McClelland, 1958; McClelland, Clark, Roby, & Atkinson, 1949), for power (McClelland & Watson, 1973; Winter, 1973; see also Fiske & Berdahl, Chapter 29, this volume), and for affiliation (Baumeister & Leary, 1995; Schachter, 1959), but many of the others have been less well studied, if not forgotten. What is remembered is that the list was too long. However, another look at Murray's (1938) analysis shows that the question of exactly how many needs he identified can be answered in more than one way. Although Murray did list 20 manifest needs, they were presented in nine groups. These nine groups were not named, but we have taken the liberty of doing so in Figure 20.1. Murray also noted that most of these manifest needs were represented by four basic reaction systems, also shown in Figure 20.1. In the spirit of Maslow's nesting boxes metaphor, the number of needs identified by Murray could be 20, 9, or 4 (see Figure 20.1). Or five: Costa and McCrae's (1988) analysis of the Jackson (1984) Personality Research Form (PRF) suggests that Murray's needs as measured by the PRF can largely be captured by the Big Five set of personality dimensions.

As we turn to the six theories chosen for comparison, it is worth remembering that the number of basic human needs is to a large extent dependent on the level of definitional generality or specificity being used by the theorist. All these theories present a set of basic motives ranging in number from three to five, but it is instructive to remember Maslow's nesting-boxes metaphor, as illustrated by the various ways of counting up the manifest
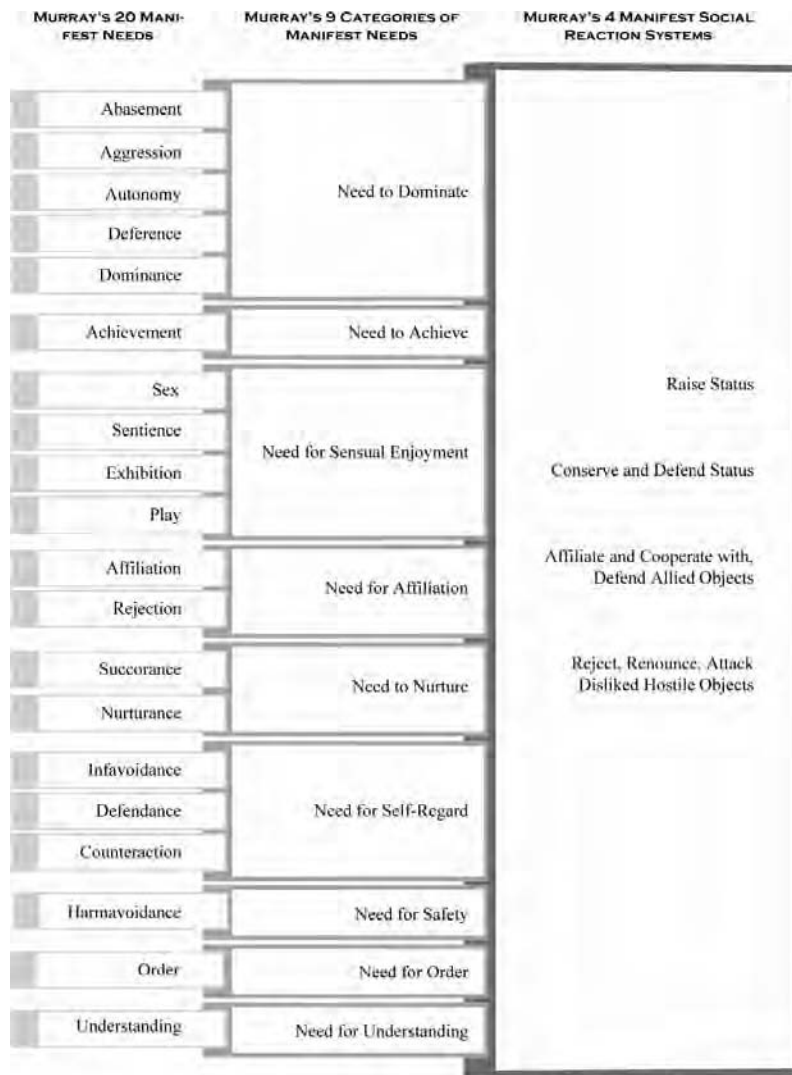
FIGURE 20.1. Murray's (1938) manifest needs represented as nesting boxes: 20, 9, or 4?

needs identified by Murray. In some cases, we can easily imagine reducing the set of basic needs that is proposed to fewer, or expanding it to more if we are willing to move to a more specified level of detail.

Theoretical Comparisons

Six need theories will be compared: Maslow's (1943) hierarchy of needs; core social motives theory (Fiske, 2004; Stevens & Fiske, 1995); terror management theory (Pyszczynski et al., 1997); attachment theory (Bowlby, 1969); cognitive-experiential self-theory (Epstein, 1992); and self-determination theory (Deci & Ryan, 1980, 2000). We focus on two general dimensions of these theories: how they structure the set of basic needs, and how those needs are distributed across levels of analysis. Although these are certainly not the only theories of basic needs we could have chosen to compare, they are representative and include among them several current theories that have led to a considerable amount of research.

STRUCTURAL ASSUMPTIONS IN THEORIES OF BASIC HUMAN NEEDS

One of the ways in which theories about basic human needs differ lies in the proposed structure of those needs. These structural assumptions vary considerably. We have included for comparison a *hierarchical model*; several theories that specify a *root primary need* to which other needs are related; a theory that proposes a *system of checks and balances* across a conscious/nonconscious divide; and a theory including a proposed *set of independent needs*. Figure 20.2 illustrates these structural variations.

Hierarchical Structure

Maslow's Hierarchy of Needs

It could be said, with only a bit of exaggeration, that Maslow's hierarchy of needs is a theory that everyone knows, and no one uses. Typically portrayed in introduc-
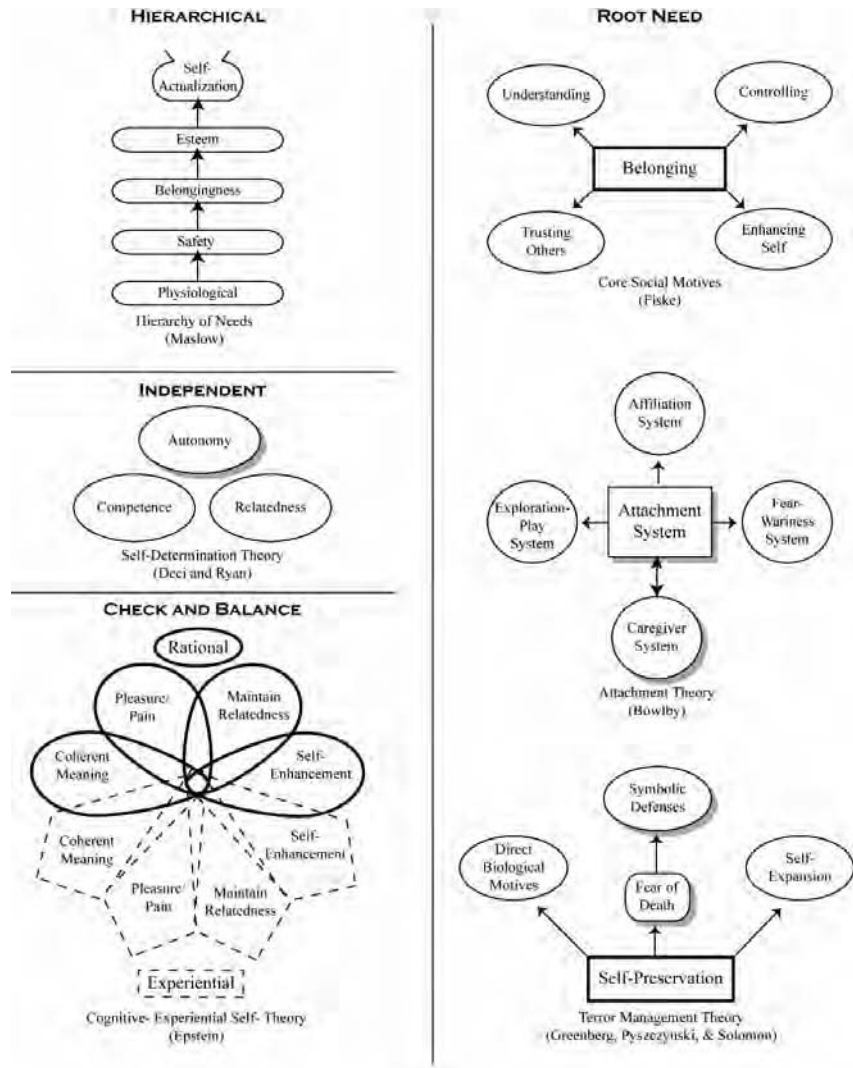
FIGURE 20.2. Structural form of six theories of basic human needs. Shadowing indicates the needs that have generated the most empirical research.

tory textbooks in a familiar pyramidal figure, the basic structural assumption is that some needs take precedence over others, and that those more fundamental to existence must be satisfied before others will be addressed (Maslow, 1943). The theory is often portrayed, and we have done it again in Figure 20.2, as having five levels representing five types of needs. In fact, Maslow discussed a much more nuanced division of needs, so that one could argue about the actual number of needs that he thought were fundamental, as his "nesting boxes" quote at the beginning of the earlier section on Murray suggests. But regardless of the number of needs that are specified, the invariance of any such hierarchy is easy to call into question. We have ample examples of needs higher in the hierarchy taking precedence over those said to have priority (in the case of basic sustenance we find people starving themselves to death for the sake of appearance, to save a loved one, or in service to a cause). In actual operation, it becomes difficult to see how in

specific settings this hierarchy could either hold up or make testable predictions, particularly at the higher levels of the hierarchy. This may be why the theory has not generated a great deal of empirical research.

There are other current theories that use the notion of a hierarchy. Steele's (1988) self-affirmation theory is an example. When threatened with an inconsistency (Steele & Liu, 1981, 1983) or a failure (Liu & Steele, 1986), self-affirmation theory predicts that these concerns can be taken care of at the level at which discomfort was created (achieve consistency, regain control) or at the higher-level basic need to feel good about oneself (affirm self). Note that in this model, the most important need is the one highest in the hierarchy. This also illustrates an advantage of more targeted and bounded theories. In such theories predictions and explanations are limited to the particular conditions under which the processes specified in the model should operate. By not claiming to capture ev-

ery aspect of basic human needs these theories have the advantage of being more easily testable, even if the overall import of the analysis is on a scale less grand.

Maslow's theory has been most influential in organizational psychology and related fields as a model of motivation but has not itself motivated a great deal of empirical research. The concept of self-actualization, however, has been influential in the construction of other current theories, most notably self-determination theory. Our discussion of that theory comes later, but we note here that Vallerand (2000) has offered a hierarchical version of self-determination theory.

Root Need Structure

Another basic kind of theoretical structure is one in which a single need is identified as the most important one. This *root need* is either more important than the others, is the one to which the others are closely related, or the one from which the others are derived. Three different versions of root need models are discussed.

Core Social Motives Theory

Based on a comprehensive literature review of a wide variety of writings on basic needs and motives, Stevens and Fiske (1995) argued that there was overall agreement on five basic needs. Fiske (2002, 2004) continued to develop and elaborate this set of basic needs, or core social motives, using the BUC(K)ET acronym as a mnemonic device for the five motives: *belonging, understanding, controlling, enhancing self*, and *trusting* (leaving the "K" for students to play with if they would like to imagine a sixth motive). In this system, *belonging* is the root need, the essential core social motive. The others are all said to be in service to, facilitating, or making possible effective functioning in social groups. No structural relation among the five motives is specified, but we have visually arranged this theory in Figure 20.2 with belongingness at the center and the other four motives arranged around and connected to it.

As implied by its name, this theory is specifically designed to apply to needs that arise in social settings. "Core social motives describe fundamental, underlying psychological processes that impel people's thinking, feeling, and behaving in situations involving other people" (Fiske, 2004, p. 14). A basic assumption of this theory is that underlying all the basic needs is an evolutionary process that has led to these characteristics of human nature because they promote survival of the individual through belonging in groups. Although this kind of imagined evolutionary, survival-oriented thinking is not logically a required aspect of a theory of basic needs with a root need structure, in fact such thinking has been employed in the development of all three of the root need theories in Figure 20.2.

Core social motives theory has been used primarily as an organizing structure for understanding a wide variety of social–psychological research (Fiske, 2004). At this writing, it had not yet been used extensively to generate and test new research questions, and may need some further development to generate testable hypotheses.

Attachment Theory

We should begin by making it clear that attachment theory was not put forward as a theory of basic needs. Instead, attachment theory (Bowlby, 1969) was developed as a goal-corrected systems version of control theory (cf. Bandura, 1997; Carver & Scheier, 1981, 1990, 1998, 2000), without reference to needs. In fact, Bowlby explicitly objected to use of the need concept. His reasons included his apprehension about the difficulties inherent in using what we would call the strong form of the term (i.e., required for continued survival of the organism) given that humans can survive if not thrive even when several of the attachment systems remain unsatisfied. He was also concerned that the use of the need concept, in which action is designed to create some preordained future state, might invite in "the fallacy of teleology" (Bowlby, 1969, p. 137).

Even so, we have included attachment theory in our set of comparisons. Specifying an attachment system as the central organizing principle, this theory is based on deprivation and deviation, with self-correcting control systems that could easily be thought of as need based and motivational in character, even if Bowlby preferred not to do so. The attachment, caregiver, and affiliation systems could all be thought of under the heading of belongingness or relatedness. Viewed in this fashion, belongingness is seen to be a complex set of compatible but quite different processes (nesting boxes). Most of the work with adult attachment has focused on styles of attachment (secure, anxious, ambivalent) as they play out in adult close relationships (Ainsworth & Wittig, 1969; Hazan & Shaver, 1994; Reis & Patrick, 1996; Rholes & Simpson, 2004). Attachment theory also includes an exploration/play system that provides a dimension of comparison with self-determination theory and terror management theory. It includes as well a fear-wariness system that has been traditionally addressed with motivational analyses, and this provides a general point of contact with the pleasure/pain aspect of cognitive-experiential self-theory. The theoretical root need structure of this developmental theory also introduces a developmental change aspect to our discussion.

Bowlby did relate attachment and closeness to the survival needs of nurturance and security. This distinction forms the basis for the survival functions underlying promotion (nurturance) and prevention (security) focus in regulatory focus theory (Higgins, 1997, 1998, 1999) and in the previous self-discrepancy theory (Higgins, 1987, 1989). Regulatory focus theory is a kind of need distinction model, without strong claims about evolution but employing Bowlby's notion of survival through nurturance and security. The emphasis in this theory is on how the need-related orientations (promotion vs. prevention) are strategically carried out rather than on the orientations themselves. In this way, it is like terror management theory's emphasis on the defensive systems based on sur-

vival needs and the apprehension of mortality rather than the on needs themselves.

Terror Management Theory

Terror management theory (TMT; Pyszczynski et al., 1997, 2000) assumes that *self-preservation* is the root need or master motive for all other needs, including those based on tissue deficits and the derived *symbolic-defensive* and *self-expansive* desires. This root need, survival, is said to be the overriding desired end state. But given the fact of mortality, this creates an existential crisis for humans who are perhaps uniquely aware of their inevitable demise (Becker, 1973; Freud, 1933; Rank, 1936/1976). "According to TMT, the self-preservation instinct—the goal of staying alive—is the superordinate goal toward which all behavior is oriented. All other motives are, in one way or another, derived from and subservient to their "prime directive" (Pyszczynski et al., 1997, p. 5).

The assumption that there is a single underlying motive from which all others are derived is similar to other single motive approaches (e.g., the hedonic assumption of approach and avoidance of basic learning theories; the rational self-interest assumption in economic analyses; and the reproductive/species survival assumption of evolutionary analyses). However, at the social-psychological level of analysis it has the interesting character of motivating primarily defensive behavior designed to distract from or comfort in the presence of mortality salience. This also could be considered to be a hierarchical structure of a different kind. The root need, combined with the knowledge of the inevitability of death, creates a uniquely human need to deal with the knowledge of one's own inevitable mortality, leading to two fundamental defensive needs.

TMT specifies three sets of motives. *Direct motives* include the need for food, water, and temperature regulation as well as instinctive reactions such as flinching from noise and recoiling from pain. But the bulk of the empirical research inspired by TMT is focused on the *defensive motives*:

> Instinct for self-preservation → fear of death
> → defensive motives:
>   (a) pursuit of self-esteem
>   (b) faith in the cultural worldview

The empirical research literature on these defensive motives is extensive, enough so to have its own *Handbook of Experimental Existential Psychology* (Greenberg, Koole, & Pyszczynski, 2004).

More recently a growth component has been added to TMT, perhaps in an attempt to address the ground covered by self-determination theory (SDT). Unlike SDT, these *self-expansive motives* are also connected to the root need for survival. Although fear of death apparently motivates only the symbolic defensive system, self-preservation, underlies all three systems.

> A human being with a capacity to do nothing other than maintain an interior homeostasis and defend against physi-

cal and psychological threats would have little chance of long-term survival in a complex and changing environment. Such an animal must also be inclined to explore, assimilate new information, and integrate that information with its existing conception of the world, because survival depends on the development of an adequate understanding of the environment and a complex set of skills for interaction with that environment. Thus it seems clear that a motive for growth and expansion of one's capacities (e.g., Deci & Ryan, 1991; Maslow, 1943; Rank, 1932/1989; White, 1959) would make good evolutionary sense in that it would greatly increase the animals *sic*) chances of surviving long enough to reproduce. (Pyszczynski et al., 1997, p. 6)

The addition of self-expansive motives to TMT creates some interesting conflicts and contradictions. An organism motivated to explore, grow, and expand is a risk-taking organism. Existential terror would seem to be at odds with such an inclination. One possible solution to this problem would be to adopt the perspective of evolutionary theory (Buss, 1997), in which reproductive success rather than individual survival is considered to be the fundamental driver and crucial selector in evolutionary change. Or one might argue, with Woodworth (1958), that people wish to survive so that they can go about their business, rather than going about their business in order to survive—and in evolutionary terms their business is to reproduce for the sake of their species. However, this view tends to subordinate what is supposed to be the master motive, self-preservation. The potential for tension between the defensive and self-expansive systems provides fertile ground for future research in TMT.

Check and Balance Structure:
Cognitive–Experiential Self-Theory

Cognitive–experiential self-theory (CEST) is a global theory of personality, heavily influenced by psychodynamic thinking (Epstein, 1992, 1993, 1994; Epstein & Pacini, 1999). Epstein proposes that people process information via two fundamentally different (although related) systems. These two systems—*rational* and *experiential*—function in the formation of schemas, and in the more general information processing required for individuals to adapt to their surroundings. Although these two systems may be combined in the completion of a single act, they operate according to different rules. The experiential system is affected predominantly by emotions, relying on intuition and heuristic cues; it functions automatically (without the conscious deliberation characteristic of the rational system), organizing experiences and directing behaviors. By contrast, the rational system is assumed to be wholly conscious and affect-free, driven instead by analytical thought and socially mediated knowledge. From a psychodynamic perspective, the influence of the experiential system on the rational system is parallel to that of the unconscious on rational, waking thought. However, Epstein's specification of the "preconscious" modifies the psychodynamic unconscious with the cognitive unconscious based on contemporary experimental psychology.

CEST assumes that there are four fundamental human needs: (1) to maximize pleasure and minimize pain; (2) to maintain a stable, coherent conceptual system for organizing experience; (3) to maintain relatedness to others; and (4) to maintain a positive sense of self-esteem. Different from the other theories in our comparison set, CEST assumes these four needs function in a system of checks and balances at both the rational and experiential levels, helping to keep behavior within adaptive limits. For example, if one need is fulfilled at the expense of the others, these other needs become more insistent, serving to moderate the strength of the first need. Consequently, behavior is influenced simultaneously by all four needs and tends not to be dominated by any one need in particular. The specification of checks and balances among needs is in essence the opposite of the assumptions of hierarchical theories, in which some needs take precedence over and may have to be satisfied before others. It also differs from the root need theories in that balance among needs rather than the precedence of a root need is assumed.

Evidence for the utility of CEST comes from studies on the nature of inferences in the two systems (Kirkpatrick & Epstein, 1992); on inconsistencies in the literature on depressive realism (Pacini, Muir, & Epstein, 1998); on the use of the theory in justice research (Krauss, Lieberman, & Olson, 2004; Lieberman, 2002); and in understanding individual differences as revealed in thinking in the two systems (Epstein, Pacini, Denes-Raj, & Heier, 1996) and in temperament (Teglasi & Epstein, 1998).

The distinction between a conscious, deliberative, and rational system as opposed to a nonconscious, intuitive, and emotional system is seen in a number of other contemporary theories. For example, Metcalfe and Mischel (1999) have proposed a distinction between "hot–cool," *know* and *go* systems that determine self-control. Strack and Deutsch (2004 and Chapter 17, this volume) proposed a *reflective–impulsive* model that is very similar in overall conception to the rational–experiential distinction, although Strack and Deutsch's model relies much more on associative network assumptions to understand the operation of the impulsive system. McClelland, Koestner, and Weinberger (1989) argued for a distinction between *implicit motives* and *self-attributed motives.* At the level of emotion and preference, Zajonc (1980) demonstrated that *preferences need no inferences* in a seminal paper that argued for a distinction between affective and cognitive systems. However, none of these other theories are about the set of fundamental needs. In that regard, it is the articulation of four basic motives operating among each other and across the cognitive–experiential divide that sets CEST apart from these other approaches.

Independent List Structure: Self-Determination Theory

Deci and Ryan (1980, 1985, 1991, 2000) proposed three basic human needs: *autonomy*, *competence*, and *relatedness*. Deci and Ryan (2000) have also provided a very clear elaboration of how they use the concept "need":

human needs specify the necessary conditions for psychological health or well-being and their satisfaction is hypothesized to be associated with the most effective functioning" and that "we assert that there are not instances of optimal, healthy development in which a need for autonomy, relatedness, or competence was neglected, whether or not the individuals consciously valued those needs. In short, psychological health requires satisfaction of all three needs; one or two are not enough. (p. 229)

We have classified this theory as one with an independent need structure because although it is clearly stated that all three needs must be satisfied, the theory does not specify any structural organization among those needs. There is no hierarchical structure, no root need that is said to be more basic or more important than the others, no system of checks and balances. All three needs must be satisfied for optimal functioning according to the theory, but each need has its independent set of requirements.

Over the past 30 years SDT has generated an extensive empirical literature, more so than any of the other theories we have chosen for this review (for reviews, see Deci, Koestner, & Ryan, 1999; Sansone & Harackiewicz, 2000), including its own *Handbook of Self-Determination Research* (Deci & Ryan, 2002). We have shaded the need for autonomy in Figure 20.2 because that is the need that gave the theory its name, and it is the one that has received the earliest and most consistent research attention. The work of deCharmes (1968) on internal and external perceived loci of control was part of the original thinking about the need for autonomy. The concept of autonomy can also be related to Brehm's (1966, 1993) work on reactance theory with its concept of *behavioral freedoms*. When a perceived behavioral freedom is eliminated or threatened with elimination, reactance motivation increases the attractiveness of that freedom and motivates the person to reestablish that freedom. In this context, SDT suggests that the set of behavioral freedoms needs to be more than zero for the individual to thrive. From another perspective, in the current context of theories of self-regulation and self-control (cf. Wegner & Wenzlaff, 1996), the need for autonomy might be taken to mean that individuals need to engage in autonomous self-regulatory activities to a sufficient extent or well-being will suffer. From this viewpoint, self-regulation is not only something that humans can do, it is something they need to do.

The need for competence, in the tradition of White's (1959) analysis of effectance motivation, refers to a need for effective interaction with the environment, and in a different sense to Murray's (1938) need for achievement. In the early versions of SDT, the combination of autonomy and competence was found to underlie intrinsic motivation (Deci, 1975) and both generated a considerable amount of research on the nature of competence motivation (Boggiano & Pittman, 1992; Elliot & Dweck, 2005). Research combining the individual difference approach to achievement motivation (e.g., McClelland & Watson, 1973) with an experimental approach to research on intrinsic and extrinsic motivational orientations (e.g.,

Pittman, Boggiano, & Ruble, 1983) has proven to be a particularly fruitful avenue for empirical research on the effects of intrinsic, achievement, and competence motivation (Harackiewicz, Manderlink, & Sansone, 1992; Harackiewicz, Sansone, & Manderlink, 1985; Sansone & Harackiewicz, 2000).

The need for relatedness has not received as much empirical attention, perhaps because it is a more recent addition to SDT. In the research on intrinsic motivation, relatedness has been studied through its role as a source of support for autonomy and competence (e.g., Ryan & Grolnick, 1986; Ryan, Stiller, & Lynch, 1994). Here there is a clear connection with attachment theory, particularly with Ainsworth & Wittig's (1969) use of the "Strange Situation" in research on the exploration–play system (see Figure 20.2). In this work, the attachment system provides a secure base for exploration. The security provided by satisfaction of the need for relatedness is seen in SDT as an important influence on the ability of persons to engage in the pursuit of autonomy and competence (Ryan & LaGuardia, 2000). Relatedness has played a more direct role in the analysis of the process of internalization. The extent to which extrinsically imposed rules and regulations, cultural prescriptions, and shared habits of thought and action become integrated and part of a person's way of satisfying intrinsic motivation depends very much on the nature of the relationship between the individual and socializing agents. Internalization is most likely when relatedness needs are satisfied in a way that also promotes feelings of autonomy and competence (Deci & Ryan, 2000; Grolnick & Ryan, 1989).

SDT makes clear predictions about the overall effects of satisfying or failing to satisfy basic needs on individual functioning. Satisfaction of the three basic needs should promote optimal development, while failure to satisfy any one of the needs should interfere with that development. These predictions have been tested in studies of the relation between need satisfaction and well-being. In addition to positive affect, Deci and Ryan (2000, p. 243) define well-being as "an organismic function in which the person detects the presence or absence of vitality, psychological flexibility, and a deep inner sense of wellness (Ryan, Deci, & Grolnick, 1995; Ryan & Frederick, 1997)." These predictions have been tested with measures of well-being in studies relating daily experiences with autonomy and competence (Sheldon, Ryan, & Reis, 1996), and with autonomy, competence, and relatedness (Reis, Sheldon, Gable, Roscoe, & Ryan, 2000).

The theory can be and has been criticized for being too optimistic about human nature (e.g., Pyszczynski et al., 2000), and also for not allowing sufficiently for the possibility of joy in the pursuit of extrinsic goals (e.g., Sansone, 2000).

LEVELS OF ANALYSIS IN THEORIES OF BASIC HUMAN NEEDS

By limiting themselves to three to five basic needs, all of these theories are comparable in their level of generality, in the sense of Maslow's metaphor of the nesting boxes.

But in other important ways, their comparability is problematic. One important dimension of difference lies in the level of analysis at which each of the proposed basic needs is assumed to operate. Scientific disciplines are distinguished by their predominant level of analysis. While physicists typically work at the level of atomic and subatomic particles, biologists are more likely to work at the level of the cell or the organism. Psychologists take the perspective of the individual, while sociologists and economists tend to develop their analyses at the level of social groups. Although all six of our social–psychological theories of basic human needs do take the perspective of the individual, within that perspective they still can be understood to vary in the levels of analysis they employ. To illustrate these kinds of differences, we discuss three different levels of analysis as they are represented across these six theories: basic or biological-level needs, needs operating at the level of the individual, and needs operating at the level of the individual in social groups (see Figure 20.3). We recognize and want to state clearly that this rather crude classification scheme runs the risk, perhaps the certainty, of distortion through oversimplification. Our purpose in using these broad classifications is primarily to illustrate that social–psychological theories of needs are not always operating at comparable levels of analysis.

Theoretical Comparisons within Levels of Analysis

Basic/Biological-Level Processes

At the level of basic or biological processes are needs that are probably not so different in their fundamental action across a wide variety of species. In addition to the need for food, water, temperature regulation, and oxygen, this level of analysis includes needs involving basic processes such as fight-or-flight mechanisms of survival and fundamental psychological processes of learning and change based on classical and instrumental conditioning. This is a level of analysis that is typically assumed to be important but not commonly chosen for study by social psychologists. It is, however, represented in several of the theories under consideration.

Maslow's hierarchy starts at its base with needs at the physiological level. These most fundamental needs are assumed to take precedence over all others. In addition, the need for safety and security can be considered to fall into this level of analysis, at least partially if not entirely. One of the four basic needs in CEST, pleasure/pain, also can be classified as a basic biological level need. In attachment theory, we could think of the fear-wariness system as operating at this level. In TMT, the survival motive and its resulting need for self-preservation also operates at this level. We think that two of our six theories, SDT and core social motives, do not address needs at this level of analysis (see Figure 20.3).

Although four of the six theories do specify needs at this level, none of them has focused empirically on those needs. The three theories that have generated a substantial amount of research have not done so at this level. That is to be expected given that these are the theories of
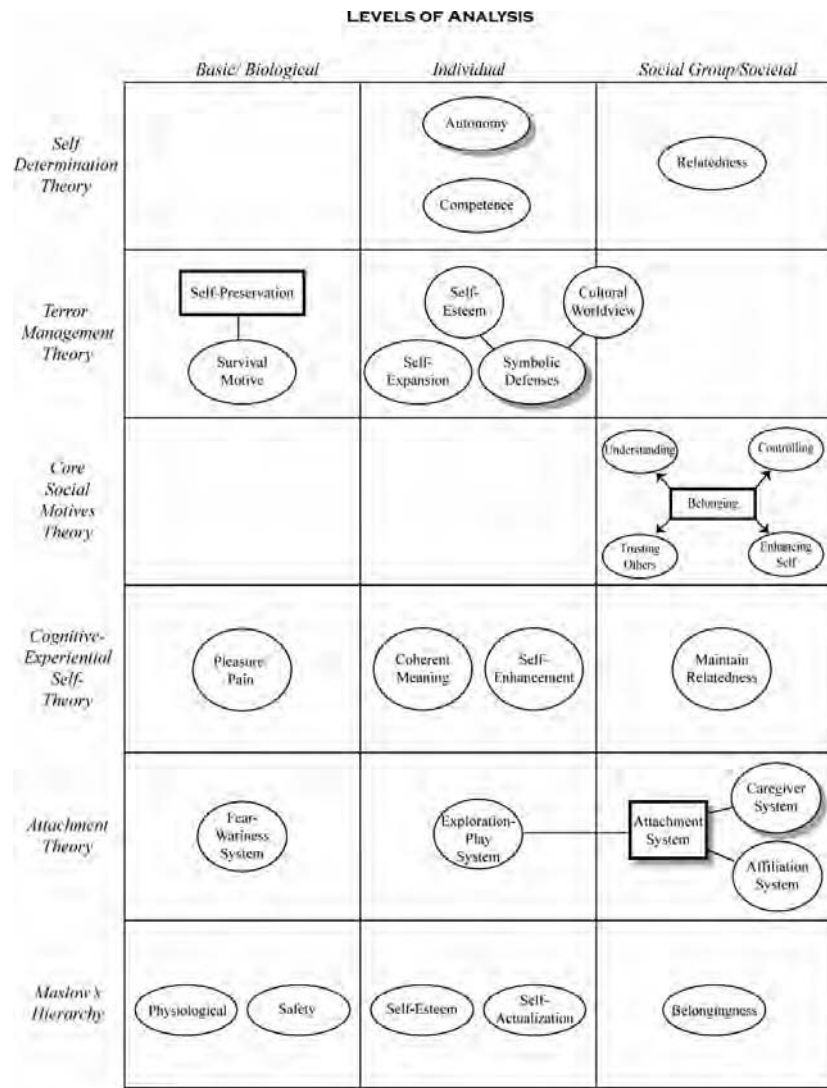
LEVELS OF ANALYSIS

FIGURE 20.3. Levels of analysis of six theories of basic human needs. Shadowing indicates the needs that have generated the most empirical research.

social and personality psychologists. However, in two of the theories the needs at the basic/biological-level of analysis are given particular theoretical importance. In Maslow's hierarchy, it is the level of need that takes precedence over all others. In TMT, it is the home of the "master motive," survival. None of the other theories give such fundamental precedence to needs at the basic/biological level, instead locating their most important or root needs at the individual or the social group level of analysis.

Individual-Level Processes

More familiar in social–psychological theories in general are processes assumed to operate at the level of the individual. Many traditional motivational theories in social psychology such as cognitive dissonance theory (Festinger, 1957) and other consistency-based ap-

proaches (cf. Abelson et al., 1968) could be classified as individual-level analyses. The implied assumption here is that although these processes are embedded in each person's social world, and therefore affect and are affected by others, they are fundamental aspects of individual human functioning that would be present and important to understand even in the absence of social considerations. Indeed, the intensive study of such processes sometimes leads to questions and arguments about what it is that is social about such research (in other words, is it really social psychology).

Not surprisingly, five of the six theories specify basic needs at this level of analysis. In Maslow's hierarchy, self-esteem and self-actualization, the two needs at the top of the hierarchy, are individual-level needs. In attachment theory, the exploration–play system is where individuals express their curiosity about the environment. In CEST coherent meaning and self-enhancement are individual-

level needs. TMT includes the symbolic defenses for managing existential terror and the self-expansion system at this level of analysis. SDT has two of its three basic needs, autonomy and competence, at the individual level of analysis. Only core social motives theory makes no claims about individual-level needs (see Figure 20.3).

Also not surprisingly, this is the level at which most of the empirical research generated by three of these theories has been done. SDT has generated the most research on the needs for autonomy and competence. TMT has focused primarily on the operation of the symbolic defenses in the face of mortality salience as they are expressed in self-esteem and embracing the cultural worldview. We have placed the cultural worldview to straddle the individual and social group levels, but primarily the research focus has been on how and when individuals use their version of the cultural worldview to manage their personal terror. The empirical research guided by CEST has also been concentrated on the needs for coherent meaning and for self-enhancement. It is probably fair to characterize these three theories as primarily individual-level approaches, at least in terms of how they have been tested in the research literature.

One common thread worth noting is that in attachment theory, Maslow's hierarchy, and SDT, security of attachment, belongingness, or relatedness does or can provide the basis for effective satisfaction of needs at the individual level: exploration–play in attachment theory, self-esteem and self-actualization in Maslow's hierarchy, and at least as represented in some of the research on relatedness inspired by SDT, for satisfaction of autonomy and competence needs.

Social-Level Processes

At the social level of analysis are needs that depend on the individual being embedded in a social environment. They are still processes operating within an individual, but they depend on and are oriented toward social groups. Social comparison theory (Festinger, 1954) could be considered to be an early example of a theory oriented toward this level of analysis. Theories of social identity (e.g., Brewer, 1991; Brewer & Pickett, 1999; Tajfel & Turner, 1979; see Deaux, 1996, for a review) and the need for shared reality (Asch, 1951, 1956; Hardin & Higgins, 1996; Sherif, 1935, 1936) are also examples of theorizing at this level. The distinction between this level of analysis and that of the individual is that the social group is required for such processes to operate and make sense.

All six theories have at least one need or basic process at this level. Maslow's hierarchy specifies a need for belongingness in the middle of the hierarchy. CEST includes a need to maintain relatedness, and SDT also includes a need for relatedness. Attachment theory is primarily about the social level of analysis, including the attachment, affiliation, and caregiver systems. Core social motives theory is all about the social level of analysis, with all five needs related to functioning in social groups. Of all the theories, TMT is least oriented to this level of analysis. Maintaining a cultural worldview requires a community from which to learn and with which to share this source of comfort in the face of the knowledge of mortality, but we think that the generation of the shared culture from the combination of many individual personal existential problems and the waxing and waning of the use of defensive terror management mechanisms with the salience of mortality within the person still pulls the theoretical emphasis toward the individual level of analysis (see Figure 20.3).

Although there seems to be a great deal of agreement about what is important at this level of analysis, only attachment theory has generated much actual research on such needs; in addition to the extensive developmental work on attachment and its associated systems in infancy there is a burgeoning literature on adult attachment processes in close relationships (Reis & Patrick, 1996; Rholes & Simpson, 2004). Maslow's hierarchy and core social motives theory have not been used to generate empirical research; relatedness is the least intensively studied need in both SDT and CEST; and TMT is primarily oriented to the individual level, although there is a considerable amount of research on when individuals will embrace cultural values. The potential generative impact of assuming that there is a need for relatedness or belongingness (Baumeister & Leary, 1995) is not yet apparent in most of these theories.

Comparing theories with attention to differences in level of analysis helps to reveal, at least in part, why they are so different in the needs that they specify, and perhaps also why their theoretical structures are so different. Comparing an essentially social group/societal-level theory (e.g., core social motives) with an essentially individual-level theory (e.g., SDT or TMT) is more difficult than might otherwise be apparent in part because of these differences in levels of analysis.

Theoretical Comparisons across Levels of Analysis

Briefly comparing the theories across levels of analysis, Figure 20.3 reveals that the range of levels represented within individual theories varies considerably. Maslow's hierarchy and CEST both cover the range from basic/biological to social group/societal. In contrast, core social motives theory is completely contained at the level of the social group. Attachment theory, although not entirely at the social level of analysis, has its root motive and the systems that have been most heavily researched at that level. SDT and TMT, if judged by the research they have generated, are essentially individual-level theories.

As an example of what can be noticed by attending to levels of analysis, we note that in Maslow's theory as we go up the hierarchy, we don't go up the levels of analysis in a linear fashion. Instead, we move from the basic/biological level to the social group level, and only then to the individual level where self-esteem and self-actualization are at the top of the hierarchy. Seeing belongingness as the foundation for self-esteem and self-actualization is consistent with attachment theory and perhaps with SDT, but would seem to be at odds with TMT. Another concordance among attachment theory, SDT, and Maslow's hierarchy can be seen at the individual level of

analysis, where exploration and play, autonomy and competence, and self-actualization appear to be addressing comparable dimensions.

Positive and Negative Psychology?

SDT and TMT have generated a considerable amount of research. Both are essentially individual level of analysis models. But their assumptions about the nature of basic human needs and human nature are polar opposites. In the recently proposed language of positive psychology (e.g., Seligman & Csikszentmihalyi, 2000) and, by implication, negative psychology (e.g., Seligman, 1975), SDT is positive about human nature. On the other hand, if a Positive Psychology Hall of Fame were ever created, TMT would be unlikely to occupy a place of prominence.

> . . . it is part of the adaptive design of the human organism to engage interesting activities, to exercise capacities, to pursue connectedness in social groups, and to integrate intrapsychic and interpersonal experiences into a relative unity. (Deci & Ryan, 2000, p. 229)

> . . . many psychological needs are ultimately rooted in the existential dilemma into which our species was born." "Knowledge of the inevitability of death gives rise to the potential for paralyzing terror, which would make continued goal-directed behavior impossible." ". . . this terror is managed by a dual-component cultural anxiety buffer, consisting of . . . (a) an individual's personalized version of the cultural worldview, . . . and (b) self-esteem or a sense of personal value . . . (Pyszczynski et al., 1997, p. 2)

Perhaps a better to way to describe the difference between the two theories would be to say that SDT is a growth model while TMT is a defensive model. One is based on self-actualization, the other on managing anxiety. Both approaches have a long tradition in psychology, and both are probably capturing important aspects of human nature. Self-expansive motives have been added to TMT, but they are still justified as survival based, and do not have the optimistic self-actualizing quality of SDT's approach. The darker and brighter aspects of SDT and TMT have been discussed by the principles (Pyszczynski et al., 2000; Ryan & Deci, 2000), and further comparisons and contrasts, as well as empirical tests, are likely.

Basic Needs: Where Do We Stand?

Although it is tempting to try to come up with *the* basic set of human needs, we think that the theoretical differences we have revealed in structure and in levels of analysis make any simple attempt at synthesis of these theories premature. One might try thinking about, for example, the relations between autonomy and competence (SDT) and understanding and control (core social motives), but these theories are operating at two different levels of analysis. One could similarly wish to compare the need for self-esteem (TMT) and self-enhancement (CEST) and perhaps to think they might be the same thing. But in one theory, self-esteem is a derivative of the core motive, while in the other it is one of four equally balanced needs

maintained across a cognitive–experiential divide. Probably the differences in theoretical structure and in levels of analysis will need to be addressed before an overall integration or winnowing would be possible. Even so, some commonalities can be discerned. The most obvious of these is that five of the six theories share, at the level of the social group, an assumed need for relatedness, belonging, or affiliation.

Are theorists and researchers working on evaluating or integrating these disparate theories of basic human needs? Theoretical discussions and empirical comparisons among several of these theories have begun to emerge. At the empirical level, there have been comparisons or integrations of TMT and CEST (Simon et al., 1997), attachment theory and TMT (Mikulincer & Florian, 2000), and SDT, attachment theory, and TMT (Arndt, Routledge, Greenberg, & Sheldon, 2005). CEST has been used to examine self-verification theory and self-enhancement (Morling & Epstein, 1997; Swann, 1984, 1990; Swann, Stein-Seroussi, & Geisler, 1992), while an SDT-inspired approach to self-concordance uses a rational–experiential consistency model with features similar to CEST (Sheldon & Elliot, 1999; Sheldon et al., 2004). And at the theoretical level, discussions among SDT and TMT, as well as several other self-theories (Crocker & Park, 2004; Leary, 2004) have been published (Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004a, 2004b; Ryan & Deci, 2004). There have also been several investigations of cross-cultural evidence for the existence of a set of universal human needs (Grouzet et al., 2005; Sheldon et al., 2001; Sheldon et al., 2004).

In keeping with the beginning of this chapter, we should ask to what extent these theories tell us what it is that is distinctive about human nature. Have we learned how humans differ from other organisms? Many of the basic needs proposed by these various theories could easily be characteristic of other species as well. For example, basic attachment processes may not be uniquely human, nor perhaps are needs related to living in social groups, so in that sense we may not have made much progress in learning what is distinctive about human beings. TMT is based on what may be a unique, or if not unique at least unusual, ability to understand the concept of mortality and to deal with it through defensive symbolism. If self-reflexive consciousness is uniquely human (Higgins, 2005; Terrace & Metcalfe, 2005), then the rational–experiential distinction of CEST may also be uniquely human. If self-reflexive self-regulation is uniquely human, then the complex relations among autonomy, competence, and relatedness that play out in the extent to which reasons for task engagement are external, introjected, integrated, or internal (SDT) may be a uniquely human issue. It is an interesting and open question to ask more generally if self-actualization and the desire for autonomy and competence are uniquely human. The question of whether the best way, or even a good way, to understand what is uniquely human is through understanding basic human needs is still open.

Is the specification of basic human needs a useful psychological approach? On their own, several theories based on such specifications have led to extensive pro-

grams of empirical research, most notably SDT and
TMT. Although there is at this time considerable dis-
agreement among the various theories of basic human
needs, this disagreement also has had a generative im-
pact on theory and research. Much work remains to be
done. Differences in theoretical structure, levels of analy-
sis, and the set of basic human needs will need to be ad-
dressed. But it seems likely that this first handbook re-
view of *basic human needs* will not be the last.

REFERENCES

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, T. M. (1968). *Theories of cognitive consistency: A sourcebook*. Skokie, IL: Rand McNally.

Ainsworth, M. D. S., & Wittig, B. A. (1969). Attachment and exploratory behavior of one-year-olds in a strange situation. In B. Foss (Ed.), *Determinants of infant behaviour* (Vol. 4, pp. 129–173). New York: Barnes & Noble.

Arndt, J., Routledge, C., Greenberg, J., & Sheldon, K. M. (2005). Illuminating the dark side of creative expression: Assimilation needs and the consequences of creative action following mortality salience. *Personality and Social Psychology Bulletin*, *31*, 1327–1339.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership, and men.* Pittsburgh, PA: Carnegie Press.

Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, *70*(9, Whole No. 416).

Atkinson, J. W. (Ed.). (1958). *Motives in fantasy, action, and society*. Princeton, NJ: Van Nostrand.

Atkinson, J. W., & Raynor, J. O. (1978). *Personality, motivation, and achievement*. New York: Wiley.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Becker, E. (1973). *The denial of death*. New York: Free Press.

Berlyne, D. D. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.

Boggiano, A. K., & Pittman, T. S. (Eds.) (1992). *Achievement and motivation: A social–developmental analysis*. New York: Cambridge University Press.

Bowlby, J. (1969). *Attachment and loss: Volume 1: Attachment*. London: Hogarth Press.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Brehm, J. W. (1993). Control, its loss, and psychological reactance. In G. Weary, F. Gleicher, & K. Marsh (Eds.), *Control motivation and social cognition* (pp. 3–30). New York: Springer-Verlag.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Brewer, M. B., & Pickett, C. L. (1999). Distinctiveness motives as a source of the social self. In T. Tyler, R. Kramer, & O. John (Eds.), *The psychology of the social self* (pp. 71–87). Mahwah, NJ: Erlbaum.

Buss, D. M. (1997). Human social motivation in evolutionary perspective: Grounding terror management theory. *Psychological Inquiry*, *8*, 22–26.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Carver, C. S., & Scheier, M. F. (2000). Autonomy and self-regulation. *Psychological Inquiry*, *11*, 284–291.

Cofer, C. N., & Appley, M. H. (1964). *Motivation: Theory and research.* New York: Wiley.

Costa, P. T., & McCrae, R. R. (1988). From catalogue to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, *55*, 258–265.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, *130*, 392–414.

deCharmes, R. (1968). *Personal causation*. New York: Academic Press.

Deaux, K. (1996). Social identification. In E. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 777–798). New York: Guilford Press.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.

Deci, E. L., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 39–80). New York: Academic Press.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. Dientsbier (Ed.), *Nebraska Symposium on Motivation: Vol. 38. Perspectives on motivation* (pp. 237–288). Lincoln: University of Nebraska Press.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Deci, E. L., & Ryan, R. M. (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Deeg, D., J., van Zonneveld, R. J., van der Maas, P. J., & Habbema, J. D. (1989). Medical and social predictors of longevity in the elderly: Total predictive value and interdependence. *Social Science and Medicine*, *29*, 1271–1280.

Doob, L. W. (1947). The behavior of attitudes. *Psychological Review*, *54*, 135–156.

Dweck, C. S. (1996). Implicit theories as organizers of goals and behavior. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 69–90). New York: Guilford Press.

Elliott, A. J., & Dweck, C. S. (2005). *Handbook of competence and motivation.* New York: Guilford Press.

Epstein, S. (1992). Coping ability, negative self-evaluation, and overgeneralization: Experiment and theory. *Journal of Personality and Social Psychology*, *62*, 826–836.

Epstein, S. (1993). Bereavement from the perspective of cognitive-experiential self-theory. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of bereavement* (pp. 112–125). Cambridge, UK: Cambridge University Press.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, *49*, 709–724.

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive–experiential self-theory. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462–482). New York: Guilford Press.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical-rational thinking. *Journal of Personality and Social Psychology*, *71*, 390–405.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58* 203–210.

Fiske, S. T. (2002). Five core social motives, plus or minus five. In S.

Spencer, S. Fein, M. Zanna, & J. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 233–246). Mahwah, NJ: Erlbaum.

Fiske, S. T. (2004). *Social beings: A core motives approach to social psychology*. New York: Wiley.

Freud, S. (1933). *New introductory lectures on psycho-analysis*. New York: Norton.

Gilbert, D., Fiske, S. T., & Lindzey, G. (1998). *The handbook of social psychology* (4th ed.). Boston: McGraw-Hill.

Gove, P. (Ed.). (1963). *Webster's seventh new collegiate dictionary*. Springfield, MA: G. & C. Merriam.

Gollwitzer, P. M., & Moskowitz, G. B. (1996). Goal effects on action and cognition. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399). New York: Guilford Press.

Greenberg, J., Koole, S. L., & Pyszczynski, T. (2004). *Handbook of experimental existential psychology*. New York: Guilford Press.

Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, *81*, 143–154.

Grouzet, F. M., Kasser, T., Ahuvia, A., Dols, J. M. F., Kim, Y., et al. (2005). The structure of goal contents across 15 cultures. *Journal of Personality and Social Psychology*, *89*, 800–816.

Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1992). Competence processes and achievement motivation: Implications for intrinsic motivation. In A. K. Boggiano & T. S. Pittman (Eds.), *Achievement and motivation: A social–developmental analysis* (pp. 115–137). New York: Cambridge University Press.

Harackiewicz, J. M., Sansone, C., & Manderlink, G. (1985). Competence, achievement orientation, and intrinsic motivation. *Journal of Personality and Social Psychology*, *48*, 493–508.

Hardin, C. D., & Higgins, E. T. (1996). How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Volume 3. The interpersonal context* (pp. 28–84). New York: Guilford Press.

Harlow, H. F. (1953). Mice, monkeys, men, and motives. *Psychological Review*, *60*, 23–32.

Hazan, C., & Shaver, P. R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry*, *5*, 1–22.

Henchy, T., & Glass, D. C. (1968). Evaluation apprehension and the social facilitation of dominant and subordinate responses. *Journal of Personality and Social Psychology*, *10*, 446–454.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1989). Self-discrepancy theory: What patterns of self-beliefs cause people to suffer? In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 93–136). San Diego, CA: Academic Press.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). San Diego, CA: Academic Press.

Higgins, E. T. (1999). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 503–525). New York: Guilford Press.

Higgins, E. T. (2005). humans as applied motivation scientists: Self-consciousness from "shared reality" to "becoming." In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflexive consciousness* (pp. 157–173). New York: Oxford University Press.

Higgins, E. T., & Kruglanski, A. W. (Eds.). (1996). *Social psychology: A handbook of basic principles*. New York: Guilford Press.

Higgins, E. T., & Kruglanski, A. W. (Eds.). (2000). *Motivational science: Social and personality perspectives*. Philadelphia: Psychology Press.

Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

Hunt, J. M. (1965). Intrinsic motivation and its role in psychological development. *Nebraska Symposium on Motivation Series*, *13*, 189–282.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd ed.). Port Huron, MI: Research Psychologists Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrika*, *47*, 263–291.

Kirkpatrick, L. A., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: Further evidence for two conceptual systems. *Journal of Personality and Social Psychology*, *63*, 534–544.

Krauss, D. A., Lieberman, J. D., & Olson, J. (2004). The effects of rational and experiential information processing of expert testimony in death penalty cases. *Behavioral Sciences and the Law*, *22*, 801–822.

Kruglanski, A. W. (1996). Motivated social cognition: Principles of the interface. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 493–520). New York: Guilford Press.

Leary, M. R. (2004). The function of self-esteem in terror management theory and sociometer theory: Comment on Pyszczynski et al. (2004). *Psychological Bulletin*, *130*, 478–482.

Lieberman, J. D. (2002). Head over heart or heart over head? Cognitive-experiential self-theory and extralegal heuristics in juror decision making. *Journal of Applied Social Psychology*, *32*, 2526–2553.

Lindzey, G. (1954). *Handbook of social psychology*. Cambridge, MA: Addison-Wesley.

Lindzey, G., & Aronson, E. (1968). *The handbook of social psychology* (2nd ed.). Reading, MA: Addison-Wesley.

Lindzey, G., & Aronson, E. (1985). *The handbook of social psychology* (3rd ed.). New York: Random House.

Liu, T. J., & Steele, C. M. (1986). Attribution analysis as self affirmation. *Journal of Personality and Social Psychology*, *51*, 531–540.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396.

Maslow, A. H. (1970). *Motivation and personality*. New York: Harper & Row.

McClelland, D. C. (1958). Risk-taking in children with high and low need for achievement. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 306–321). Princeton, NJ: Van Nostrand.

McClelland, D. C., Clark, R. A., Roby, T. B., & Atkinson, J. W. (1949). The projective expression of needs. IV: The effect of need for achievement and thematic apperception. *Journal of Experimental Psychology*, *39*, 242–255.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702.

McClelland, D. C., & Watson, R. I. (1973). Power motivation and risk-taking behavior. *Journal of Personality*, *41*, 121–139.

McDougall, W. (1926). *An introduction to social psychology* (rev. ed.). Boston: Luce.

Mencken, H. L. (1949). *A Mencken chrestomathy*. New York: Knopf.

Metcalfe, J., & Mischel, W. (1999). A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.

Mikulincer, M., & Florian, V. (2000). Exploring individual differences in reactions to mortality salience: Does attachment style regulate terror management mechanisms? *Journal of Personality and Social Psychology*, *79*, 260–273.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.

Morling, B., & Epstein, S. (1997). Compromises produced by the dialectic between self-verification and self-enhancement. *Journal of Personality and Social Psychology*, *73*, 1268–1283.

Murchison, C. (1935). *A handbook of social psychology*. Worcester, MA: Clark University Press.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Pacini, R., Muir, F., & Epstein, S. (1998). Depressive realism from the perspective of cognitive-experiential self-theory. *Journal of Personality and Social Psychology*, *74*, 1056–1068.

Pavlov, I. P. (1927). *Conditioned reflexes*. New York: Oxford University Press.

Pittman, T. S. (1993). Control motivation and attitude change. In G. Weary, F. Gleicher, & K. Marsh (Eds.), *Control motivation and social cognition* (pp. 157–175). New York: Springer-Verlag.

Pittman, T. S. (1998). Motivation. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 549–590). Boston: McGraw-Hill.

Pittman, T. S., Boggiano, A. K., & Ruble, D. N. (1983). Intrinsic and extrinsic motivational orientations: Limiting conditions on the undermining and enhancing effects of reward on intrinsic motivation. In J. Levine & M. Wang (Eds.), *Teacher and student perceptions: implications for learning* (pp. 319–340). Hillsdale, NJ: Erlbaum.

Pittman, T. S., & D'Agostino, P. R. (1989). Motivation and cognition: Control deprivation and the nature of subsequent information processing. *Journal of Experimental Social Psychology*, *25*, 465–480.

Pittman, T. S., & Heller, J. F. (1987). Social motivation. In M. Rosenzweig & L. Porter (Eds.), *Annual review of psychology* (Vol. 38, pp. 461–489). Palo Alto, CA: Annual Reviews.

Pyszczynski, T., Greenberg, J., & Solomon, S. (1997). Why do we need what we need? A terror management perspective on the roots of human social motivation. *Psychological Inquiry*, *8*, 1–20.

Pyszczynski, T., Greenberg, J., & Solomon, S. (2000). Toward a dialectical analysis of growth and defensive motives. *Psychological Inquiry*, *11*, 301–305.

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J., & Schimel, J. (2004a). Converging toward an integrated theory of self-esteem: Reply to Crocker and Nuer (2004), Ryan and Deci (2004), and Leary (2004). *Psychological Bulletin*, *130*, 483–488.

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J., & Schimel, J. (2004b). Why do people need self-esteem? A theoretical and empirical review. *Psychological Bulletin*, *130*, 435–468.

Rank, O. (1976). *Will therapy and truth and reality*. New York: Knopf. (Original work published 1936)

Reis, H. T., & Patrick, B. C. (1996). Attachment and intimacy: Component processes. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (523–563). New York: Guilford Press.

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, *26*, 419–435.

Rholes, W. S., & Simpson, J. A. (2004). *Adult attachment: Theory, research, and clinical implications.* New York: Guilford Press.

Ryan, R. M., & Deci, E. L. (2000). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Inquiry*, *11*, 319–338.

Ryan, R. M., & Deci, E. L. (2004). Avoiding death or engaging life as accounts of meaning and culture: Comment on Pyszczynski et al. (2004). *Psychological Bulletin*, *130*, 473–477.

Ryan, R. M., Deci, E. L., & Grolnick, W. S. (1995). Autonomy, relatedness, and the self: Their relation to development and psychopathology. In D. Cicchetti & J. Cohen (Eds.), *Developmental psychopathology* (Vol. 1, pp. 618–655). New York: Wiley.

Ryan, R. M., & Frederick, C. M. (1997). On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality*, *65*, 529–565.

Ryan, R. M., & Grolnick, W. S. (1986). Origins and pawns in the classroom: Self-report and projective assessments of individual difference in children's perceptions. *Journal of Personality and Social Psychology*, *50*, 550–558.

Ryan, R. M., & LaGuardia, L. G. (2000). What is being optimized over development? A self-determination theory perspective on basic psychological needs across the life span. In S. Qualls & R. Abeles (Eds.), *Dialogues on psychology and aging* (pp. 145–172). Washington, DC: American Psychological Association.

Ryan, R. M., Stiller, J., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *Journal of Early Adolescence*, *14*, 226–249.

Sansone, C. (2000). The "how" of goal pursuit: Interest and self-regulation. *Psychological Inquiry*, *11*, 306–309.

Sansone, C., & Harackiewicz, J. M. (2000). *Intrinsic and extrinsic motivation; The search for optimal motivation and performance*. San Diego, CA: Academic Press.

Schachter, S. (1959). *The psychology of affiliation; Experimental studies of the sources of gregariousness*. Stanford, CA: Stanford University Press.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55*, 5–14.

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*, 482–497.

Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology*, *80*, 325–339.

Sheldon, K. M., Elliot, A. J., Ryan, R. M., Chirkov, V., Kim, Y., Wu, C., et al. (2004). Self-concordance and subjective well-being in four cultures. *Journal of Cross-Cultural Psychology*, *35*, 209–223.

Sheldon, K. M., Ryan, R. M., & Reis, H. T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin*, *22*, 1270–1279.

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, No. 187.

Sherif, M. (1936). *The psychology of social norms.* New York: Harper.

Simon, L., Greenberg, J., Harmon-Jones, E., Solomon, S., Pyszczynski, T., Arndt, J., et al. (1997). Terror management theory and cognitive-experiential self-theory: Evidence that terror management occurs in the experiential system. *Journal of Personality and Social Psychology*, *72*, 1132–1146.

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1981). Selection by consequences. *Science*, *213*, 501–504.

Soloman, S., Greenberg, J., & Pyszczynski, T. (1997). Return of the living dead. *Psychological Inquiry*, *8*, 59–71.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 261–302). San Diego, CA: Academic Press.

Steele, C. M., & Liu, T. J. (1981). Making the dissonant act unreflective of the self: Dissonance avoidance and the expectancy of a value-affirming response. *Personality and Social Psychology Bulletin*, *7*, 393–397.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*, 5–19.

Stevens, L. E., & Fiske, S. T. (1995). Motivation and cognition in social life: A social survival perspective. *Social Cognition*, *13*, 189–214.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Bulletin*, *8*, 220–247.

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review*, *91*, 457–477.

Swann, W. B., Jr. (1990). To be adored or to be known? The interplay of self-enhancement and self-verification. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition:*

Vol. 2. Foundations of social behavior (pp. 408–448). New York: Guilford Press.

Swann, W. B., Jr., Stein-Seroussi, A., & Geisler, R. B. (1992). Why people self-verify. *Journal of Personality and Social Psychology, 62*, 392–401.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Teglasi, H., & Epstein, S. (1998). Temperament and personality theory: The perspective of cognitive-experiential self-theory. *School Psychology Review, 27*, 534–550.

Terrace, H. S., & Metcalfe, J. (2005). *The missing link in cognition: Origins of self-reflexive consciousness.* New York: Oxford University Press.

Vallerand, R. J. (2000). Deci and Ryan's self-determination theory: A view from the hierarchical model of intrinsic and extrinsic motivation. *Psychological Inquiry, 11*, 312–318.

Watson, J. B. (1930). *Behaviorism.* New York: Norton.

Weary, G., Gleicher, F., & Marsh, K. L. (1993). *Control motivation and social cognition.* New York: Springer-Verlag.

Wegner, D. M., & Wenzlaff, R. M. (1996). Mental control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 466–492). New York: Guilford Press.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*, 297–333.

Winter, D. G. (1973). *The power motive.* New York: Free Press.

Woodworth, R. L. (1958). *Dynamics of behavior.* New York: Holt.

Zajonc, R. B. (1965). Social facilitation. *Science, 149*, 269–274.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.

The Goal Construct
in Social Psychology

AYELET FISHBACH
MELISSA J. FERGUSON

Goals constitute the focal points around which human behavior is organized. Much of what people think about, feel, and do revolves around the goals they are trying to meet, or those goals they have already met or dismissed. Goals can influence major life decisions such as choosing one's career path, as well as more mundane everyday choices, such as which book to read. Goals guide one's behavioral responses to the social environment, such as whether one responds to a provocation by being competitive, collaborative, or resigned, for instance. And goals, and the ways in which people pursue them, also determine people's evaluations, moods, and emotional experience both during a pursuit and after a pursuit has been completed or abandoned. The scholarship on goals in social psychology has reflected the centrality of goals in people's lives, and consequently the goal construct has been defined, examined, and challenged, iteratively, throughout almost the entire century of empirical psychology (e.g., Ach, 1935; Atkinson, 1964; Austin & Vancouver, 1996; Bandura, 1986; Bargh, 1990; Carver & Scheier, 1998; Deci & Ryan, 1985; Gollwitzer, 1990; Higgins, 1997; James, 1890; Kruglanski, 1996; Lewin, 1926; Locke & Latham, 1990; Mischel, Cantor, & Feldman, 1996).

In the current chapter, we propose a contemporary framework for understanding what goals are and how they influence human experience and behavior. In particular, we address how goals are activated, the characteristics of their operation, and the ways in which they interact with one another. We anchor the framework with a set of definitional assumptions about the structure and content of goals. In support of our framework, we draw primarily on research conducted over the last decade that is characterized by its social-cognitive approach. By adopting this approach to the study of goals we also emphasize the implicit nature of motivation, including the ways in which goals can become activated outside conscious intention and operate according to a variety of implicit mechanisms. This stands in contrast with much of the traditional research on goals, which has focused on the conscious processes involved in setting a goal and striving toward its completion (e.g., Carver & Scheier, 1981; Gollwitzer, 1999; Locke & Latham, 1990).

We organize the chapter into four major parts. We consider in the first part ("What Is a Goal?") a working definition of goals as well as a set of assumptions underlying goal research. We then move onto the second part ("On the Activation of a Goal"), which considers theory and findings on the determinants of goal activation. We discuss in the third part ("On the Operation of a Goal") the various characteristics of active goal operation that involve goal-related knowledge activation, evaluations, and affective experience. In the fourth and final part ("On the Interaction among Goals"), we turn to an arguably more realistic view of goals—one that assumes that people are constantly switching their attention and motivation from goal to goal, depending on a host of situational and personal variables (e.g., Atkinson & Birch, 1970). Any given goal pursuit potentially interferes with other possible pursuits, and thus we examine the special

challenges that simultaneous goal pursuits pose and the ways in which interaction and interdependence among goals occur. Our broadest objective in this chapter is to develop a goal framework that both grounds previous work as well as generates new questions and research directions.

WHAT IS A GOAL?

We define a goal as *a cognitive representation of a desired endpoint that impacts evaluations, emotions and behaviors.* Aspects of this definition have been echoed in goal literature throughout the past 50 years (e.g., see Carver & Scheier, 1981; Gollwitzer & Moskowitz, 1996; Higgins & Kruglanski, 2000; Locke & Latham, 1990; Sorrentino & Higgins, 1986). In what follows we explicitly consider a set of more detailed assumptions about goals that underlie this definition and much of the recent work on this topic. These assumptions can be organized into those that concern the *structure* of a goal in memory versus those that involve the *content* of goal representations.

The Structure of Goals

Researchers have long assumed that goals exist as cognitive representations in memory (Bargh, 1990; Hull, 1931; Kruglanski, 1996; Tolman, 1932), even if various theoretical treatments of goals over the last century have varied in terms of explicitly mentioning this point. We argue that although there is a general consensus that goals exist in memory, an explicit consideration of this point inevitably leads to certain implications, which have not been as widely discussed or tested. The fact that goals exist as knowledge structures suggests (at least) three characteristics. First, as a memory construct, a goal necessarily *fluctuates in accessibility* (i.e., its activation potential; Higgins, 1996). This means that the likelihood of the goal being activated will vary across time and situations according to its accessibility at the moment.

Another characteristic concerns the *multiple memories* underlying any given goal. In particular, rather than a goal consisting of a unitary, discrete construct, it consists of a wide array of interconnected memories that are related to that goal (e.g., means of attainment and opportunities) and become associated with one another through a variety of ways. For example, the interconnection among memories underlying the goal of riding a bike might develop through direct experience (e.g., when the bike tilts left, shift weight to the right) as well as semantic and episodic knowledge (e.g., bike riding is a form of exercise and recreation perfect for a sunny summer afternoon).

The fact that goals consist of many memories that are interconnected naturally leads to the third characteristic of goals. Namely, the memories of a goal become activated according to classical *knowledge activation* processes (Anderson, 1983; Anderson & Reder, 1999; Collins & Loftus, 1975; Neely, 1977, 1991; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). In particular, it has long been postulated that the activation of a given memory will influence the activation of those memories with which it is connected. The nature of this influence can be either excitatory or inhibitory. With excitatory connections, as one memory of a goal construct becomes activated, and therefore, relatively more accessible, those memories interconnected with it should become activated and accessible as well. In this way, making one component of a goal construct more accessible can render much of the construct as a whole more accessible. For instance, the activation of a single memory concerning the goal of achievement could automatically lead to the activation (i.e., greater accessibility) of many other memories associated with achievement (see also research on stereotype activation; e.g., Bargh, Chen, & Burrows, 1996; Devine, 1989). But other connections among goal memories are inhibitory in nature, such that the activation of one goal automatically leads to the inhibition (i.e., lower accessibility) of another, competing goal. For instance, the activation of a central goal (e.g., academic achievement) might inhibit another tempting goal (e.g., partying—Fishbach, Friedman, & Kruglanski, 2003; Shah, Friedman, & Kruglanski, 2002).

Notably, the link between any two memories may not be bidirectional. Just because the activation of one goal memory can render accessible an associated memory, the same facilitative effect may not emerge in the reverse direction. For example, when considering the relationships among competing goals, whereas an immediately tempting goal can activate an overriding, more important goal, the reverse is not necessarily true. In fact, some recent research suggests that the same important goal might actually inhibit the tempting one (see Fishbach et al., 2003). In this sense, the connection among any two goal-related memories cannot be inferred merely on the basis of how one memory influences the activation of the other.

These three characteristics of goal structure (i.e., varying accessibility, multiple memories, and excitatory and inhibitory links) would be consistent with, and explained by, numerous types of cognitive models of memory, including simple associative networks as well as connectionist models, for example. A consideration of the types of cognitive architecture that might be able to explain and reproduce goal phenomena is beyond the scope of this chapter and we consider it to be one of the next challenges that social-cognitive psychologists will face in the near future, just as has been the case with research on attitudes and stereotypes (e.g., Bassili & Brown, 2005; Mitchell, Nosek, & Banaji, 2003; Smith, 1996).

The Content of Goals

Beyond these structural characteristics, what type of knowledge is reflected by goal memories? The answer to this question directly builds on our definition of goals as representations of desired endpoints that direct behavior, evaluation, and emotions. Below we consider in more detail what this view implies about the nature of goal memories.

Ends and Means

First and foremost, goals contain information about end states. End states are the reference points toward which behavior is directed. One notable feature of end states is that they can vary in their abstractness (Hommel, Muesseler, Aschersleben, & Prinz, 2001; Jeannerod, 1997; Kornblum, Hasbroucq, & Osman, 1990; Kruglanski et al., 2002; Miller, Galanter, & Pribram, 1960; Powers, 1973). For example, a goal may involve an end state that entails something tangible and perceptual in the world (e.g., having a cup of coffee) or one that is relatively more abstract and conceptual in nature (e.g., achievement).

Goals entail more than just end states, however. They also include the variety of behaviors, plans, and objects that enable one to reach that end state. For instance, the goal of getting a cup of coffee might entail temporally ordered, procedural information about first grinding coffee beans and then putting them into a filter in a coffee machine (see Norman, 1981), and the goal of achievement might include behaviors such as studying at the library and paying attention in class (Aarts & Dijksterhuis, 2000; Bandura, 1997; Bargh & Gollwitzer, 1994; Carver & Scheier, 1998; Custers & Aarts, 2005; Emmons, 1992; Schank & Abelson, 1977; Shah & Kruglanski, 2002, 2003; Vallacher & Wegner, 1985; Vallerand & Ratelle, 2002; Wilensky, 1983). The behaviors and objects associated with an end state can also vary in abstractness. For instance, the end state of achievement might include the specific behavior of neatly writing lecture notes as well as the relatively more general behavior of being punctual.

When considering ends versus means, it quickly becomes apparent that almost any end state can be understood as a means for a higher-order end state. For instance, the means of studying in order to attain academic success could itself constitute an end state with its own associated means (e.g., take notes and go to library). In such a hierarchical organization, the terms "end state" and "means" are clearly meaningful only in relation to one another. Despite the relative nature of the terms, they are nevertheless useful in that they identify the point toward which a person is striving, and the specific ways in which that person might succeed. In this way, the "end state" organizes one's behavior, whereas the variety of means can be somewhat interchangeable or substitutable, and an inability to utilize one means does not necessarily imply that the end state is abandoned (e.g., Kruglanski et al., 2002; Tesser, Martin, & Cornell, 1996).

Evaluative Information

We assume that a goal consists of an overall end state and the behaviors, objects, and plans needed for attaining it. But is that all a goal is? Just because someone possesses knowledge about how to put a tree house together, for instance, does not mean that that knowledge constitutes a goal. This leads to a second important aspect of the content of goal constructs—the end state (and its associated means) has to be *desirable* (Carver & Scheier, 1981; Custers & Aarts, 2005; Kruglanski et al., 2002; Peak,

1955; Pervin, 1989; Shah et al., 2002; Young, 1961). By definition, a goal that is desirable must be associated in some way with positive affect. We argue that, in line with the long-standing notion that people are motivated to approach pleasure and avoid pain (Arnold, 1960; Bogardus, 1931; Corwin, 1921; Doob, 1947; Frijda, 1986; Lang, 1984; Lazarus, 1991; Lewin, 1935; Mowrer, 1960; Osgood, 1953; Thurstone, 1931; Young, 1959), the positivity in a goal representation is what imbues the construct with its motivational force. In other words, the primary reason that goals influence and guide behavior is because the positivity associated with them is inherently motivating (see research on expectancy-value models, Atkinson, 1974; Tolman, 1932).

Although we define goals as desirable end states, and therefore assume that they must include positive evaluative information in their representation, it is not yet clear exactly how goals become positive. For instance, a goal might become positive and desirable in a conscious and intentional manner, such as when a person sees a friend playing a complicated, fun game and wants to learn it in order to join in. Or, a goal can become desirable in a more implicit, nonconscious fashion, such as through repeated pairings (i.e., conditioning) of a given activity and consequent reward experiences. Recent research has provided support for the latter claim. Custers and Aarts (2005) first implicitly conditioned a goal (e.g., playing a puzzle) with positive evaluations by creating a computer task in which they paired aspects of a task (e.g., the words "puzzle" and "number") with positive words (e.g., "happy"). They found that participants who had received positive (vs. neutral) conditioning of the puzzle words subsequently showed greater motivation to begin the puzzle task.

What Distinguishes a Goal Construct from Other Social Psychological Constructs?

We have noted so far that a goal construct varies in accessibility, consists of many interconnected memories, and operates according to classic knowledge activation principles. These memories refer to ends and means and also contain positive information. But, given these characteristics, how is a goal construct distinct from other types of knowledge structures?

Goals have been distinguished from other hypothetical constructs primarily by the nature of their effects on behavior (Aarts, Gollwitzer, & Hassin, 2004; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Troetschel, 2001; Fitzsimons & Bargh, 2003; Kawada, Oettingen, Gollwitzer, & Bargh, 2004; Shah & Kruglanski, 2003; see review by Förster and Liberman, Chapter 9, this volume). In particular, the strength, or activation, of a goal only dissipates when the goal has been reached, whereas the activation of semantic constructs dissipates at a constant rate from the moment of activation (Atkinson & Birch, 1970; Gollwitzer & Moskowitz, 1996; Lewin, 1936; McClelland, Atkinson, Clark, & Lowell, 1953). Specifically, whereas Lewin (1936) suggested that a goal will stay active until the discrepancy between the actual and desired state is reduced, others have argued that the goal

strength will actually increase over time until it is met (Atkinson & Birch, 1970; McClelland et al., 1953), or when the pursuit becomes too difficult to sustain (Brehm & Self, 1989; Wright, 1996). This suggests, for example, that when the *goal* of achievement has been activated, the person will increase his or her efforts for a while until the goal has been met (or until he or she encounters an insurmountable obstacle). In contrast, when mere *semantic knowledge* about achievement has been activated, that activation should rapidly decay over time such that the person may quickly show less evidence of that activated knowledge in perception or judgment (see Bargh et al., 2001).

It follows that a cue for a goal (e.g., the word "achievement") does not always influence behavior in a goal-related fashion; rather, its influence depends on other variables such as the nature of the task and whether the goal is applicable to it. In addition, whereas all goals include semantic knowledge, not all semantic constructs are goals (i.e., have motivational force) or positivity associated with them. As we consider how a goal might become activated and then operate, we review the ways in which researchers have distinguished between goals versus other types of constructs.

ON THE ACTIVATION OF A GOAL

What determines whether a given goal is activated and then guides behavior? The main theme of classical goal research has been that goals are enacted when people deliberately and purposively decide to adopt them (Bandura, 1986; Carver & Scheier, 1998; Deci & Ryan, 1985; Gollwitzer, 1990; Locke & Latham, 1990; see also Mischel et al., 1996, for a review). This would suggest that a goal becomes activated via conscious, intentional thought. For instance, a person might consciously consider whether to intentionally pursue the goal of being funny while at a dinner party.

However, research over the last decade on how goals become activated suggests a different perspective. Many of the insights in this work follow from the definition of goals, and the assumptions regarding their structure in memory in particular. We noted earlier that goals consist of interconnected memories that become activated (i.e., more accessible) according to knowledge activation principles. This means that the perception of any stimulus that is strongly associated with the goal should be sufficient for the goal to become activated (Bargh, 1990; Bargh & Barndollar, 1996; Bargh et al., 2001; Gollwitzer, 1999; Jacoby & Kelley, 1987; Kruglanski, 1996; Shah & Kruglanski, 2003; see also McClelland, Koestner, & Weinberger, 1989). Importantly, the perception of a stimulus does not have to be conscious (e.g., Greenwald, 1992; Greenwald, Draine, & Abrams, 1996). And, even if people's perception of a stimulus *is* conscious, they may not be aware that it has activated a whole array of associated memories, including goal constructs (see Ferguson & Bargh, 2004a).

In general, by considering goals as constructs in memory, recent goal research acknowledges the possibility of

nonconscious goal activation. We review below the kinds of stimuli that are capable of triggering goal activation. This range of stimuli must, by necessity, be associated with that goal. In this way, not only does our review address the ways in which goals can become activated, it also further reveals the kinds of stimuli that are part of the goal construct.

It is also important to note that although we concentrate in this section on the ways in which goals are activated, the findings also necessarily speak to the operation of a goal. That is, we infer the activation of a goal from how the goal influences behavior, judgment, attitudes, and emotions. Although goal activation and goal operation are often empirically difficult to disentangle, we assume that goal activation precedes goal operation. Therefore, we emphasize in the next section the minimal requirements for a goal to be activated, and we then turn our attention to the types and kinds of downstream consequences of activation in subsequent sections.

Priming by End States and Means

In one of the first tests of how a goal can become activated and influential without the person's awareness or intention, Chartrand and Bargh (1996) subtly primed participants with either a person impression or memory goal. They administered to participants a scrambled sentence task in which participants had to create grammatically correct four-word sentences out of groups of five scrambled words (Srull & Wyer, 1979). Some sentences included words related to forming an impression (e.g., judge, impression, and personality), while others contained words related to memorization (e.g., remember, recall, and retain). Participants were then asked to read through a set of behaviors about a fictional target and were given a surprise recall test afterwards. The results showed that those who had merely read a few words related to forming an impression in fact processed and integrated the behavioral information about the target in a way similar to when someone is intentionally trying to form an impression. That is, they formed more clusters of the behaviors around personality traits and were also more likely to show deeper processing of those behaviors that were inconsistent with the overall personality theme (e.g., see Hamilton & Sherman, 1996; Stangor & MacMillan, 1992). This was one of the first demonstrations of how information-processing goals can become nonconsciously activated and influential.

But, what is the behavioral evidence that a goal is nonconsciously activated? Bargh and colleagues (2001) tested for goal activation by first asking participants to complete a word-search puzzle. Whereas for some participants some of the words were related to achievement (e.g., strive, achieve, and master), for others none of the words were related to this goal. After this subtle exposure to the notion of achievement, participants were asked to complete a series of other word-search puzzles. Those who were exposed to achievement words found significantly more words than those in the control condition. These findings demonstrate that by simply reading words related to a given end state, a person is likely to

perform goal-congruent actions unknowingly and unintentionally.

How might a nonconsciously activated goal compare with one that is consciously activated? To examine this question, participants in another study (Bargh et al., 2001) were explicitly told to cooperate, were subtly primed with cooperation words via a scrambled sentence task, or were not primed in any way. Each participant then played a resource management game with another participant in which they had to fish from a lake while ensuring that the lake did not become depleted. The results showed that those in the consciously activated goal condition, as well as those in the nonconsciously activated goal condition, showed more cooperation than those in the control condition. It appears that a nonconscious prime can have an effect similar to a conscious prime on goal-congruent behavior.

Another predominant issue concerns the evidence for the activation of a goal versus some other construct. That is, in these tests of nonconscious goal activation, how do we know that a goal was activated, versus perhaps simply semantic concepts related to the goal? For example, was participants' achievement behavior due to the influence of the *goal* of achievement or simply the *semantic concept* of achievement? Perhaps the priming task simply increased the concept of achievement, and then participants interpreted the situation as achievement related and acted accordingly. Recall that whereas the activation of semantic concepts decreases over time, the activation of goals increases over time until the goals are attained. Accordingly, Bargh and colleagues (2001) nonconsciously primed participants with achievement and then asked them to complete either a semantic task of evaluating an ambiguously achieving target (Higgins, 1996) or a goal task of solving a set of word-search puzzles. Participants also completed the measure either immediately after the priming or after a 5-minute delay. In the immediate condition, those in the priming condition who did the goal task performed better than those in the control condition, and those in the priming condition who completed the semantic judgment task rated the target as more achieving than those in the control condition. The critical question concerned the effects for those in the delay condition. If nothing but the semantic concept of achievement was activated, the effects for both the judgment task and the goal task should have decayed. However, if the goal of achievement was actually activated (in addition to semantic knowledge), the effect on the goal-relevant task should have increased over time. The pattern of results confirmed this, suggesting that the goal of achievement was indeed activated.

More recent research suggests that in addition to end states, goals can also be nonconsciously activated by relevant means and strategies. Shah and Kruglanski (2003) showed that people who were subliminally primed with a recently learned behavioral strategy showed evidence of pursuing the goal related to the strategy. In one study, before completing an anagram task, participants learned a strategy for solving anagrams. Those participants who were subliminally primed with the name of that strategy ("first–last," which refers to determining initially

whether the first and last letters of the letter string anchor any known words) showed a greater accessibility of words related to anagrams and also exhibited more persistence and better performance. These findings suggest that the perception of (even recently learned) means can activate the goal associated with that means.

Whereas the work described earlier showed that a goal can be nonconsciously activated by semantic cues (i.e., words) closely related to end state or means, what other ways might goals become triggered by the environment? We suggested that the perception of any stimulus that is associated with the goal should be sufficient for the goal to become activated. Because people live in a social environment, a large proportion of these stimuli are social stimuli. Indeed, a bevy of studies has now uncovered some of the main categories of social stimuli that lead to goal activation.

Priming by Relationship Partners

Goals can include the representation of individuals (e.g., a parent and a teacher) who expect the person to pursue the goal as well as the representation of individuals who pursue that goal themselves. For instance, a person's goal of making money might include representations of that person's father, who expects that person to make money, as well as representations of a best friend who is obsessed with making money. If goals include representations of others, the perception of a relationship partner can automatically activate those goals associated with that partner.

As a demonstration of this principle, Shah (2003) has shown that being subtly reminded of a significant other can activate the significant other's expectations, which can then influence the person's own expectations and performance. Shah demonstrated that participants who were subliminally primed with the name of a significant other who had high expectations for the person (e.g., a father) on an anagram task actually persisted longer and performed better than those not primed. In a similar line of research, Fitzsimons and Bargh (2003) claimed that people normatively have achievement goals for impressing their mothers. They accordingly found that those who were reminded of their mothers in a subtle way achieved more on a word-search puzzle than those not reminded.

Relationship partners can further activate the emotional experience that is included in the goal representation. For example, Higgins and colleagues have shown that people can adopt a style that emphasizes nurturance needs (a promotion focus) or one that emphasizes security needs (a prevention focus; see Higgins, 1997). Based on this theory, Shah (2003) showed that a significant other's regulatory focus can also influence one's own reactions to the task according to regulatory focus. For example, those whose fathers *hoped* that they would do well on academic tasks (a promotion goal), and who were primed with words related to father, experienced cheerfulness when given positive feedback on an anagram task and dejection when given negative feedback on the task, in line with the ways in which regulatory focus influences

emotion-specific reactions. Those whose fathers *expected* them to do well on academic tasks (a prevention goal), and who were primed with father-related words, experienced relaxation when given positive feedback on an anagram task and agitation when given negative feedback, again in line with research on how regulatory focus influences emotions (e.g., Higgins, Shah, & Friedman, 1997).

Priming by Group Members

In addition to relationship partners activating goals, the perception of (unfamiliar) group members can also activate the goals that the perceiver tends to pursue when in the presence of those group members (Cesario, Plaks, & Higgins, 2006). When one encounters another person, an automatic preparation to interact with that person, either in an approach or an avoidance manner, is activated, depending on that person's implicit attitudes toward that group. The result is that the perception of a group member activates one's goals toward that group (in addition to stereotypes) and these goals influence behavior. As a demonstration of this principle, Cesario and colleagues (2006) primed participants with gay or straight men and then introduced a mild provocation when the computer failed and participants' data were supposedly lost (a paradigm developed by Bargh, Chaiken, Raymond, & Hymes, 1996). The degree to which participants then interacted with the experimenter in a hostile manner constituted the main dependent measure. If the contents of the gay stereotype are most influential, then those who were primed with gay should behave in a more passive manner after the provocation (given that gay men are stereotyped as passive; e.g., Herek, 2000, 2002) compared with those not primed. However, if one's goal to interact with the group member is activated and assuming that most people have negative implicit attitudes toward gay men, those primed with gay men should be more hostile toward the experimenter than those not primed. The results favored the latter hypothesis—priming gay men activated the goal to act with hostility.

Priming by a Stranger's Goal Pursuit

In addition to relationship partners and group members, the perception of another person engaging with goal-related actions might be sufficient to trigger the goal related to these actions, even if the actor is unfamiliar. This is because people infer other people's goals from their actions, and these inferred goals have implications for one's own behavior (Aarts et al., 2004; Aarts, Hassin, & Ferguson, 2005). As a demonstration, Aarts and colleagues (2004) gave participants a vignette about a target person's behavior (which implied a goal), and then participants were placed in a setting where they could behave in line with that goal or not. For instance, in one study, male participants either read about a target person who was trying to pick up women in a bar (implying the goal of seeking casual sex) or read a control vignette that did not imply the goal. Participants were then asked to provide feedback on one of the experimental tasks to the experimenter, who was described to half of the partici-

pants as female and to the other half of participants as male. Because men who are sexually interested in women tend to show more helping behavior toward them (e.g., Baumeister & Tice, 2001; Buss, 1988), those who had read the vignette implying the goal of casual sex gave more feedback (i.e., showed more helping behavior) toward the female experimenter but not the male. These findings show that merely observing someone else's behavior can activate the goal associated with the behavior.

Notably, these "goal contagion" effects reflected the influence of a goal rather than the influence of simple behavior priming (e.g., Bargh et al., 1996; Dijksterhuis & van Knippenberg, 1998). Namely, because the dependent measure (giving feedback to a female experimenter) was sufficiently semantically distinct from the primed behavior (picking up women in a bar), the effect was probably due to an overarching goal that contained both behaviors as means.

Summary

The research we have described in this section shows how a goal can become activated (and influential) on the basis of the mere (conscious or nonconscious) perception of a goal-related stimulus. There is precedent for this notion in classic goal research, which assumes that the degree to which a person is consciously thinking about a goal determines the likelihood that the person will pursue it (e.g., Bandura, 1986; Deci & Ryan, 1985; Gollwitzer, 1990; James, 1890; Lewin, 1935; Locke & Latham, 1990; Mischel et al., 1996). The present analysis expands on classical research, by showing that if accessibility is in fact the underlying mechanism, goals should be able to be activated by even the nonconscious perception of goal-related stimuli.

Importantly, the claim that a goal's influence will depend on its accessibility in memory does not imply that people will behave in line with whatever memories have recently been activated. Once a goal is activated, its effect on behavior still conforms to the principle of applicability (Higgins, 1996). Increased accessibility of a construct via priming simply means that it will be more likely to be applied to a stimulus that is *relevant to that construct*. In the research reviewed earlier, participants were primed with cues for a certain goal and were then placed in a situation that "afforded" the relevant goal pursuit to some degree. The degree to which a particular task is goal related determines the extent to which an accessible goal guides behavior.

In our discussion of goal activation, we inferred activation based on the downstream behavioral effects of goals (e.g., puzzle performance, helping behavior, and hostile behavior). In this way, this research joins a litany of other classic findings showing how goals influence behavior. However, in thinking about the downstream consequences of goals on behavior, we now move away from merely documenting overt, behavioral effects to identifying more subtle effects that perhaps might mediate between a goal and overt behavior. Specifically, we are interested in examining the ways in which an activated

goal, and in particular a nonconsciously activated goal, influences knowledge accessibility, evaluations, and emotions, and we review these influences in the next section.

ON THE OPERATION OF A GOAL

In this section we identify the characteristics of goal pursuit, including goal-relevant knowledge accessibility, goal-relevant evaluations, goal-relevant moods and emotions, and, of course, goal-relevant choices and behaviors. Just as we did in the section on the activation of a goal, we develop the current section on the operation of a goal based on the definition of the goal construct that we outlined in the beginning of the chapter. In particular, throughout the following section we note how some of the characteristics of goal operation derive directly from our assumptions about the content and structure of the goal concept. For instance, because goals contain information on evaluations and behaviors, the operation of goals can be characterized by changes in the evaluation of goal-related stimuli and the enactment of goal-congruent behaviors.

Goal-Relevant Knowledge Accessibility

We proposed earlier that increased accessibility of goal-related knowledge is what it means for a goal to be activated. In addition, the accessibility of goal-related knowledge can also be understood as a *consequence* of goals. This suggests that goal-relevant knowledge should be more accessible during the pursuit of that goal, compared with when the pursuit is over or has not been initiated. For example, the activation of the hunger goal should increase the accessibility of knowledge that is related to that goal, such as restaurants. In this way, the increased accessibility of restaurants simultaneously represents what it means for a hunger goal to be activated and one type of downstream consequences of goal activation.

There is a long history of the theoretical notion that the (conscious) activation of a goal influences the types of knowledge that become accessible (Ach, 1935; Bargh, 1997; Bruner, 1957; Gollwitzer, 1996; Jones & Thibaut, 1958; Klinger, 1996; Kruglanski, 1996; Kuhl, 1987; McClelland & Atkinson, 1948). Some of the precedent for this started with the New Look research movement. In contrast with the classic view of perception in the first half of the 20th century that perception was entirely driven by the stimulus (Stevens, 1951), New Look research showed that people's perceptions are influenced by the value of the stimulus being perceived (Bruner, 1957; Bruner & Postman, 1948; Jones & Thibaut, 1958; McClelland & Atkinson, 1948; for a review, see Greenwald, 1992). For example, in the classic experiment by Bruner and Postman (1948), poor children overestimated the size of coins to a greater degree than rich children, for whom the money was presumably less intensely desired. In a review of the New Look research, Bruner (1957) argued that what people want, need, and desire can influence the accessibility of knowledge, and thus how they see the world around them. Each nonconscious

act of perception is an act of categorization, with multiple categories being available for a given stimulus. People's needs and motives can influence the accessibility of those categories and thus make them "perceptually ready" to categorize, or perceive, stimuli in certain ways. For instance, when people are looking at an ambiguous object in the distance that looks like a storefront but could be a restaurant façade, they should be more likely to "see" a restaurant when they are hungry than when they are not (see Bruner, 1957; see also Glenberg, 1997).

Recent evidence provides more methodologically rigorous support for the theoretical claim of the New Look that an active goal increases the accessibility of related knowledge (Aarts et al., 2001; Balcetis & Dunning, 2006; Förster, Liberman, & Higgins, 2005; Moskowitz, 2002). For instance, Moskowitz (2002) tested whether knowledge that is related to an active goal automatically captures attention. Based on self-completion theory, people who receive negative feedback about an important self-relevant domain should be especially motivated to reestablish competence in that domain. Accordingly, Moskowitz (2002) reasoned that athletes who think about one of their recent athletic failures (e.g., missing a crucial foul shot) should be highly motivated to reclaim or prove their competence as athletes, and if so, those who have recently thought about failure should demonstrate the strongest accessibility of knowledge related to their goal of athleticism. Participants thought about either a recent failure or success in athletics or nothing at all, and then they completed a computer task in which there were distractors either related or unrelated to athleticism (e.g., athletic, fast, and agile). Those participants who had been thinking about failure, and thus who presumably had particularly accessible goal-related knowledge, responded more slowly to the focal task when the distractors were athletic-related versus unrelated. Apparently, when a goal is activated, stimuli related to the fulfillment of that goal become highly accessible and automatically attract attention.

But, part of our argument (also consistent with the New Look research) is that the accessibility of goal knowledge should influence the stimuli in the environment to which people pay attention. Does this happen? Aarts, Dijksterhuis, and De Vries (2001) manipulated participants' thirst by asking some of them to consume salty snacks. Participants then completed a lexical decision task in which some of the words were beverages or items used to drink beverages (e.g., juice, soda, and bottle). The results showed that those who had been manipulated to be thirsty showed significantly greater accessibility of drinking-related words, compared with control words, and compared with nonthirsty participants. Aarts and colleagues then showed in a second study that thirsty participants were more likely than nonthirsty participants to recall drinking-related objects. These studies demonstrate that the goal of quenching thirst can render accessible knowledge concerning stimuli, actions, and concepts related to sating that goal, just as Bruner (1957) and others argued, and, importantly, that greater accessibility then determines the objects to which people attend in their environment.

Although an active goal increases the accessibility of knowledge related to that goal, which then influences the stimuli that are noticed, does it influence what people actually see in the world, as New Look researchers claimed? Recent research by Balcetis and Dunning (2006) has provided support for this notion. In one study, participants were told that they were going to be randomly assigned by the computer to one of two conditions. In one (desirable) condition, they would be asked to taste a glass of fresh orange juice, and in the other (undesirable) condition, they would have to sip an unappealing, green vegetable drink. They were told that the computer would randomly present either a number or letter to them, and that either a number or letter (depending on counterbalancing) would mean that they were assigned to the OJ condition. The computer then flashed the well-known, ambiguous "B/13" figure, and then there was a message indicating computer failure. The experimenter, who had not seen what was flashed, asked the participant what he or she saw on the screen. Whereas those for whom the *number* meant the desirable condition were more likely to see the 13, those for whom the *letter* meant the desirable condition were more likely to see the B. A series of additional experiments demonstrated (using a variety of implicit measures) that the effect was not due to response bias but, rather, reflected what participants actually perceived. On the basis of this work, we conclude that what someone wants does influence how they disambiguate stimuli in the world; critically, this seems to happen because what someone wants influences the types of knowledge that are accessible in memory, which then serve to capture any ambiguous stimuli relevant to that knowledge (see Bruner, 1957; Higgins, 1996).

Interestingly, goal pursuit is not simply characterized by accessible knowledge during the pursuit; the completion of a pursuit leads to the inhibition of related knowledge. Recently, Förster and colleagues (2005; see also Liberman, Förster, & Higgins, in press) have demonstrated this point. They asked participants to search for a picture of a pair of glasses on a computer screen and found that during the search, but before participants found the target, the accessibility of words related to glasses was greater compared with the accessibility for those who were not searching for the target. This is in line with the findings we just described. However, once participants found the target, the accessibility declined below the level for control participants. This work is consistent with work in cognitive science showing that knowledge related to fulfilled intentions becomes inhibited (Goschke & Kuhl, 1993; Liberman & Förster, 2000; Marsh, Hicks, & Bryan, 1999).

Goal-Relevant Evaluations

We argued in the previous section that the accessibility of goal-related knowledge can be understood as evidence of goal activation as well as a consequence of goal activation. In a similar way, the effects of goals on *evaluations* of stimuli in the environment can be conceptualized both as evidence that those stimuli are relevant to an active goal and as effects of that active goal. Indeed, we argue in this framework that the evaluations that follow from goal pursuit reveal the nature of the associations in memory between the goal construct, means and objects, and evaluative information. We therefore suggest that the "effects of a goal" on evaluation and emotion also speak to the content of the respective goal construct.

How then does active goal pursuit influence the way in which people evaluate stimuli related to that goal? In one way, the answer to this question is obvious and straightforward, and seems self-evidently true. When people are actively pursuing a goal, by definition they want (desire) those things that can help them achieve the goal, and similarly should not want those things that prevent them from reaching the goal. For example, being thirsty makes water more desirable and positive because it can alleviate one's thirst, and salty things more undesirable because they can exacerbate one's thirst (see also Loewenstein, 1996). Thus one consequence of goal operation is more positive evaluations of those stimuli that can facilitate the goal, and perhaps more negative evaluations of those stimuli that can thwart the goal (Brendl & Higgins, 1996; Cabanac, 1971; James, 1890; Lazarus, 1991; Lewin, 1926, 1935; Markman & Brendl, 2000; Rosenberg, 1956; Shah & Higgins, 2001).

In what follows, we explore how goals influence evaluations but focus in particular on studies that used implicit rather than explicit measures of evaluation. There are two reasons for this focus. First, implicit measures capture changes in evaluations that are not contaminated by people's response biases, self-presentation pressures, or demand effects. In this way, any changes in implicit evaluation as a function of goal pursuit can be regarded as spontaneous and likely to occur in "real-world," nonlaboratory settings. Second, research has shown that explicit and implicit evaluations are not identical; not only might they rely on different memories and underlying processes (e.g., Gawronski & Strack, 2004; Hofman, Gawronski, Gshwendner, Le, & Schmidt, 2005), they also seem to guide different types of behaviors (e.g., Asendorpf, Banse, & Mücke, 2002; Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Egloff & Schmukle, 2002; Fazio, 1990; Wilson, Lindsey, & Schooler, 2000). Whereas explicit evaluations seem to guide behaviors of which the person is aware, and that are easy to guide and monitor, implicit evaluations seem to direct behaviors that are less intentional and relatively more difficult to control and monitor. Given that implicit evaluations influence people's subtle and unintentional behaviors, any effect of goals on implicit evaluations would explain and demonstrate one way in which goals can guide people's behavior in a subtle and nonconscious manner.

Evaluations of Stimuli Consistent with the Goal

Stimuli are evaluated implicitly in line with one's active goals (Ferguson & Bargh, 2004b; Moors & De Houwer, 2001; Moors, De Houwer, & Eelen, 2004; Sherman, Rose, Koch, Presson, & Chassin, 2003). In support of this proposition, Sherman and colleagues (2003) found, for example, that chronic cigarette smokers automatically

evaluate cigarette paraphernalia more positively when they are in need of a fix, versus when they just recently satisfied the urge. In one study, heavy smokers who had been instructed to refrain from smoking automatically evaluated smoking-related stimuli in a more positive fashion than those heavy smokers who had just recently smoked. This suggests that when a goal is activated, those stimuli that can help the person to reach the goal are automatically evaluated as positive.

But how long does this implicit positivity last? Ferguson and Bargh (2004b) showed that stimuli that are relevant to a currently active, but not recently completed, goal are implicitly evaluated as more positive than control stimuli. This suggests that the effect of a goal on implicit evaluations lasts only as long as the goal is active. In one study, participants who were still involved in a competitive word game automatically evaluated game-related words (e.g., win and achieve) as more positive than those who had never played the game, as well as those who had played the game but were already finished. This demonstrates that the automatic evaluation of stimuli is contingent upon what the perceiver is currently doing at the moment, rather than what the perceiver has just done. In another demonstration of goal-based evaluation, Ferguson and Bargh asked thirsty participants to either drink multiple beverages, thereby sating their thirst, or sample salty, dry pretzels, thereby exacerbating their thirst. The participants then automatically evaluated a series of words that varied in their relevance to thirst. The results showed that those who were still thirsty automatically evaluated words that were strongly related to the thirst goal (e.g., water and juice), but not unrelated to the thirst goal (e.g., chair), as more positive than those who had just sated their thirst.

In general, then, there is some evidence for our claim that objects and means related to a goal become more implicitly positive when that goal is active compared with when it is not. However, what about the end state itself? When someone is pursing an achievement goal, for instance, are words such as *success* and *achievement* evaluated in a more positive manner? We claim that people who are actively pursuing a goal automatically evaluate relevant end states as more positive compared with when the goal is not being pursued. In a study that tested for this possibility, Ferguson and Bargh (2004b) assumed that participants who were asked to think about recent failure in an important, relevant domain would be the most motivated to pursue that end state (reestablish their competence in the domain) compared with those who thought about success in the domain, or who thought about an unrelated topic (see research on self-completion theory; e.g., Wicklund & Gollwitzer, 1982). Participants who were athletes were thus asked to think about recent failure or success in athletics, or an unrelated topic. Their automatic evaluations of words related to the goal of improving their athleticism (e.g., athletic and agile) were then measured. As predicted, those participants who had thought about a recent failure in athletics generated the most positive automatic evaluations of the end states (and not other types of words) compared to those who had thought about success or an unrelated topic. Consistent with previous research on self-

completion theory, this effect emerged most strongly for those for whom the athletic domain was the most important—varsity athletes. The activation of a goal thus renders as positive those end states that are directly related to the goal.

Even though the evaluation of stimuli seems to depend on whether those stimuli are related in some way to people's current goals, this does not mean that stimuli that are unrelated to a current, primary goal will have no valence. People's average evaluations of stimuli should indicate the average relevance of those stimuli for the person's goals. Obviously, those stimuli that are consistently useful for a person's important goals might be evaluated as positive most of the time, whereas those that are only occasionally useful might be less consistently positive. If so, it should be the case that people's implicit evaluations of stimuli in default (non-goal-related) settings should predict the likely influence of that goal in a goal-relevant setting. Ferguson (in press) tested this by measuring participants' chronic, implicit evaluations of end states in one setting, and then testing whether those evaluations predicted participants' goal pursuit in another setting. In one study, participants' implicit evaluation of the goal to be thin was measured. A week later, participants were asked to report how much over the previous week they had avoided eating tempting foods, as well as how often they planned to do so in the upcoming week. Participants' implicit evaluations measured a week earlier significantly predicted their goal-relevant behavior, and even did so significantly above and beyond their explicit evaluation of the goal. Such findings suggest that people's chronic goals influence their evaluation of stimuli related to the respective end states.

Evaluation of Stimuli Inconsistent with the Goal

The activation of a goal representation might also lead to more negative evaluations of stimuli that undermine that goal (e.g., Ferguson, 2006; Fishbach, Zhang, & Trope, 2006). For example, participants who were consciously or nonconsciously primed with a goal construct (e.g., academic pursuits) implicitly generated negative evaluations of words that were related to another low priority goal (e.g., social life) that might undermine the primed goal (Ferguson, 2006). But importantly, whereas an active high-priority goal undermines the positive value of stimuli related to a competing low-priority goal (as in the previous case), an active low-priority goal may actually increase the positive value of stimuli related to a competing higher-priority goal, because of the motivational priorities of the person pursuing these goals. For example, reminding participants of their social goals led to a more positive evaluation of academic pursuits among students who strived toward academic excellence and considered it more important than social activities (Fishbach, Zhang, & Trope, 2006; Trope & Fishbach, 2000). In the section "On the Interaction among Goals," we discuss these patterns of influence between conflicting goals in more details.

Are there any variables that might determine when negative evaluation of goal-undermining stimuli is most

likely to occur? One possibility is that the extent to which it occurs depends on whether the person can effectively self-regulate in the focal goal domain. The findings from Ferguson (2006) and Fishbach, Zhang, and Trope (2006) together suggest that negative goal-related evaluations emerge most strongly for those who are skilled in the focal domain. For example, when participants were nonconsciously primed with academic concerns (e.g., grades), they automatically evaluated social temptations as more negative—especially so if they had relatively high grade point average (GPA) scores. This suggests that the degree to which goals might shift automatic evaluations of pertinent stimuli in some cases depends on the person's skill level and experience in the relevant goal domain.

We further argue that the activation of a goal can have repercussions for the evaluation of stimuli that are irrelevant to the goal. Recent work by Brendl and colleagues (Brendl, Markman, & Messner, 2003; Markman & Brendl, 2000) has suggested that such "devaluation effects" occur when the activation of a given goal (e.g., hunger) renders as negative those objects (e.g., movie tickets) that might draw resources away from the focal goal. From this perspective, even though movie tickets do not directly undermine the goal of getting food, they indirectly do so by drawing limited resources away from the focal pursuit (see also Shah et al., 2002). To test this idea, they asked smokers who had or had not recently smoked to purchase raffle tickets for a prize of either cash or cigarettes. A devaluation effect occurred such that deprived smokers bought fewer tickets for the cash prize than those smokers who were not deprived. In this way, the active goal to smoke led to a lower evaluation of cash. We conclude that the activation of a goal may make stimuli that are not directly relevant to the overall goal less positive.

Goal-Relevant Moods and Emotions

Beyond evaluations of specific stimuli, how might the operation of a goal influence one's affective state more generally? There are at least two ways to approach this question. It is possible to consider the ways in which goal pursuit might influence people's moods and emotion both *during* the pursuit as well as *after* the pursuit has been completed. We first consider the former, and then move to the latter.

Considering our earlier argument that during goal pursuit the related end state and associated means should be evaluated as more explicitly and implicitly positive, it seems possible that the positivity associated with a specific stimulus (e.g., a means) might extend to a more general affective state, such as a mood or emotion. This possibility was supported in research by Fishbach, Shah, and Kruglanski (2004). These researchers documented a transfer of emotions from goal to related means in proportion to the degree of association between the means and expected goal attainment. In particular, while pursuing a given means, people experience some of the emotions that characterized goal attainment. For example, in one of their studies, participants self-generated a goal

(e.g., making friends), and one versus two activities that serve this goal attainment (e.g., joining a fraternity and being helpful to people). Listing a second activity was expected to dilute the association between the goal and the first activity, thereby decreasing the magnitude of the emotional transfer. Accordingly, participants perceived the first activity listed as more enjoyable when it was the only activity listed compared with it being the first of two activities listed. In another study, it was shown that the quality of feelings (promotion- or prevention-type affect) experienced toward social figures who also serve the attainment of means (e.g., a hair designer and a tax consultant) varied as function of the type of goals they were helpful in mediating.

People also experience general affective states during goal pursuit as a result of feedback processes, a possibility posed by cybernetic models of behavioral control. For example, Carver and Scheier (1990, 1998) have argued that people monitor the discrepancy between the desired end state and their current status, and that their mood can be an important part of the feedback for such monitoring. Specifically, when people are progressing faster than they expected, a positive mood will be generated. A negative mood, on the other hand, should result when one's progress is slower than expected. Theoretically, this means that as long as mood is associated with goal performance, a negative mood should prompt people to increase their efforts and pursuit, while a positive mood should signal that people should relax their efforts given that they are moving more quickly than they planned (see Carver, 2003).

What about moods and emotions that emerge *after* the termination of a goal pursuit? In one way, an answer to this question is straightforward. Psychologists have long recognized that there are general affective consequences for attaining desirable things and failing to do so. Those who attain things that they view as desirable feel good; indeed, things are desirable precisely because they promise to deliver pleasure or an escape from pain. And, by extension, those who fail to reach something desirable will undoubtedly feel bad. Although people may not be able to accurately calibrate the actual extent to which they will feel good or bad once they reach or fail to reach a goal (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000), it is well established that such affective experience transpires in this way (e.g., Bandura, 1989, 1991; Carver & Scheier, 1990, 1999; Clore, 1994; Frijda, 1996; Higgins, 1999).

Moreover, the nature of a given goal pursuit influences moods and emotions (e.g., Higgins et al., 1997). Specifically, different goals will lead to different emotional responses to completing the pursuit. A focus on reducing the discrepancy between one's actual and ideal self (a promotion focus) leads to feelings of cheerfulness in the case of success and dejection in the case of failure. In contrast, a focus on reducing the discrepancy between one's actual and "ought" self (a prevention focus) leads to feelings of calmness in the case of success and anxiety in the case of failure.

We therefore suggest that goal pursuit can influence more generalized affective states in addition to evalua-

tions of specific stimuli. Furthermore, the termination of a goal pursuit induces certain affective states. One important question, however, is whether these "effects on affect" can also be considered part of the goal construct; that is, whether they should be considered both part of what it means for the goal construct to be activated in memory as well as the consequences of goal operation. We argued earlier that implicit effects on knowledge activation and evaluations can reveal the content of the goal construct, and we extend this logic to more generalized affective states. Goal constructs include the positive emotions that characterize goal attainment as well as the negative emotions that characterize goal failure. These emotions may be associated with the end state as well as with the related means of attainment and be part of the goal structure (Fishbach et al, 2004; Higgins, 1997). In addition, emotions are downstream consequences of goal activation and goal pursuit, as we reviewed here.

Goal-Relevant Behavior

Goals influence how people choose to react and behave toward the world (e.g., Bandura, 1986; Carver & Sheier, 1998; Deci & Ryan, 1985; Fiske, 1989; Gollwitzer, 1990; Locke & Latham, 1990; Miller et al., 1960; Mischel et al., 1996; Norman & Shallice, 1986; Shallice, 1972). The research that we have reviewed in this chapter so far shows that even nonconsciously activated goals influence overt behavior, including achievement, cooperation, helping, expressing anger, seeking casual sex, and much more.

In addition to such overt behavioral effects, goals also influence more subtle types of action. Thus, Fishbach and Shah (2006) demonstrated that people possess implicit behavioral dispositions (approach, avoid) toward stimuli that are consistently desirable (high-priority goal stimuli) or undesirable (low-priority temptations). They first asked participants to generate words related to important goals and words related to associated, undermining temptations (e.g., studying, exercising, vs. movies, alcohol). They then measured participants' implicit behavioral tendencies toward those stimuli by asking participants to push or pull a standard joystick in response to each of those stimuli. Given that previous research has shown that pulling movements are faster in response to desirable stimuli, and pushing movements are faster in response to undesirable stimuli (e.g., Solarz, 1960), Fishbach and Shah hypothesized that participants would show implicit behavioral responses in accord with the desirability of the goal-related stimuli. The results showed that participants were in fact faster to pull (vs. push) a joystick toward them in response to a goal-related word; they were also faster to push (vs. pull) the joystick away from them in response to a temptation-related word. These implicit behavioral dispositions predict explicit behavior and successful self-regulation.

It should be noted that although plenty of the research we have reviewed examined the effects of goals on knowledge activation, evaluations, and emotions, it is ultimately concerned with predicting behavior. This research is grounded on the assumption that such phenomena mediate between the goal and more overt behavior. For instance, the accessibility of knowledge should eventually translate into how the person behaves (Higgins, 1996). Similarly, a large and extensive literature details how evaluative and affective experiences lead to behavioral effects (e.g., Albarracín, Johnson, & Zanna, 2005; Carver & Scheier, 1990). In this way recent work has emphasized the (often implicit) mediators at work in goal pursuit.

Summary

In this section we discussed the characteristics of goal operation, including those that involve knowledge activation, evaluations, moods and emotions, and behavior. We now turn to a new direction in the study of goals. This next section addresses how multiple goals interact, and it includes topics such as goal competition and self-control. Just as most of the characteristics we considered in the previous two sections depend on the definitional assumptions about the structure and content of goals, so too does the theory and research in the next section. In particular, this theory relies on the assumption that goals are often interconnected with one another and may contain facilitative as well as inhibitory links.

ON THE INTERACTION AMONG GOALS

Soren Kierkegaard, the Danish existentialist philosopher, instructed his readers to will only one thing (Kierkegaard, 1938). However, according to modern goal research, it is unclear whether people wish or can ever follow his recommendation (e.g., Kruglanski et al., 2002). Indeed, in previous sections we discussed how a variety of stimuli that people might naturally encounter in everyday situations, including various semantic stimuli (words), objects, relationship partners, and strangers, can activate goals. This suggests that in a typical and richly complex social environment, in which there undoubtedly exist multiple cues for different goals, the coactivation of simultaneous goals seems inevitable. In addition, people also at times consciously choose to pursue several goals simultaneously (e.g., career and family). In the face of such numerous competing pursuits, a person necessarily has to prioritize the pursuits and resolve goal conflict in order to best ensure the successful attainment of as many goals as possible (Cantor & Langston, 1989; Emmons & King, 1988; Higgins, 1997; Markus & Ruvolo, 1989; Shah, 2005). Which of multiple goals deserves priority? And when does a person decide to emphasize the pursuit of a single goal versus balance between the pursuits of several goals?

As is evident, an integral part of understanding how goals operate is an understanding of how *multiple* goals interact with each other and together influence behavior, evaluation, and emotion. Virtually all of our earlier discussion dealt with the requirements for the activation of a single goal and the characteristics of the operation of that goal. In this section, we discuss the challenge presented by multiple goals and how the interaction among goals poses a special problem for decision making and

choice. We specifically distinguish between three configurations of multiple goals. First, we consider the implications of pursuing multiple goals that are of similar centrality to the individual. We then move on to theorizing about situations in which a person is confronted with multiple goals of different centrality, which, therefore, pose a potential self-control conflict between a central goal with delayed benefits and a less central goal with immediate benefits. Finally, we look beyond the impact of several goals on a *single* action to the effects of multiple goals on a *sequence* of actions that unfold over time.

Just as in previous sections, much of the principles we consider in these areas of research derive from the definitional assumptions concerning the structure and content of goal representations that we described at the outset of the chapter. In particular, it is assumed that many goals have been activated simultaneously in the past, or are related with each other in semantic or emotional meaning. We therefore argue that many goals themselves are interconnected in memory, just as are the memories associated with a single end state. This implies then that the activation of a given goal can automatically facilitate other compatible goals or perhaps inhibit competing goals. This assumption lies at the heart of much of the research on multiple goals.

Multiple Goals of Similar Centrality

How does a person manage multiple goals of approximately equal centrality that conflict with one another? We identify two assumptions that govern research on the effect of multiple goals of similar centrality: *goal competition* and *multiple goal attainment.* In what follows, we discuss their implications for behavior, evaluation, and emotional experience.

Goal Competition

One underlying assumption of goal research is that simultaneously activated goals compete for limited motivational resources. And, because resources are limited, the pursuit of a given goal will inevitably pull resources away from another goal. In particular, goals compete for attention, commitment, and effort (Anderson et al., 2004; Baumeister, Bratslavsky, Muraven, & Tice, 1998; Förster et al., 2005; Shah et al., 2002; Shah & Kruglanski, 2002).

In one demonstration of goal competition, Shah and Kruglanski (2002) found that priming participants with a background goal (vs. a control word) undermined their commitment to the focal goal, which then hindered the development of effective means for goal pursuit and dampened participants' emotional responses to positive and negative feedback about their goal progress. In one study, participants expected to perform two consecutive tasks corresponding to two goals. While working toward the first task (i.e., the focal goal), they were subliminally primed with the name of the second task they expected to perform later (i.e., the background goal) or with a control prime. The activation of the background goal led to a decline in persistence on the first task, lower performance success, and lower emotional reactivity to success

and failure feedback. In other words, the activation of an alternative goal pulled away motivational resources from the focal goal.

Because goals compete for attentional resources, the activation of one, focal goal can sometimes lead to the inhibition of another, alternative goal in memory; in this way, the focal goal "shields" itself from alternative ones by directly reducing the accessibility of alternative goals in memory (Shah et al., 2002). Empirically, this inhibition is often reflected in the slowing down of lexical decision times to concepts that represent alternative goals. For example, Shah and colleagues (2002) demonstrated that when a goal-related concept (e.g., "study" vs. control word) was subliminally primed, it slowed down the lexical decision time to concepts related to alternative goals (e.g., "jogging"). The degree of inhibition of alternative goals was moderated by participants' commitment to the focal goal they were currently pursuing, such that only highly committed individuals (i.e., those who indicated that the goal is important) inhibited completing goals. In addition, because goals compete with each other, there is a self-regulatory advantage for inhibiting focal goals once they are accomplished, because by inhibiting completed goals, a person frees up resources to be used for new goal pursuits (Förster et al., 2005; Liberman & Förster, 2000).

An underlying assumption in research on goal competition is that goals acquire their motivational force from a limited pool of motivational resources. In other words, any act of self-regulation is by definition resource depleting (Baumeister, Heatherton, & Tice, 1994; Muraven & Baumeister, 2000; Vohs & Heatherton, 2000). Research on *ego depletion* has provided ample demonstrations for the depleting nature of self-regulatory acts across many self-regulatory domains (see Baumeister, Schmeichel, & Vohs, Chapter 22, this volume). For example, participants who were asked to control their emotional responses to an upsetting movie (vs. watching that movie with no goal in mind) were subsequently less able to persist on holding a handgrip. Or, in another study, participants who suppressed forbidden thoughts (vs. no suppression condition) were subsequently less likely to persist on trying to solve unsolvable anagrams (Muraven, Tice, & Baumeister, 1998).

But because goal pursuits are resource depleting, people withdraw from a current, effortful goal in order to save their resources for another upcoming, goal-related task. For example, dieting students might stop trying to control their food intake just before they undergo an important academic test. In general, over a lifetime's worth of experience with regulating limited motivational resources, people may develop strategies of resource conservation and resource management, which are designed to save self-regulatory recourses for future goal pursuits (Shah, 2005). These resource management processes may further operate outside conscious awareness. Shah and his colleagues found that participants who were subliminally primed with the name of an upcoming difficult task (vs. nonword control) were less likely to put effort into the present task, took longer breaks, and consumed more juice, which they were told was helpful for the subsequent task (Shah, Brazy, & Jungbluth, 2005). This work

suggests that resource management is often strategic (while still nonconscious) and can follow different patterns of self-regulation, such that lower efforts follow or precede actual physiological depletion. Because resource management is strategic, the extent of decline in goal performance also depends on one's lay belief that another act of self-regulation is or will be depleting (Mukhopadhyay & Johar, 2005).

Taken together, there is converging evidence for the phenomenon of goal competition. How does goal competition influence a person's evaluations, emotions, and behavior? First, beyond the effects of the activation and operation of a single goal on evaluations, goal competition presents some consequences for patterns of evaluations more generally. One such consequence is instability of evaluations over time. Because various goals wax and wane in accessibility, the evaluations of objects related to those goals (means and hindrances) will also fluctuate. This means that a decision that is made according to the goal relevance of options at one point in time may not be as optimal at a later time when the goal relevance of those same options has changed. This can be particularly troublesome if the accessibility of the goals and the corresponding fluctuation of evaluations all take place nonconsciously, without the person's awareness. For example, a person who selected flight tickets based on low price may find this selection incompatible with another, competing goal of saving time, which becomes salient later on. Because the person may be unaware of this goal conflict, he or she may experience little satisfaction with the choice and may regret it if the accessible goal has changed from saving money to saving time. In this way, the fluctuating nature of goal activation might sometimes introduce negative emotional consequences and mean that people are often somewhat dissatisfied with their choices.

Second, what implications does goal competition have for behavior? With respect to behavioral effects, a normative choice theory (e.g., the multiattribute utility theory, or MAUT) entails that when people want to make a single choice in a way that will meet several goals (e.g., ordering food that is healthy, tasty, and not too expensive), they should integrate these various goals by weighing their relative importance (e.g., Baron, 2000; Keeney & Raiffa, 1976). However, our analysis attests that the relative weight of a goal in the decision process is not fixed, and therefore integration is rarely optimal. That is, because multiple goals that are brought into a decision process can directly interfere with the attainment of each other, people may tend to overemphasize a focal goal in their decision while discarding other background goals that are temporarily inhibited by the focal goal. For example, when primed with "ease," students may choose to work on a project that is easy while completely overlooking other goals, such as their level of interest in any particular project.

Multiple Goal Attainment

We assume that the pursuit of multiple goals is characterized by a desire for *multiple goal attainment*. According to this assumption, given the presence of several salient goals and limited motivational resources, self-regulators search for attainment means that are *multifinal*, that is, means that are linked to the attainment of several goals simultaneously (Kruglanski et al., 2002). For example, a person may prefer to dine out (vs. dine in) in order to satisfy both hunger and various social motives (to see and be seen, etc.), or commuters may choose to commute by bike (vs. car) in order to save money and keep in shape.

What are the implications of the assumption that people try to find means that can meet as many active goals as possible? Multifinal means are by definition scarce because they constitute a subset of the original set of means to a goal and are therefore more difficult to find. Thus, when individuals wish to achieve multiple goals, any increase in the number of accessible goals negatively affects the number of satisfactory means, thus elevating the difficulty of the search (Kruglanski et al., 2002; Tversky, 1972). For example, while many restaurants will satisfy one's hunger, somewhat fewer of them will provide an interesting scene, and fewer still are also not too expensive. In general, when holding multiple goals people end up searching longer for satisfying means and they also end up choosing "compromise" options that are less effective at satisfying each goal separately (Simonson, 1989). Moreover, because compromise options imply that none of the goals is met very strongly, people may at times choose to abandon the search for multifinal means altogether and focus on only one goal.

The search for multifinal means also has consequences for evaluation, emotions, and behavior. The preference for multifinal means may have an adverse effect on the evaluation of the selected choice options if these options are only partially associated with the attainment of any salient goal (e.g., when people order food that is moderately tasty and moderately healthy). We argued earlier that goal-facilitating stimuli acquire positive value and goal-thwarting stimuli acquire negative value (Brendl et al., 2003; Ferguson & Bargh, 2004; Fishbach et al., 2004). However, in the course of pursuing multiple goals, an attainment means to one goal can potentially interfere with satisfying another goal, and hence, although this means may be positively evaluated because its facilitation of one goal, it might also tend to be negatively evaluated because it hinders another goal. Thus, even though a given means to an active, focal goal should be particularly positive because it facilitates that goal, the simultaneous activation of another goal, one that the given means cannot facilitate, can end up dampening the positivity of that means. One consequence of this is that the quest for multifinal means may undermine the evaluation of a given available means and lead to choice deferral and decision aversion because none of the means seems satisfying (Dhar, 1996, 1997; Iyengar & Lepper, 2000; Tversky & Shafir, 1992). As one example of this notion, Iyengar and Lepper (2000) found that students are more likely to choose a class assignment when offered a limited array of a few options that activate fewer goals, compared with when more options are presented. It also follows that holding a single goal (or fewer goals) should lead to the positive evaluation of means and decision-seeking behav-

iors related to this goal. For example, a student who wishes to select an interesting project to work on would be less likely to defer her choice and be more satisfied with the selected project than her classmate, who might share equal interest in selecting an interesting and easy project.

In terms of the emotional experience of goal pursuit more generally, the quest for multiple goal attainment can lead to mixed emotions and ambivalence when people strive toward incongruent ends (e.g., academics and leisure) and a means to one end (e.g., a textbook) acts as a hindrance to another. Under these circumstances, the same object or activity may be experienced both positively and negatively at the same time and end up seeming ambivalent (e.g., Cacioppo, Gardner, & Berntson, 1999). For example, a student who works on an easy but uninteresting project would be both satisfied and unsatisfied with her choice.

It was shown that the preference for multifinal means has further behavioral implications, and, in general, people prefer choice alternatives that partially meet, or strike a compromise between, several goals at once rather than ones that fully meet or highlight a single goal (e.g., Simonson, 1989). As a demonstration, Simonson asked participants to evaluate several consumption products (e.g., apartment, calculator, and television). Participants exhibited a greater preference for options that struck a compromise between several goals (e.g., large/small size and low price) than those that accomplished a single goal (e.g., provided low price).

Because in a multifinal choice the number of activated goals is inversely related to the number of acceptable means, it follows that there should be a negative relationship between the number of goals and the number of acceptable means that a person would choose to pursue. This pattern was demonstrated in a study conducted around lunchtime by Köpetz, Fishbach, and Kruglanski (2006), in which participants listed three goals that they had for that day (vs. goals already accomplished on that day), other than getting lunch, before indicating the number of different lunch options that they would consider. Compared to participants in the control (accomplished goals) condition, those for whom actual goal alternatives were activated listed significantly fewer food options in which they were interested.

Though highly desirable to have, multifinal means may suffer a disadvantage as well in that they may be perceived as less effective and instrumental to goal attainment. This may be so because multifinal means can be objectively less effective. But this may also be because perceived effectiveness of a given means to goal attainment is determined in part by the strength of the association between that means and the goal, with stronger associations leading to higher perceived effectiveness. When the number of goals attached to a given means increases, each association becomes weaker, as demonstrated by a lower retrieval rate of the associated goal when the means is activated (Anderson, 1983; Anderson & Reder, 1999). The result is a dilution of the means–goal association, which may reduce the perceived effectiveness of the means with respect to the goal. In a demonstration of such a *dilution effect*, Zhang, Fishbach, and Kruglanski (in press) found that when participants considered the different goals (e.g., building muscles and losing weight) that a single means (e.g., working out) could satisfy, an increase in the number of goals resulted in a reduction in the perception of the instrumentality of the means with respect to each goal.

As a result of a dilution effect, means that are connected with multiple goals are also less likely to be chosen and pursued when a single (vs. multiple) goal needs to be fulfilled. For example, participants were less likely to use the writing function of a pen that had also been used as a laser pointer (vs. was not used as a laser pointer) when they only needed to write (Zhang et al., in press). It appears that multifinal means are desirable when the individual foresees the pursuit of multiple goals, but those same means are judged as less effective and they are less likely to be selected when the individual focuses on a single goal.

Self-Control Conflicts

We have identified two underlying mechanisms for managing multiple goals that are of similar centrality: goal competition and multiple goal attainment. But people often hold multiple goals that differ in their importance or centrality, and these goals can impose a self-control dilemma. In what follows, we address such a situation.

People face a self-control problem when the attainment of their central, higher-order goals comes at the expense of foregoing low-order desires or temptations (Ariely & Wertenbroch, 2002; Baumeister, Heatherton, et al., 1994; Dhar & Wertenbroch, 2000; Gollwitzer & Moskowitz, 1996; Kivetz & Simonson, 2002; Kuhl & Beckmann, 1985; Loewenstein, 1996; Metcalfe & Mischel, 1999; Rachlin, 1997). For example, the pursuit of academic excellence, professional success, or fitness and general health comes with the expense of foregoing low-order although salient goals (e.g., partying, taking long vacations, or consumption of fatty foods, respectively). As these examples demonstrate, temptations are defined within a given situation and with respect to the higher-order goals at hand. For example, while going on vacations interferes with pursuing professional success, thoughts about one's career can undermine one's ability to relax and enjoy a vacation. This context-specific definition of temptations suggests that when individuals strive toward multiple goals, any goal can potentially constitute an interfering temptation with respect to another, currently more central, goal. In response to self-control dilemmas, people exercise self-control (Dhar & Wertenbroch, 2000; Gollwitzer, 1999; Kivetz & Simonson, 2002; Kuhl, 1986; Muraven & Baumeister, 2000), and these self-control operations influence behavior, evaluation, and emotion.

The Operation of Self-Control through Construal

What do self-control operations entail? One category of such operations involves the construal of the self-control conflict in abstract (vs. concrete) terms. For example, in

one of the first systematic studies of self-control operations, Walter Mischel and his colleagues found that an abstract representation of the immediate reward (e.g., a small candy) helped children wait for the delayed, preferred reward (e.g., a large candy) (Mischel, 1964; Mischel & Mischel, 1983; Mischel, Shoda, & Rodriguez, 1989). According to Mischel and colleagues, abstract representations facilitate success at self-control because they activate a "cool" (cognitive and evaluative) system, while suppressing a "hot" (emotional and operating) system (Metcalfe & Mischel, 1999; Mischel & Ayduk, 2004). When a person is in an evaluative mode, rather than an action mode, the person is more likely to follow a higher-order goal. As a recent demonstration of this idea, Kross, Ayduk, and Mischel (2005) manipulated abstractness by asking participants to elaborate on the "why" versus "how" aspects of their experience (Freitas, Gollwitzer, & Trope, 2004; Vallacher & Wegner, 1987). They found that participants displayed improved self-control in coping with anger-provoking experiences when they had elaborated on why they had the experience (an abstract construal) as opposed to how they exactly felt (a concrete construal).

In addition, abstract processing increases success at self-control by directing people's attention to their central, high-order (vs. low-order) goals (e.g., Fujita, Trope, Liberman, & Levin-Sagi, 2006; Rachlin, 2000). According to a construal-level analysis (e.g., Trope & Liberman, 2003), abstract processing is associated with high-level construal and it can facilitate success at self-control by directing people's attention to high-order goals. In support of this analysis, Fujita and colleagues (2006) found that asking participants to generate superordinate category labels (abstract processing) versus subordinate exemplars (concrete processing) for a variety of common objects, increased participants' subsequent motivation to undergo a difficult yet important test. Presumably, abstract processing allows one to successfully ignore the immediate aversiveness of adhering to high-order goals.

The Operation of Self-Control through Evaluation and Emotion

Another category of self-control operations includes *counteractive control* processes, which offset the influence of temptations on adherence to a central goal. Of particular interest, counteractive control processes influence the evaluation of and the affective experience of choice alternatives related to a central goal and less central temptations when these are in conflict (Fishbach & Trope, 2005; Trope & Fishbach, 2000, 2005).

Research on counteractive control attests that when people anticipate a self-control problem, they proactively increase the desirability of adhering to a goal relative to yielding to temptation. The presence of tempting alternative may thus influence goal-directed behavior in two opposite directions: directly, the perception of tempting alternatives decreases the likelihood of adhering to a more central goal; but, indirectly, the perception of tempting alternatives triggers the operation of counteractive control, which then acts to increase the likelihood

of adhering to the goal. For example, an invitation to go out on the night before an important exam directly decreases the likelihood of studying, but it may further set into action counteractive bolstering of the value of studying, which increases the likelihood of engaging with this activity. As a result of counteractive control such invitation has no effect on studying for the exam overall.

Some of the counteractive control operations that people employ involve changes in the actual choice situation. For example, people may impose penalties on themselves for failing to adhere to an important goal (e.g., failing to abstain from smoking), or they may eliminate certain choice alternatives such as cigarettes or fatty foods from their environment, thus making their decision irreversible (Ainslie, 1975; Green & Rachlin, 1996; Rachlin & Green, 1972; Schelling, 1978, 1984; Strotz, 1956; Thaler, 1991; Thaler & Shefrin, 1981). In addition, people counteract temptations by changing the positive evaluation of adhering to their goals and pursuing temptations (see also Kuhl, 1986; Mischel, 1984), and they further change the perceived emotional significance of goals and temptations.

To demonstrate changes in evaluation in response to temptation, Trope and Fishbach (2000) offered participants an opportunity to take a diagnostic test that was described as requiring abstinence from food containing glucose for either a long or a short period (3 days vs. 6 hours). Participants evaluated the test more positively when it required a long (vs. short) period of glucose abstinence (i.e., when the temptation to forego the test was stronger). They also found that whereas the length of the abstinence directly decreased interest in the test, indirectly it increased interest in undergoing the test, by increasing its positive evaluation. Other studies demonstrated similar effects on the emotional reactivity to succeeding on goal-related activities. When facing strong versus weak temptations, participants reported that goal pursuits were associated with more intense pride while failing on goal pursuits was associated with more intense guilt.

Bolstering the value and emotional reactivity of a goal in response to a temptation can be deliberate and may require some level of conscious awareness, intentionality, and processing resources (Baumeister et al., 1998; Mischel, 1996; Muraven & Baumeister, 2000; Trope & Neter, 1994; Vohs & Heatherton, 2000). However, our analysis suggests that goals and the process of self-regulation may not require consciousness and intentionality, and it follows that processes of self-control and overcoming temptations can also proceed nonconsciously (Ferguson, 2006, in press; Fishbach et al., 2003; Gollwitzer, Bayer, & McCulloch, 2005; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). One such implicit strategy involves the activation of goal representations in response to cues for temptations (Fishbach et al., 2003). For example, Fishbach et al. assessed the lexical decision time to respond to words representing a potential goal following the subliminal presentation of words representing potential temptations. They found that subliminal temptation primes (e.g., "drugs" vs. control words) facilitated the lexical times for goal-related targets (e.g.,

"bible"). In addition, goal-related primes (vs. control words) inhibited the lexical time for temptation-related targets, and these implicit and asymmetrical activation patterns were shown to increase success at self-control.

Other implicit self-control operations involve changes in the implicit positivity of goals and temptations. For example, Fishbach, Zhang, and Trope (2006) documented an implicit negative evaluation of temptations and an implicit positive evaluation of goals when these two were in a conflict. In one study, dieters (vs. non dieters) responded faster to positive concepts after being subliminally primed with words related to dieting (e.g., diet), and they responded faster to negative concepts after being subliminally primed with words related to food (e.g., cake). We claimed that such changes in implicit positivity directly influence behavior (e.g., Ferguson & Bargh, 2004). Indeed, as indicated earlier, Fishbach and Shah (2006) documented a similar tendency to automatically approach stimuli related to a goal (through faster pulling responses) and automatically avoid stimuli related to temptation (through faster pulling responses). These implicit approach and avoidance responses predicted the attainment of high-order interests. For instance, the rate of responding by pulling a joystick in response to academic targets (e.g., "library") and by pushing a joystick in response to nonacademic, tempting targets (e.g., "party") predicted student participants' GPA scores.

The Bidirectional Relationship between Emotions and Self-Control

We have thus far claimed that self-control operations involve changes in evaluation and emotions. Here we consider more generally the relationship between emotions and success at self-control. We suggest that the resolution of a self-control conflict has implications for one's emotional experience, and in addition, people's emotional experience and mood influence how they resolve a self-control conflict. In what follows we address these influences.

First, with regard to the effect of self-control on people's emotions, whereas the successful resolution of a self-control conflict is characterized by the experience of feelings such as pride, a failure at self-control is characterized by feelings such as shame and guilt. These emotions (e.g., pride vs. guilt) are high-level, self-conscious emotions that people experience when they engage in a self-control behavior directed toward higher-order goals, and they are qualitatively different from more basic emotions such as happiness and fear that are low level and signal immediate rewards or punishments (e.g., "hot" feelings; Metcalfe & Mischel, 1999). Presumably, part of the reason that people adhere to high-order goals is because they wish to experience positive self-conscious feelings and avoid negative self-conscious feelings (Beer & Keltner, 2004; Giner-Sorolla, 2001; Tangney, Miller, Flicker, & Barlow, 1996; Tracy & Robins, 2004). In support of this notion, guilt is associated with failing to maintain social relationships and with overeating, and therefore, considering one's possible guilty feelings leads to improving social relationships (Baumeister, Stillwell, & Heatherton, 1994) and reducing the amount of fatty food eaten by dieting individuals (Giner-Sorolla, 2001).

But how do existing affective states influence the subsequent motivation to exercise self-control? This second question refers to the effect of emotions on self-control, and previous research poses an apparent contradiction in addressing it. Some research has claimed that positive mood undermines self-control (e.g., Wegener & Petty, 1994, 2001), while others have claimed that positive mood improves self-control (e.g., Aspinwall, 1998; Raghunathan & Trope, 2002). Specifically, researchers have claimed that positive mood impairs self-control because happy (vs. unhappy) people prefer activities that prolong the quest for positive mood. For example, Isen and Simmonds (1978) reported that participants in a happy mood were less helpful than those in a neutral mood when the helping behavior involved reading unpleasant information. Similarly, Wegener and Patty (1994) found that happy (vs. neutral or unhappy) participants chose to see more happy films but not more interesting films. Conversely, other mood researchers found that positive mood is often "used" for accomplishing tasks that have immediate costs and require self-control (Aspinwall & Taylor, 1997; Raghunathan & Trope, 2002; Trope & Pomerantz, 1998). For example, research on the delay of gratification attests that happy (vs. unhappy) children are better able to wait for a delayed, preferred reward than for an immediate, less preferred reward (Moore, Clyburn, & Underwood, 1976; Schwarz & Pollack, 1977). In addition, research on negative feedback seeking (i.e., feedback about a person's shortcomings) reveals that people take an increased interest in this potentially useful information when positive mood is induced. For example, caffeine drinkers who were induced to feel good were more attentive to negative information about the health effects of caffeine (Raghunathan & Trope, 2002, see also Trope & Neter, 1994). Also consistent with this latter possibility, there is research showing impaired self-control ability during negative mood states (Leith & Baumeister, 1996; Tice, Bratslavsky, & Baumeister, 2001).

How can these areas of work be reconciled? Our view assumes that people can use their mood as information about the task at hand (e.g., Schwarz & Clore, 2003) and, in particular, we suggest that moods are seen as signals to either adopt or reject any accessible goal. That is, while the experience of positive mood signals to people that they should approach a stimulus, the experience of negative mood signals to them that they should avoid a stimulus (e.g., Cacioppo et al., 1999; Higgins, 1997; Larsen, McGraw, & Cacioppo, 2001; Lazarus, 1991). Accessible goals are one set of stimuli that people need to decide whether to approach or avoid. Thus, it follows that a positive mood should increase people's tendency to adopt any accessible goal, whether the goal is high order (e.g., self-improvement) or low order (e.g., mood management). In this way, happy people should perform better on self-control tasks when they hold an accessible high-order goal but perform poorly when they hold an accessible low-order goal.

In support of this analysis, Fishbach and Labroo (in press) found that when self-improvement goals were accessible, happy (vs. unhappy) participants invested more effort in a task that furthered the goal, even if the task was unpleasant or demanding. Conversely, when mood management goals were made more accessible, happy people invested less effort than unhappy people. In one study that tested for charity donations, happy (vs. unhappy) participants were asked to describe what they generally do to "be better" (high-level, self-improvement) versus "feel better" (low-level, mood enhancement). They were then asked to participate in a local charity campaign that promoted protecting young children from injury or death by improving children's product safety. Happy (vs. unhappy) participants donated more money when they had considered the self-improvement goal but not when they considered the mood management goal. Other studies replicated the effect of mood on self-control by nonconsiously priming self-improvement or mood management goals, which further demonstrates that the direction of the relationship between mood and success at self-control depends on a person's accessible goal.

The Pursuit of Multiple Goals in a Choice Sequence

The previous sections refer to situations that involve the consideration of multiple goals of either similar or different centrality, which influence the selection of an action that secures their attainment. Notably however, few goals can be completed by the execution of a single action; rather, goals frequently require taking several actions that maintain goal pursuit over time. The challenge that individuals face over repeated choice situations is to decide between emphasizing, or highlighting, the pursuit of a single goal and balancing between several goals. In this section, we address this challenge and consider how the specific strategy of goal pursuit (highlighting a single goal vs. balancing among several goals) that an individual employs for actions that unfold over time may influence their immediate behavior, evaluations, and emotional experience.

As stated previously, when individuals simultaneously hold multiple goals that they wish to pursue over time, self-regulation may follow one of two possible dynamics: highlighting the pursuit of a single goal in several consecutive actions versus balancing among several potentially incongruent goals across several actions (e.g., Fishbach & Dhar, 2005, in press; Fishbach, Dhar, & Zhang, 2006). For example, consider a person who chooses to dine out and wishes to both save money and seek pleasure. In the absence of compromise options, that person can balance between these conflicting goals by choosing an expensive appetizer and a less expensive entrée, or, the person can highlight one of these goals (e.g., by choosing an expensive appetizer and an expensive entrée). *Choice highlighting* refers to a dynamic of self-regulation where pursuing one goal enhances the commitment to this particular goal relative to competing ones and motivates complementary actions over time. *Choice balancing* refers to a dynamic of self-regulation where pursuing one goal liber-

ates the individual to pursue other, conflicting goals at the next opportunity (Dhar & Simonson, 1999; Fishbach & Dhar, 2005; Fishbach, Dhar, & Zhang, 2006).

What then determines a person's interest in choice highlighting versus choice balancing? One factor is how the person interprets the meaning of an initial action that is congruent with one of the goals. It is possible that a person could interpret such an action as indicating a strong commitment to the respective goal. If so, such an interpretation would then increase the motivation to make similar, complementary actions and to inhibit any competing goals (Aronson, 1997; Atkinson & Raynor, 1978; Bem, 1972; Feather, 1990; Festinger, 1957; Locke & Latham, 1990). The following choices then would be considered *choice highlighting* because the person would be prioritizing one goal over the others. On the other hand, it is also possible that a person might interpret that initial choice as indicating progress toward that goal. If so, that person might consequently relax his or her efforts toward the goal and begin to attend to the other competing goals. In this way, the interpretation of a goal-congruent action as progress signals the reduction of a discrepancy between the present state and goal attainment (Carver & Scheier, 1998; Miller et al., 1960; Powers, 1973). The person's choices would thereafter be considered *choice balancing* because he or she would be attempting to pursue multiple goals as much as possible, rather than focusing on a particular goal.

Research by Fishbach and Dhar (2005) demonstrated that people do indeed make inferences concerning goal commitment or goal progress, and these inferences activate different dynamics of self-regulation when there are multiple goals at stake. As an illustration, these researchers found that when initial academic success was interpreted as indicating greater commitment to academic goals, students were subsequently more interested in pursuing additional academic tasks and they were less interested in pursuing incongruent leisure activities. Yet, this same level of initial academic performance decreased interest in additional academic tasks and increased interest in balancing between initial success and subsequent choice of leisure activities when students inferred that progress had been made on the academic goals.

In addition to an initial goal-congruent action being able to be interpreted in multiple ways, an initial failure to pursue a goal is also open to multiple interpretations. Such a failure can signal either a lack of sufficient commitment to a goal or a lack of progress toward the attainment of that goal (Fishbach, Dhar, & Zhang, 2006). If people infer low goal commitment based on an initial failure, they tend to subsequently highlight this failure by disengaging from the goal (Cochran & Tessser, 1996; Soman & Cheema, 2004). If, however, following failure people infer a lack of progress toward the goal to which commitment remains intact, they tend to balance between the initial failure and their subsequent greater motivation to work harder by choosing additional actions that pursue this goal (e.g., see research on self-completion theory—Brunstein & Gollwitzer, 1996;

Wicklund & Gollwitzer, 1982). Thus, for example, failure on an exam decreases the subsequent motivation to study if it signals low commitment to doing well academically but increases the subsequent motivation to study if it signals the absence of progress toward the goal of academic excellence.

Previous goal research has often focused on one of these dynamics only. Thus, as an example of choice balancing, Shah and Kruglanski (2002) examined goal substitution. In one study, these researchers framed two anagram tasks as relating either to the same goal or to different goals. In one condition, one task was linked to promotion goal and the other task was linked to prevention goal (see Higgins, 1997); in the other condition the tasks were linked to the same prevention or promotion goal. They found that success at the first task decreased performance on the second when it served the same (vs. different) regulatory goal, because participants experienced goal attainment. But failure at the first task increased performance on the second task if both served the same (vs. different) goals, because participants did not experience attainment. Such substitution was shown to lead to ironic results when people substitute intention for action (e.g., Prelec & Bodner, 2003; Tesser et al., 1996). For example, Monin and Miller (2001) gave participants an opportunity to disagree with blatantly sexist statements, and those who received the opportunity (vs. not) were later more willing to favor a man for a stereotypically male job, presumably because the first task was sufficient to establish their moral credentials.

In yet another demonstration of ironic substitution, Fishbach and Dhar (2005) found that an initial sense of successful weight loss increased dieters' tendency to indulge. In their study, dieting participants were asked to draw a line that represented the distance between their current and ideal weight on a scale that either had –5 lbs. or –25 lbs. as its maximum discrepancy. Providing a scale with a wide range (–25 lbs.) created an illusion of smaller discrepancy (e.g., 4% vs. 20%, for a person who would like to lose 1 lb.), which led to greater perceived goal attainment. As a result, those who completed a wide (vs. narrow) scale were more likely to choose a chocolate bar over a low-calorie snack on a subsequent, supposedly unrelated choice task.

But how does substitution influence everyday behavior? People's intuitive belief in balancing between multiple goals leads them to seek variety and switch among goals when choosing items such as foods or leisure activities (Drolet, 2002; Ratner, Kahn, & Kahneman, 1999; Simonson, 1990). As a result, people sometimes end up choosing the less preferred item that is associated with the less valuable goal and that undermines choice satisfaction. According to the current framework, a variety-seeking behavior is driven by individuals' beliefs about satiation and maximizing the attainment from multiple goals. Therefore, for example, people incorporate more variety when simultaneously choosing several items than when choosing one item at a time (Simonson, 1990), because of their overestimation of the rate at which they will experience attainment (Read & Loewenstein, 1995).

However, people also demonstrate choice highlighting when they infer commitment and end up performing congruent behaviors. For example, research by Fishbach, Ratner, and Zhang (2006) demonstrated that variety-seeking behavior is attenuated and even reversed (as indicated by a greater preference for a previously selected item in a sequence) if participants consider their stable preferences based on their initial choice rather than the extent of satiation on that goal. In general, consistency theories in social psychology documented a desire to express congruency in a behavioral sequence; thus once a person engages in an initial action, the person feels that she should pursue similar actions (e.g., Aronson, 1997; Bem, 1972; Cialdini, Trost, & Newsom, 1995; Heider, 1958). For example, once participants agreed to display a small sign to advocate driving safety, they were more likely to display a larger sign to advocate the same goal compared with those who did not display the small sign (Freedman & Fraser, 1966). Other researchers have further indicated that behavioral consistency is associated with emotional benefits (Aronson, 1997; Festinger, 1957).

Future Plans Influence Present Actions

We described the effect of past actions on the present preference for actions that pursue the same or different goals, but what about the effect of future, planned actions? Do these actions also influence which goals a person decides to pursue in the present? There is some evidence that planned actions do influence present ones (Bandura, 1997; Oettingen & Mayer, 2002; Taylor & Brown, 1988). Thus, research on self-efficacy (Bandura, 1997) and positive illusions (Taylor & Brown, 1988) attest that exaggerated beliefs in actions that will be taken in the future lead to higher motivation to work harder on that goal in the present (see also Atkinson, 1964; Weiner, 1979). But others suggested that future plans can also undermine the motivation to work on a goal in the present. For instance, Oettingen and Mayer (2002) found that positive *expectations* of future goal pursuit lead to greater effort and successful performance on a focal goal in the present. But the reverse was true for positive *fantasies*, which are images depicting future goal attainment. Fantasies predicted lower effort on a focal goal in the present. As a demonstration, in one study college students who expected to start a relationship with a person were more likely to start an intimate relationship compared with those who experienced positive fantasies about future romantic success.

But regardless of the direction of the influence on present actions (more vs. less goal pursuit), what is the relative impact of future plans compared with past actions? Building on the observation that people are unrealistically optimistic (Buehler, Griffin, & Ross, 1994; Weinstein, 1989; Zauberman & Lynch, 2005) and therefore believe more goal-congruent activities will be accomplished in the future than in the past, it is possible that future plans have a greater impact on immediate goal pursuits than retrospection on past pursuits (Zhang,

Fishbach, & Dhar, 2006). The direction of the impact should then depend on the framing of the goal pursuit as indicating commitment versus progress. When people consider their level of goal-commitment, thinking about plans for future (vs. past pursuits) leads to greater persistence on the goal in the present. Conversely, when people consider their level of goal progress, thinking about future plans (vs. past pursuits) justifies disengagement from the focal goal in the present. As a demonstration, Zhang and colleagues (in press) asked gym members to estimate either the frequency of their exercise in the coming year or the frequency of their actual exercise regimen last year. Those who considered future (vs. past) exercise were more likely to consume healthy food in the present, if the exercise was framed as increasing commitment to the health goal. But envisioning future (vs. recalling past) exercise decreased the relative preference for healthy food in the present when the exercise was framed as increasing progress toward the health goal.

When Do People Highlight versus Balance Multiple Goals?

We described evidence in support of people's preference for making congruent choices that highlight a single goal when they consider their goal commitment, and people's preference for making incongruent choices that balance between different goals when they consider their goal progress. Several variables determine the relative focus on commitment versus progress. First, these inferences may be determined by situational cues, such as framing questions that direct one's attention to different aspects of goal-related actions. For example, Zhang and colleagues (in press) manipulated the degree of optimism that gym members experienced (following Taylor, Pham, Rivkin, & Armor, 1998) before asking them whether by exercising they are "getting closer" to their workout objectives (progress frame), or whether they are "feeling more committed" to their workout objectives (commitment frame). High levels of optimism had opposite consequences for the subsequent interest in healthy eating: dampening the interest among those who focused on the progress from their actions and increasing the interest among those who focused on the commitment from their actions.

Second, the degree to which individuals infer progress or commitment from their actions depends on their relative attention to the concrete aspects of the action in comparison with the corresponding abstract goal that initiated this action. When people consider the attainment of the action itself, they may experience some of the benefits associated with goal fulfillment, which motivates them to move temporarily away from the goal. On the other hand, when the focus is on an overall, more abstract goal, the same level of successful attainment provides evidence for a person's higher commitment to, and identification with, the goal more than it indicates progress. Fishbach, Dhar, and Zhang (2006) tested this idea by giving participants an opportunity to work on two independent verbal ability tests that represented actions to an academic achievement goal. The first test had correct solutions,

whereas the second test was unsolvable. They found that those who received high (vs. low) success feedback on the first test exhibited lower motivation to persist on a second similar but unsolvable test. This pattern replicated Shah and Kruglanski's (2002) findings on substitution. However, when in another condition an overall achievement goal was nonconsciously primed, high (vs. low) success feedback elicited greater motivation to persist on the second test, because success signaled greater commitment.

In another study, Fishbach, Dhar, and Zhang (2006) tested temporal distance (e.g., Trope & Liberman, 2003) as another variable that determines the relative focus on the action itself (for proximal actions) versus the abstract goal that initiated it (for distant actions). They found that actions that were scheduled in the near future signaled their own attainment, whereas actions that were scheduled in the distant future signaled commitment to an overall goal. For example, studying for an exam in the present signaled the accomplishment of an academic task whereas studying in the future signaled the commitment to an academic goal. These inferences in turn increased the amount of time that participants intended to invest on additional actions to an overall goal that were scheduled in the distant versus proximal future (e.g., study for a second exam).

Third, with regard to goals with an obvious end state, the relative focus on commitment versus progress may depend on whether a person attends to the amount of goal pursuit that was accomplished, as opposed to the remaining amount of goal pursuit that is required to meet the goal. Whereas completed actions establish a sense of commitment by signaling to the person that the goal is important, actions that are yet to be taken highlight the amount of progress that is still needed for goal accomplishment. For example, in the decision to participate in a charity campaign, learning about the amount of seed money that was collected thus far provides information regarding the importance of the campaign, which establishes commitment, whereas learning about the amount of money that is needed to complete the campaign goal provides information that establishes a sense of goal progress. It follows that uncommitted individuals, who wish to assess whether the goal is important, would be more influenced by learning about accomplished actions, whereas committed self-regulators, who wish to assess the required efforts in order to accomplish the goal, would be more influenced by considering the remaining distance for goal completion. These predictions were recently tested by Koo and Fishbach (2006) who conducted a field study as part of an HIV/AIDS initiative. Participants in their study were potential donors who were either committed individuals who donated money before or uncommitted individuals who did not donate money before. Uncommitted participants were more likely to donate and donated higher amounts when they read about the amount of money raised thus far as opposed to the amount of money that is still required, whereas committed participants were more likely to donate and donated higher amounts when they read about the amount of money still required than the amount of money that was raised.

Effects on Evaluations and Emotions

These aforementioned dynamics of multiple goal pursuit have further implications for evaluation and emotion. We proposed that in self-control situations, people secure the attainment of an important goal by increasing the positive evaluation of the high-order goal relative to temptations (e.g., Fishbach & Trope, 2005; Trope & Fishbach, 2000). But what if people perceive an opportunity to balance between the goal and temptations and, hence, view these options as complementary rather than competing? For example, a dieter may choose to balance between low- and high-calorie foods, or choose to highlight a choice of low-calorie foods. We next explore how each of these dynamics influences the evaluation of choice options.

When people plan to highlight the pursuit of a single goal across several actions, they should generate a positive evaluation of objects or means related to this goal, and a negative evaluation of objects or means related to competing alternatives (i.e., temptations). Conversely, when people wish to balance between goals and temptations that they see as complementary rather than competing, they should express a more positive evaluation of objects or means related to the tempting option relative to those that are related to the goal option. The reason for this latter evaluative pattern is that goals (relative to temptation) offer delayed benefits (Ainslie, 1975; Rachlin, 1997; Thaler & Shefrin, 1981), and therefore when people expect to balance, they prefer to pursue the temptation in the present and postpone goal pursuits for the future and, thus, maximize the attainment from both. For example, people may choose to indulge today and start a diet tomorrow and therefore express a positive evaluation of fatty food in the present.

In studies that demonstrated these evaluative patterns, Fishbach and Zhang (2006) manipulated the perception of items related to goals and temptations as complementing each other versus as competing with each other. Complementary items were presented in one choice set and competing items were presented in two different choice sets. They found, for example, that when healthy and unhealthy foods are included on one menu, participants saw them as complementing and planned to balance between them. As a result, the value of unhealthy foods was higher relative to the value of healthy foods. However, when these foods were presented apart in two different menus, participants saw them as competing with each other and planned to highlight the consumption of healthy food. As a result, the value of healthy foods was higher. Importantly, when these items were evaluated in isolation (i.e., in the absence of cues for alternative goals), they had similar value.

The evaluation of items related to multiple goals has further influence on the emotional experience that characterizes the self-regulatory process and goal attainment. That is, when people wish to highlight the pursuit of a single goal in a sequence, actions related to this goal are associated with positive emotions and actions that interfere with it are associated with negative emotions. However, when people wish to balance between several goals,

actions directed toward one goal can interfere with the attainment of another goal and, hence, might be less associated with positive emotions. Similarly, actions that interfere with the initial goal can advance the pursuit of another goal and be less emotionally negative. For example, socializing before an important exam is less guilt provoking if a student plans to balance between academic and social pursuits. The result is that when people consider the pursuit of multiple goals across several actions and over time, the emotional experience from goal-related actions is less intense.

Summary

Research reviewed in this section addresses the phenomena surrounding those situations in which multiple goals are at stake. We considered the effects of goals that are of similar centrality as opposed to goals that vary in their relative centrality and impose a self-control dilemma. We also described research on how multiple goals interact when a person only considers a single act of self-regulation, as opposed to when a person considers the pursuit of multiple goals over time and across several decisions. Based on research reviewed here, we suggest that multiple goals (vs. a single goal) present unique implications for people's behaviors, evaluations, and emotions. We further propose that these effects follow from our definition of the goal construct that we outlined in the first section.

CONCLUDING REMARKS

Multiple researchers across various domains of psychology have documented the wide-ranging effects of goals on behavior, attitudes and evaluations, and emotions and moods. In this chapter, we sought to identify the main principles from this literature by focusing on how goals become activated in the first place, the mechanisms that underlie and enable their operation, and the ways in which they interact with one another. Our analysis was grounded in basic definitional assumptions about goals concerning their structure in memory and the nature of memories assumed to be relevant to goals. We attempted to showcase throughout the chapter how many of the recent findings we reviewed derive from these definitional assumptions.

One central distinction between past research and the current framework concerns the degree to which people are aware of goal activation and pursuit. Throughout most of the last century of empirical and theoretical psychology, goals have been commonly understood as objects, states, or experiences that people consciously want or do not want (e.g., Gollwitzer & Moskowitz, 1996; Locke & Latham, 1990). Such conscious desires naturally dictate people's (conscious) thoughts, emotions, and behaviors. This past research also largely focused on the various determinants and effects of specific types of goals (e.g., accuracy vs. impression formation), and different ways of approaching the same goal (attaining achievement via academic or social means). In contrast with this

work, our framework involves the consideration of goals that can become activated and operate without the person's awareness or intention, either in isolation or among other goal pursuits, a move that reflects much research in social cognition over the last two decades (e.g., Bargh & Chartrand, 1999; Kruglanski et al., 2002). With the assumption that goals essentially consist of constructs in memory that operate according to basic principles of knowledge activation (e.g., Higgins, 1996) comes the potential for such constructs to be activated in memory without the person's awareness. And, just as a given thought, emotion, and action can be prompted by processing that remains implicit, so too can goal pursuit. In this way, people's choices of actions, emotions, and evaluations can be driven by goals of which they are unaware.

It is noteworthy that even though this recent framework differs in arguably substantive ways from much traditional research on goals, it nevertheless follows directly from classical research in social psychology more broadly. In particular, the view that goals can become activated and influential merely on perception of features of the environment follows from the tradition in social psychology to understand and document the power of situational forces to influence human behavior (since Asch, 1952; Cartwright, 1959; Lewin, 1935; Milgram, 1963). In this way, some of the recent work on goals provides a fuller picture of how goals might be selected merely as a function of the prompts and triggers in people's everyday surroundings.

REFERENCES

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, *78*(1), 53–63.

Aarts, H., Dijksterhuis, A., & De Vries, P. (2001). On the psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology*, *92*(4), 631.

Aarts, H., Gollwitzer, P. M., & Hassin, R. (2004). Goal contagion: Perceiving is for pursuing. *Journal of Personality and Social Psychology*, *87*, 23–37.

Aarts, H., Hassin, R. R., & Ferguson, M. J. (2005). Spontaneous goal inferences. *Journal of Experimental Social Psychology*, *41*, 129–140.

Ach, N. (1935). Analyse des willens [Analysis of the will]. In E. Abderhalden (Ed.), *Handbuch der biolagishen arbeitsmethoden* (Vol. 6, Part E). Berlin: Urban & Schwarzenberg.

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*(4), 463–496.

Albarracín, D., Johnson, B. T., & Zanna, M. P. (Eds.). (2005). *The handbook of attitudes*. Mahwah, NJ: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*(2), 186–197.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*(3), 219–224.

Arnold, M. B. (1960). *Emotion and personality*. New York: Columbia University Press.

Aronson, E. (1997). The theory of cognitive dissonance: The evolu-

tion and vicissitudes of an idea. In C. McGarty & S. A. Haslam (Eds.), *The message of social psychology: Perspectives on mind in society* (pp. 20–35). Cambridge, MA: Blackwell.

Asch, S. E. (1952). *Social psychology*. New York: Prentice Hall.

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380–393.

Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, *22*(1), 1–32.

Aspinwall, L. G., & Taylor, S. E. (1997). A stitch in time: Self-regulation and proactive coping. *Psychological Bulletin*, *121*(3), 417–436.

Atkinson, J. (1964). *An introduction to motivation*. Oxford, UK: Van Nostrand.

Atkinson, J. W. (1974). Strength and motivation and efficiency of performance. In J. W. Atkinson & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 193–218). New York: Wiley.

Atkinson, J. W., & Birch, D. (1970). *The dynamics of action*: (1970).

Atkinson, J. W., & Raynor, J. O. (1978). *Personality, motivation, and achievement*. New York: Halsted Press.

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*(3), 338–375.

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology*, *91*, 612–625.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 93–130). New York: Guilford Press.

Bargh, J. A., & Barndollar, K. (1996). Automaticity in action: The unconscious as repository of chronic goals and motives. In G. P. M. & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 457–481). New York: Guilford Press.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*(7), 462–479.

Bargh, J. A., Chaiken S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditionally automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology*, *32*, 185–210.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. In W. D. Spaulding (Ed.), *Nebraska Symposium on Motivation* (Vol. 41, pp. 71–124). Lincoln: University of Nebraska Press.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Troetschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*(6), 1014–1027.

Baron, J. (2000). *Thinking and deciding* (3rd ed.). Cambridge, UK: Cambridge University Press.

Bassili, J. N., & Brown, R. (2005). Implicit and explicit attitudes: Research, challenges and theory. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change* (pp. 543–574). Mahwah, NJ: Erlbaum.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego, CA: Academic Press.

Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Pychological Bulletin*, *115*(2), 243–267.

Baumeister, R. F., & Tice, D. M. (2001). *The social dimension of sex*. Boston: Allyn & Bacon.

Beer, J. S., & Keltner, D. (2004). What is unique about self-conscious emotions? *Psychological Inquiry*, *15*(2), 126–128.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Bogardus, E. (1931). *Fundamentals of social psychology* (2nd ed.). New York: Appleton-Century-Crofts.

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*109–131.

Brendl, C. M., & Higgins, E. T. (1996). Principles of judging valence: What makes events positive or negative. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 95–160). New York: Academic Press.

Brendl, C., Markman, A. B., & Messner, C. (2003). The devaluation effect: Activating a need devalues unrelated objects. *Journal of Consumer Research*, *29*(4), 463–473.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*(2), 123–152.

Bruner, J. S., & Postman, L. (1948). Symbolic value as an organizing factor in perception. *Journal of Social Psychology*, *27*, 203–208.

Brunstein, J. C., & Gollwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology*, *70*(2), 395–407.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*(3), 366–381.

Buss, D. M. (1988). *Personality and the evocation of anger and upset*. Unpublished manuscript, University of Texas at Austin.

Cabanac, M. (1971). Physiological role of pleasure. *Science*, *173*, 1103–1107.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, *76*(5), 839–855.

Cantor, N., & Langston, C. A. (1989). Ups and downs of life tasks in a life transition. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 127–167). Hillsdale, NJ: Erlbaum.

Cartwright, D. (1959). *Studies in social power*. Unpublished manuscript, University of Michigan, Ann Arbor.

Carver, C. S. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion*, *17*, 241–261.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer.

Carver, C. S., & Scheier, M. F. (1990). Principles of self-regulation: Action and emotion. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Volume 2. Foundations of social behavior* (pp. 3–52). New York: Guilford Press.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Cesario, J., Plaks, J. E., & Higgins, E. (2006). Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology*, *90*(6), 893–910.

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology*, *71*(3), 464–478.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, *69*(2), 318–328.

Clore, G. L. (1994). Why emotions are never unconscious. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 285–290). New York: Oxford University Press.

Cochran, W., & Tessser, A. (1996). The "what the hell" effect: some

effects of goal proximity and goal framing on performance. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: interactions among goals, affect, and self-regulation* (pp. 99–120). Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic priming. *Psychological Review*, *82*, 407–428.

Corwin, G. (1921). Minor studies from the psychological laboratory of Cornell University. *American Journal of Psychology*, *32*, 563–570.

Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology*, *89*(2), 129–142.

Deci, E. L., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18.

Dhar, R. (1996). The effect of decision strategy on deciding to defer choice. *Journal of Behavioral Decision Making*, *9*(4), 265–281.

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research*, *24*(2), 215–231.

Dhar, R., & Simonson, I. (1999). Making complementary choices in consumption episodes: Highlighting versus balancing. *Journal of Marketing Research*, *36*(1), 29–44.

Dhar, R., & Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. *Journal of Marketing Research*, *37*(1), 60–71.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, *74*, 865–877.

Doob, L. W. (1947). The behavior of attitudes. *Psychological Review*, *54*, 135–156.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.

Drolet, A. (2002). Inherent rule variability in consumer choice: Changing rules for change's sake. *Journal of Consumer Research*, *29*(3), 293–305.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*, 1441–1455.

Emmons, R. A. (1992). Abstract versus concrete goals: Personal striving level, physical illness, and psychological well-being. *Journal of Personality and Social Psychology*, *62*(2), 292–300.

Emmons, R. A., & King, L. A. (1988). Conflict among personal strivings: Immediate and long-term implications for psychological and physical well-being. *Journal of Personality and Social Psychology*, *54*(6), 1040–1048.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 23). New York: Academic Press.

Feather, N. T. (1990). Bridging the gap between values and actions: Recent applications of the expectancy-value model. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (Vol. 2, pp. 151–192). New York: Guilford Press.

Ferguson, M. J. (2006). *On evaluative readiness: The automatic attitudes of effective self-regulation*. Unpublished manuscript, Cornell University.

Ferguson, M. J. (in press). On the automatic evaluation of end-states. *Journal of Personality and Social Psychology*.

Ferguson, M. J., & Bargh, J. A. (2004a). How social perception automatically influences behavior. *Trends in Cognitive Sciences, 8*, 33–39.

Ferguson, M. J., & Bargh, J. A. (2004b). Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology*, *87*(5), 557–572.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The lib-

erating effect of perceived goal progress on choice. *Journal of Consumer Research*, *32*, 370–377.

Fishbach, A., & Dhar, R. (in press). Dynamics of goal-based choice. In C. P. Haugtvedt, P. M. Herr, & F. R. Kardes (Eds.), *Handbook of consumer psychology*. Mahwah, NJ: Erlbaum.

Fishbach, A., Dhar, R., & Zhang, Y. (2006). Subgoals as substitutes or complements: The role of goal accessibility. *Journal of Personality and Social Psychology*, *91*(2), 232–242.

Fishbach, A., Friedman, R. S., & Kruglanski, A. W. (2003). Leading us not unto temptation: Momentary allurements elicit overriding goal activation. *Journal of Personality and Social Psychology*, *84*(2), 296–309.

Fishbach, A., & Labroo, A., A. (in press). Be better or be merry: How mood affects self-control. *Journal of Personality and Social Psychology*.

Fishbach, A., Ratner, R. K., & Zhang, Y. (2006). *Construing consumption in terms of stable preference versus satiated desire*. Unpublished manuscript, University of Chicago.

Fishbach, A., & Shah, J. Y. (2006). Self-control in action: Implicit dispositions toward goals and away from temptations. *Journal of Personality and Social Psychology*, *90*(5), 820–832.

Fishbach, A., Shah, J. Y., & Kruglanski, A. W. (2004). Emotional transfer in goal systems. *Journal of Experimental Social Psychology*, *40*, 723–738.

Fishbach, A., & Trope, Y. (2005). The substitutability of external control and self-control. *Journal of Experimental Social Psychology*, *41*(3), 256–270.

Fishbach, A., & Zhang, Y. (2006). *Together and apart: What makes choice alternatives compete versus complement*. Unpublished manuscript, University of Chicago.

Fishbach, A. Zhang, Y., & Trope, Y. (2006). *Implicit counteractive evaluations*. Unpublished manuscript, University of Chicago.

Fiske, S. T. (1989). Examining the role of intent: Toward understanding its role in stereotyping and prejudice. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 253–283). New York: Guilford Press.

Fitzsimons, G. M., & Bargh, J. A. (2003). Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology*, *84*(1), 148–163.

Förster, J., Liberman, N., & Higgins, E. (2005). Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology*, *41*(3), 220–239.

Freedman, J. L., & Fraser, C. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*, 195–202.

Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Frijda, N. H. (1996). Passions: Emotion and socially consequential behavior. In R. Kavanaugh, B. Zimmerbag, & S. Fein (Eds.), *Emotion: Interdisciplinary perspectives* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, *90*(3), 351–367.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*, 535–542.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*(3), 617–638.

Giner-Sorolla, R. (2001). Guilty pleasures and grim necessities: Affective attitudes in dilemmas of self-control. *Journal of Personality and Social Psychology*, *80*(2), 206–221.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–55.

Gollwitzer, P. M. (1990). Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 53–92). New York: Guilford Press.

Gollwitzer P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 287–312). New York: Guilford Press.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503.

Gollwitzer, P. M., Bayer, U. C., & McCulloch, K. C. (2005). The Control of the Unwanted. In R. R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 485–515). New York: Oxford University Press.

Gollwitzer, P. M., & Moskowitz, G. B. (1996). Goal effects on thought and behavior. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 361–399). New York: Guilford Press.

Goschke, T., & Kuhl, J. (1993). Representation of intentions: Persisting activation in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1211–1226.

Green, L., & Rachlin, H. (1996). Commitment using punishment. *Journal of the Experimental Analysis of Behavior*, *65*(3), 593–601.

Greenwald, A. G. (1992). New Look 3: Reclaiming unconscious cognition. *American Psychologist*, *47*, 766–779.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science*, *273*, 1699–1702.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, *103*, 336–355.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Herek, G. M. (2000). The social construction of attitudes: Functional consensus and divergence in the U.S. public's reactions to AIDS. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 325–364). Mahwah, NJ: Erlbaum.

Herek, G. M. (2002). Gender gaps in public opinion about lesbians and gay men. *Public Opinion Quarterly*, *66*, 40–66.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T. (1999). Self-discrepancy: A theory relating self and affect. In R. F. Baumeister (Ed.), *The self in social psychology: Key readings in social psychology* (pp. 150–181). Philadelphia: Psychology Press/Taylor & Francis.

Higgins, E. T., & Kruglanski, A. W. (2000). *Motivational science: Social and personality perspectives*. Philadelphia: Psychology Press.

Higgins, E., Shah, J., & Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, *72*(3), 515–525.

Higgins, T. E. (1997). Beyond pleasure and pain. *American Psychologist*, *52*(12), 1280–1300.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385.

Hommel, B., Muesseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–937.

Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review*, *38*, 487–506.

Isen, A. M., & Simmonds, S. F. (1978). The effect of feeling good on a helping task that is incompatible with good mood. *Social Psychology*, *41*(4), 346–349.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is de-

motivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79*(6), 995–1006.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13,* 314–336.

James, W. (1890). *The principles of psychology* (Vol. 2). New York: Holt.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Malden, MA: Blackwell.

Jones, E. E., & Thibaut, J. W. (1958). Interaction goals as bases of inference in interpersonal perception. In L. Petrullo & R. Tagiuri (Eds.), *Person perception and interpersonal behavior* (pp. 151–178). Stanford, CA: Stanford University Press.

Kawada, C., Oettingen, G., Gollwitzer, P. M., & Bargh, J. A. (2004). The projection of implicit and explicit goals. *Journal of Personality and Social Psychology, 86,* 545–559.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Kierkegaard, S. (1938). *Purity of heart is to will one thing*. New York: Harper & Row.

Kivetz, R., & Simonson, I. (2002). Self-control for the righteous: Toward a theory of precommitment to indulgence. *Journal of Consumer Research, 29*(2), 199–217.

Klinger, E. (1996). Emotional influences on cognitive processing, with implications for theories of both. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 197–218). New York: Guilford Press.

Koo, M., & Fishbach, A. (2006). *When do accomplished versus unaccomplished actions increase motivation?: A matter of goal commitment*. Unpublished manuscript, University of Chicago.

Köpetz, C., Fishbach, A., & Kruglanski, A. W. (2006). *Having one's cake and eating it too: The quest for multifinal means in goal pursuit*. Unpublished manuscript, University of Maryland.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis of stimulus–response compatibility—A model and taxonomy. *Psychological Review, 97,* 253–270.

Kross, E., Ayduk, O., & Mischel, W. (2005). When asking "why" does not hurt: Distinguishing rumination from reflective processing of negative emotions. *Psychological Science, 16*(9), 709–715.

Kruglanski, A. W. (1996). Goals as knowledge structures. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 599–618). New York: Guilford Press.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 331–378). New York: Academic Press.

Kuhl, J. (1986). Motivation and information processing: A new look at decision making, dynamic change, and action control. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Vol. 1. Foundations of social behavior* (pp. 404–434). New York: Guilford Press.

Kuhl, J. (1987). Action control: The maintenance of motivational states. In F. Halisch & J. Kuhl (Eds.), *Motivation, intention, and volition* (pp. 279–291). Berlin: Springer.

Kuhl, J., & Beckmann, J. (1985). *Action control from cognition to behavior*. New York: Springer-Verlag.

Lang, P. (1984). Cognition in emotion: Concept and action. In C. Izard, J. Kagan, & R. Zajonc (Eds.). *Emotion, cognition, and behavior* (pp. 196–226). New York: Cambridge University Press.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Leith, K. P., & Baumeister, R. F. (1996). Why do bad moods increase self-defeating behavior? Emotion, risk tasking, and self-regulation. *Journal of Personality and Social Psychology, 71*(6), 1250–1267.

Lewin, K. (1926). Vorsatz, Wille, und Bedürfnis [Intention, will, and need]. *Psychologische Forschung, 7,* 330–385.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Liberman, N., & Förster, J. (2000). Expression after suppression: A motivational explanation of postsuppressional rebound. *Journal of Personality and Social Psychology, 79*(2), 190–203.

Liberman, N., Förster J., & Higgins, E. T. (in press). Completed vs. interrupted priming: Reduced accessibility from post-fulfillment inhibition. *Journal of Experimental Social Psychology*.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Upper Saddle River, NJ: Prentice-Hall.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*(3), 272–292.

Markman, A. B., & Brendl, C. (2000). The influence of goals on value and choice. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 39, pp. 97–128). San Diego, CA: Academic Press.

Markus, H., & Ruvolo, A. (1989). Possible selves: Personalized representations of goals. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 211–241). Hillsdale, NJ: Erlbaum.

Marsh, R. L., Hicks, J. L., & Bryan, E. S. (1999). The activation of unrelated and canceled intentions. *Memory and Cognition, 27,* 320–327.

McClelland, D. C., & Atkinson, J. W. (1948). The projective expression of needs, I: The effect of different intensities of the hunger drive on perception. *Journal of Psychology, 25,* 205–232.

McClelland, D. C., Atkinson, J. W., Clark, R. W., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96,* 690–702.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106*(1), 3–19.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371–378.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Henry Holt.

Mischel, H. N., & Mischel, W. (1983). The development of children's knowledge of self-control strategies. *Child Development, 54*(3), 603–619.

Mischel, T. (1964). Personal constructs, rules, and the logic of clinical activity. *Psychological Review, 71*(3), 180–192.

Mischel, W. (1984). Convergences and challenges in the search for consistency. *American Psychologist, 39*(4), 351–364.

Mischel, W. (1996). From good intentions to willpower. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 197–218). New York: Guilford Press.

Mischel, W., & Ayduk, O. (2004). Willpower in a cognitive-affective processing system: The dynamics of delay of gratification. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 99–129). New York: Guilford Press.

Mischel, W., Cantor, N., & Feldman, S. (1996). Principles of self-regulation: The nature of willpower and self-control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 329–360). New York: Guilford Press.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244,* 933–938.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*(3), 455–469.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*(1), 33–43.

Moore, B. S., Clyburn, A., & Underwood, B. (1976). The role of affect in delay of gratification. *Child Development*, *47*(1), 273–276.

Moors, A., & De Houwer, J. (2001). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition and Emotion*, *15*(6), 749–766.

Moors, A., De Houwer, J., & Eelen, P. (2004). Automatic stimulus–goal comparisons: Support from motivational affective priming studies. *Cognition and Emotion*, *18*(1), 29–54.

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology*, *38*(4), 397–404.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*(1), 167–184.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.

Mukhopadhyay, A., & Johar, G. V. (2005). Where there is a will, is there a way? Effects of lay theories of self-control on setting and keeping resolutions. *Journal of Consumer Research*, *31*(4), 779–786.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247–259.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*(3), 774–789.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*, 1–15.

Norman, D. A., & Shallice, T. (1986). Attention to action: willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Vol. 4. Advances in research and theory* (pp. 1–18). New York: Plenum Press.

Oettingen, G., & Mayer, D. (2002). The motivating function of thinking about the future: Expectations versus fantasies. *Journal of Personality and Social Psychology*, *83*(5), 1198–1212.

Osgood, C. E. (1953). *Method and theory in experimental psychology*. New York: Oxford University Press.

Peak, H. (1955). Attitude and motivation. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 3, pp. 149–188). Lincoln: University of Nebraska Press.

Pervin, L. A. (1989). Goal concepts: Themes, issues, and questions. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 473–479). Hillsdale, NJ: Erlbaum.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Powers, W. T. (1973). *Behavior: The control of perception*. Oxford, UK: Aldine.

Prelec, D., & Bodner, R. (2003). Self-signaling and self-control. In G. Loewenstein & D. Read (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 277–298). New York: Russell Sage.

Rachlin, H. (1997). Self and self-control. *Annals of the New York Academy of Sciences*, *818*, 85–97.

Rachlin, H. (2000). *The science of self-control*. Cambridge, MA: Harvard University Press.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, *17*, 15–22.

Raghunathan, R., & Trope, Y. (2002). Walking the tightrope between feeling good and being accurate: Mood as a resource in processing persuasive messages. *Journal of Personality and Social Psychology*, *83*(3), 510–525.

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, *26*(1), 1–15.

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34–49.

Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *Journal of Abnormal Social Psychology*, *53*, 367–372.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Schelling, T. C. (1978). Egonomics, or the art of self-management. *American Economic Review*, *68*, 290.

Schelling, T. C. (1984). Self-command in practice, in policy, and in a theory of rational choice. *American Economic Review*, *74*(2), 1.

Schwarz, J., & Pollack, P. R. (1977). Affect and delay of gratification. *Journal of Research in Personality*, *11*(2), 147–164.

Schwarz, N., & Clore, G. L. (2003). Mood as information: 20 years later. *Psychological Inquiry*, *14*(3–4), 296–303.

Shah, J. (2003). The motivational looking glass: How significant others implicitly affect goal appraisals. *Journal of Personality and Social Psychology*, *85*(3), 424–439.

Shah, J. Y. (2005). The automatic pursuit and management of goals. *Current Directions in Psychological Science*, *14*(1), 10–13.

Shah, J. Y., Brazy, P., & Jungbluth, N. (2005). *SAVE it for later: Implicit effort regulation and the self-regulatory anticipation of volitional exertion*. Unpublished manuscript, Duke University.

Shah, J. Y., Friedman, R., & Kruglanski, A. W. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, *83*(6), 1261–1280.

Shah, J., & Higgins, E. T. (2001). Regulatory concerns and appraisal efficiency: The general impact of promotion and prevention. *Journal of Personality and Social Psychology*, *80*, 693–705.

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*(4), 368–383.

Shah, J. Y., & Kruglanski, A. W. (2003). When opportunity knocks: Bottom-up priming of goals by means and its effects on self-regulation. *Journal of Personality and Social Psychology*, *84*(6), 1109–1122.

Shallice, T. (1972). Dual functions of consciousness. *Psychological Review*, *79*, 383–393.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, *22*(1), 13–39.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. *Psychological Review*, *84*, 127–190.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174.

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, *27*(2), 150–162.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, *70*, 893–912.

Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *Journal of Experimental Psychology*, *59*, 239–245.

Soman, D., & Cheema, A. (2004). When goals are counterproductive: The effects of violation of a behavioral goal on subsequent performance. *Journal of Consumer Research*, *31*(1), 52–62.

Sorrentino, R. M., & Higgins, E. T. (Eds.). (1986). *Handbook of motivation and cognition: Foundations of social behavior*. New York: Guilford Press.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some de-

terminants and implications. *Journal of Personality and Social Psychology*, *37*(10), 1660–1672.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*, 42–61.

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1–49). New York: Wiley.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*, 166–180.

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*(6), 1256–1269.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193–210.

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist*, *53*(4), 429–439.

Tesser, A., Martin, L. L., & Cornell, D. P. (1996). On the substitutability of self-protective mechanisms. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 48–68). New York: Guilford Press.

Thaler, R. H. (1991). *Quasi rational economics*. New York: Russell Sage.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy*, *89*, 392–406.

Thurstone, L. L. (1931). The measurement of social attitudes. *Journal of Abnormal and Social Psychology*, *26*, 249–269.

Tice, D. M., Bratslavsky, E., & Baumeister, R. F. (2001). Emotional distress regulation takes precedence over impulse control: If you feel bad, do it! *Journal of Personality and Social Psychology*, *80*(1), 53–67.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts.

Tracy, J., & Robins, R. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, *15*(2), 103–125.

Trope, Y., & Fishbach, A. (2000). Counteractive self-control in overcoming temptation. *Journal of Personality and Social Psychology*, *79*(4), 493–506.

Trope, Y., & Fishbach, A. (2005). Going beyond the motivation given: Self-control and situational control over behavior. In R. R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 537–565). New York: Oxford University Press.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*(3), 403–421.

Trope, Y., & Neter, E. (1994). Reconciling competing motives in self-evaluation: The role of self-control in feedback seeking. *Journal of Personality and Social Psychology*, *66*(4), 646–657.

Trope, Y., & Pomerantz, E. M. (1998). Resolving conflicts among self-evaluative motives: Positive experiences as a resource for overcoming defensiveness. *Motivation and Emotion*, *22*(1), 53–72.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Tversky, A., & Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, *3*(6), 358–361.

Vallacher, R. R., & Wegner, D. M. (1985). *A theory of action identification*. Hillsdale, NJ: Erlbaum.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*(1), 3–15.

Vallerand, R. J., & Ratelle, C. (2002). Intrinsic and extrinsic motivation: A hierarchical model. In E. L. Deci & R. M. Ryan (Eds.), *The handbook of self-determination research* (pp. 37–63). Rochester, NY: University of Rochester Press.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*(3), 249–254.

Wegener, D. T., & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality and Social Psychology*, *66*(6), 1034–1048.

Wegener, D. T., & Petty, R. E. (2001). Understanding effects of mood through the elaboration likelihood and flexible correction models. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 177–210). Mahwah, NJ: Erlbaum.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, *71*(1), 3–25.

Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science*, *246*, 1232–1233.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Erlbaum.

Wilensky, R. (1983). *Planning and understanding: A computational approach to human reasoning*. Reading, MA: Addison-Wesley.

Wilson, T. D., Lindsey S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*(5), 821–836.

Wright, R. A. (1996). Brehm's theory of motivation as a model of effort and cardiovascular response. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 424–453). New York: Guilford Press.

Young, P. T. (1959). The role of affective processes in learning and motivation. *Psychological Review*, *66*, 104–125

Young, P. T. (1961). *Motivation and emotion*. New York: Wiley.

Zauberman, G., & Lynch, J. G., Jr. (2005). Resource slack and propensity to discount delayed investments of time versus money. *Journal of Experimental Psychology: General*, *134*(1), 23–37.

Zhang, Y., Fishbach, A., & Dhar, R. (in press). When thinking beats doing: The role of optimistic expectations in goal-based choice. *Journal of Consumer Research*.

Zhang, Y., Fishbach, A., & Kruglanski, A. W. (in press). The dilution model: How additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*.

CHAPTER 22

Self-Regulation
and the Executive Function

The Self as Controlling Agent

Roy F. Baumeister
Brandon J. Schmeichel
Kathleen D. Vohs

A man squanders his money on gambling. A woman beats her child. A drunk driver causes a crash that destroys three cars and injures several people. A student postpones studying until the night before the test and gets a bad grade. A young couple engages in unprotected sex and creates an unwanted pregnancy. A delinquent shoots an acquaintance during an argument. A dieter eats seven donuts and a pint of ice cream at one sitting. An athlete trains off and on for a year without any improvement in performance. A girl breaks a promise and betrays a friend's confidence. An old man again neglects to take his daily dose of insulin and goes into diabetic shock.

What these disparate events have in common is failure of self-regulation. When self-regulation works well, it enables people to alter their behavior to conform to rules, plans, promises, ideals, and other standards. When it fails, any one of a broad range of human problems and misfortunes can arise. Self-regulation is thus a key to success in human life and, when it falls short, a contributing cause that helps explain many forms of human suffering.

In this chapter, we provide an overview of the psychology of self-regulation. We review what it is, its importance, how it functions, how it fits into the broader context of human psychological functioning, and what some of its principal applications are.

THE IMPORTANCE OF SELF-REGULATION

To appreciate the importance of self-regulation, it is necessary to consider both practical and theoretical implica-

tions. The practical ones were anticipated in the opening paragraph of this chapter, but they can be stated more systematically as follows: Most of the social and personal problems that afflict people in modern Western society have some element of self-regulatory failure at their root. This is not to say that better self-regulation would alone solve all society's problems—but it would probably go a long way toward that end.

Perhaps the problems that most obviously revolve around self-control failure are those of impulse control. Drug and alcohol addiction has multiple determinants, but to the extent that people can regulate their consumption of these problematic substances, they will be less vulnerable to addiction. Many of the problems associated with sexual behavior are fully preventable, if only people would control themselves sufficiently to minimize risks. These include the paradoxical epidemic of unwanted and out-of-wedlock pregnancy (paradoxical because those problems have proliferated in recent decades despite the concomitant, historically unprecedented availability of highly effective contraceptive methods), as well as epidemics of sexually transmitted diseases. Eating disorders likewise have remained problematic for young women, and obesity has been officially declared a national and even international health problem (the so-called globesity epidemic), as people find themselves unable to regulate the most basic human function of eating.

Self-regulation failure is less obvious but perhaps no less central to many other problems. A landmark work of criminology concluded that deficient self-control is the single most important key to understanding criminality (Gottfredson & Hirschi, 1990). Subsequent work testing

this hypothesis has confirmed the central importance of low self-control, even if it is not the only key predictor. Apparently, people become criminal because they are poor at regulating their antisocial impulses and hence violate many of society's formal (and informal) rules. This pattern helps explain many hitherto baffling aspects of criminality, including the so-called versatility of criminals (i.e., most criminals are arrested repeatedly but for different crimes).

Money problems are also often linked to self-regulatory problems. Americans often fail to save money, and the low rate of savings is a problem both for individuals, who find themselves unable to cope with unexpected financial needs, and for the society and economy as a whole, for which the low reservoir of savings creates a lack of available capital. Many people earn good incomes but suffer from heavy debt loads, often attributable to unregulated use of credit cards.

Underachievement in school and work likewise has a dimension of poor self-control. Procrastination is now generally regarded as both a cause of poor performance and a reflection of poor self-regulation, and so it is one dimension of underachievement. Poor self-regulation can contribute in other ways to underachievement, such as by making people less willing to persist in the face of failure, less able to choose effective performance settings, less able to set and reach goals, and less able to sustain effort over a period of time.

Another way of appreciating the benefits of self-regulation is to compare the lives of people with good versus bad self-control. Although such comparisons are inherently correlational, and as a result it is in principle possible that self-control is the result rather than the cause of such differences, most theorists assume that personality traits precede behaviors and are therefore more likely the cause than the consequence. A recent set of studies by Tangney, Baumeister, and Boone (2004) included a trait measure of self-control and then examined multiple indices of effective functioning. People with high scores on self-control were better off than those with low self-control on virtually all of them. They had better grades in school. They had better relationships with family and friends: less conflict and more cohesion. They were better able to understand others and scored higher on empathy. They showed better psychological adjustment, including fewer psychological problems, fewer signs of serious psychopathology, and higher self-esteem. Not surprisingly, they reported fewer impulse control problems, such as overeating and problem drinking. They had healthier emotional lives, such as being better at managing their anger and being more prone to guilt than shame. They had less juvenile delinquency.

Other work using the same scale has confirmed the benefits. Supervisors who score higher in self-control are rated more favorably (e.g., as fairer) by their subordinates (Cox, 2000). People with high self-control make better relationship partners, especially because they are better able to adapt to partners (Finkel & Campbell, 2001; Tangney et al., 2004; Vohs & Baumeister, 2004).

Probably the most dramatic and conclusive evidence of the long-term benefits of self-regulation comes from the research by Walter Mischel and his colleagues. Mischel was a pioneer of self-regulation research because of his studies on delay of gratification, beginning in the 1960s. Self-regulation is required to override the impulse to seek immediate gratification in order to obtain greater but delayed rewards. His research group then followed up the early studies, which were typically done with young children, to see how they fared on into adulthood. Four- and 5-year-olds who were able to resist the temptation of one cookie in order to eat two cookies a short while later grew up to earn better marks on the SAT, to be rated by others as rational and socially competent, and to cope with frustration and stress better than those kids who were relatively unable to resist the tempting cookie at a young age. Thus, effective self-regulation can be recognized as an important key to success in life (see Mischel & Ayduk, 2004).

If practical benefits are not enough, however, self-regulation can also be recognized as important based on its theoretical implications. It is an important key for understanding what the human self is and how it operates. An analysis of psychological and behavioral processes is inadequate without it. Perhaps the emergence of self-regulation is one of the central steps in human evolution and a crucial aspect of human nature—one of the traits that most distinguishes the human psyche from the majority of other life forms on this planet. These implications cannot be easily summarized, however, and certainly not until the theoretical context and inner processes of self-regulation have been more thoroughly elucidated. In the coming sections, we shall attempt to do that. First, however, some definitions are required.

DEFINITIONS

Self-regulation refers to the self altering its own responses or inner states. Typically this takes the form of overriding one response or behavior and replacing it with a less common but more desired response. For example, when a dedicated smoker has an urge to smoke but does not then light up a cigarette, he self-regulates his own impulses. Self-regulation also includes the ability to delay gratification, such as when a child overrides the desire to eat the cookie on her plate and waits instead for the two in the oven.

Self-regulation is one of the self's major executive functions. The executive function of the self refers to its active, intentional aspects (see Baumeister, 1998; Gazzaniga, Ivry, & Mangun, 1998) and may be thought of as that part of the self that is ultimately responsible for the actions of the individual. The other major executive function of the self is choice. Not only may a self initiate behavior or control it, but a self also is responsible for deliberating and making choices from among the universe of possible options. As we shall see, choice and self-regulation are intertwined, and they often work in concert to achieve novelty and diversity in human behavior.

Technically speaking, a self does not regulate itself directly, but it may control the behaviors, feelings, and thoughts that comprise it. In this sense, self-regulation re-

fers to the regulation of processes *by* the self. Regulation *of* the self also falls under the rubric of self-regulation, but note that this may mean the regulating is done by something (or someone) else. For example, when otherwise quite different people go to the movie theater, they tend to behave in similar ways. They sit quietly, they occasionally whisper, and they pay attention to the action on the screen. Most of this behavior occurs without much in the way of active self-regulation, although to a naive observer it may appear that the moviegoers are inhibiting their normal behavior. Instead, it is likely that the context—the movie theater, the presence of other moviegoers, the start of the movie—triggers behavior directed toward watching the movie (e.g., Schank & Abelson, 1977). Thus, the environment surrounding the self is also a powerful shaper of behavior, one that occasionally reduces the necessity of active regulation *by* the self. Thus, although self-regulation has typically implied regulation of behavior by the self in pursuit of a conscious intention or purpose, some forms of self-regulation occur without conscious awareness or active intervention by the self.

Finally, our view of self-regulation is consonant with the notion of secondary control derived from a dual-process view of control (see Rothbaum, Weisz, & Snyder, 1982). According to this view, people strive to achieve a better "fit" with their environment using either primary or secondary control strategies. Primary control involves attempts to change the world to accommodate the self, such as by donating to political candidates in order to influence policy decisions in one's favor. Secondary control strategies refer to attempts to change the self in order to fit the world, such as by regulating one's own actions not to violate current policy or law. Given the difficulties inherent in changing the world to fit one's self, secondary control probably represents the more common and more consistently successful strategy of achieving harmony between self and world.

THE BROADER CONTEXT

We said earlier that the theoretical importance of self-regulation can only be appreciated within a broader perspective of relevant contexts and concepts. In this section, we seek to describe the place and importance of self-regulation amid human psychological functioning.

The Self

Self-regulation is one important function of the human self and perhaps a significant dimension of its *raison d'etre*. In this, it is not simply one of many functions, but one of a select few that help define the self. Higgins (1996) spoke of the "sovereignty of self-regulation," referring to its preeminent importance as compared with many of the other everyday activities of the self. Self theory is incomplete without an account of self-regulation.

The activities and functions of the self, as well as the accumulated knowledge and understanding arising from research on the self, can be broadly grouped according to three main dimensions (Baumeister, 1998). These are presumably based on three basic phenomena that give rise to selfhood. The first is reflexive awareness: Consciousness can be directed toward its source, so that just as people become aware of and learn about the world, they can also become aware of and learn about themselves. The eventual upshot is a body of knowledge and belief about the self, often called the self-concept. Without this, a self would be inconceivable.

Second, the self is used to relate to others. People do not in fact develop elaborate self-concepts simply by contemplating themselves or reflecting on what they have done. Instead, they come to know themselves by interacting with others. Moreover, interpersonal relatedness is not just a root of self-knowledge but an important goal of most human functioning. Human beings essentially survive and reproduce by means of their interpersonal connections. The "need to belong" is one of the most powerful and pervasive human motivations (Baumeister & Leary, 1995), probably because evolution has designed us to achieve our biological successes through membership in groups and relationships. Throughout human evolutionary history, lone wolves have been few and far between, and they generally were less likely to pass along their genes than their more gregarious peers. Thus, the self is also a dynamic tool for connecting with others.

The third aspect of the self may be called its executive function, though it is also sometimes called the agent or agentic aspect. The first aspect of self was a knower and a known, the second a belonger or member, but this third aspect is a doer. By means of its executive function, the self exerts control over its environment (including the social environment of other people), makes decisions and choices, and also regulates itself.

Self-regulation should thus be understood in connection with the self's executive function, though it also has some relevance to self-knowledge and to interpersonal belonging. The executive function essentially does two things: It controls the self and controls the environment. Self-regulation is loosely related to decision making and choosing. We shall review research showing that self-control is directly affected by making decisions, even if the decision making is on something that has no apparent relation to the focus of self-control. Conversely, exercises in self-regulation have effects on decision making. To foreshadow, we find that making choices and exerting self-control draw on a common, limited resource, and so doing either one of them temporarily reduces one's effectiveness at the other. The connection between the two may shed light on one of the most enduring questions about human nature, namely, free will. We now turn to that.

Free Will

The magazine *The Economist* is fond of quoting Ronald Reagan's surprisingly apt characterization of an economist as someone who sees something that works in practice and wonders whether it will work in theory. In our view, this captures the approach toward choice and free will in psychology. All around us, every day, we see peo-

ple facing choices in which multiple options are really viable and possible, and they exercise some sort of strength or power to make themselves select among them. Yet, in order to be good scientists, many psychologists think they must believe that every event is caused and that the apparent exercise of choice cannot be real. And so psychologists reject the evidence of our senses and our personal experience in order to insist that people are not really choosing. The outcome of each decision must have been the only outcome that was ever really possible.

Setting metaphysics aside, let us approach the question from an evolutionary perspective (Baumeister, 2005; Dennett, 2003). If free will exists in any sense, it is almost certainly the result of evolution, and it may therefore be more advanced in human beings than in other species. What sense of free will would produce gains in terms of survival and reproduction? We (along with Baumeister, 2005, Searle, 2001 and, in some respects, Dennett, 2003) can suggest two.

The first of these is rational choice. The evolution of cognition is intricately linked to the evolution of choosing, in that organisms became more capable of selecting among behavioral options and modifying their behavior based on appraisal of their environment (Tomasello & Call, 1997). An animal that could alter its behavior to find more food or avoid newly arising dangers would survive and reproduce better than an animal that could not.

Most social sciences currently have a significant contingent of researchers whose research is based on a rational choice model. That is, they assume that people appraise their options and choose on the basis of what will further their self-interest in the long or short term. Rational analysis, which requires logical assessment (such as cost-benefit analyses) of possible outcomes, is assumed to underlie most of the decisions people make about whom to vote for (in political science) or how they invest their money (in economics). Rational analysis is a distinctively human process: As far as research has shown, no other animals engage in rational analysis, though they can make somewhat sophisticated assessments of immediate situational choices (Tomasello & Call, 1997).

Rationality, however, presupposes free will, at least in some sense. As Searle (2001) has pointed out, rational analysis is useless without free will. That is, there is no point in being able to use logic to figure out the best thing to do—if you cannot then actually do it. At best, the human capacity for logical thought would enable people to think about why what they are doing is foolish or self-defeating. If evolution created free will, it was most likely for the sake of being able to do what logic chose as the most profitable course of action.

Self-regulation is the second form of free will, if rationality is the first (and we concede that the two may be intertwined). The capacity to alter one's behavior to maximize situational payoffs, achieve long-term gains, and conform to meaningful (even abstract) standards is also highly adaptive. From an evolutionary or biological standpoint, the capacity to override an initial response and substitute another response is an immense step forward and can be powerfully adaptive. This brings up perhaps the broadest context of all.

Cultural Animals

One of us has recently argued that an adequate explanation of human psychological functioning requires a rethinking of the nature–nurture debate that has defined social sciences' ultimate explanations of human nature for decades. The two opponents in the perennial debate are nature and culture. Nature, as represented by evolutionary psychology, emphasizes similarities, specifically similarities between humans and other animals. Culture, as represented by cultural psychology, focuses on differences, especially differences among cultures.

In contrast, Baumeister (2005) proposes that we also attend to evolutionary differences and cultural similarities. That is, in what sense are humans different from other animals, and in what respects are all or most cultures similar? Crucially, Baumeister proposes that these are linked—that what all cultures have in common is also what differentiates humans from other animals. By this reasoning, the key to human nature is that evolution created us to sustain culture, in the sense of an organized network of relationships that makes the totality of its members more than the sum of its parts. Culture is the central biological strategy of human beings and the basic source of the success of the human species.

For human beings to become cultural animals, humans had to evolve to have multiple capabilities. These include language, theory of mind, reasoning—and, in some sense, free will.

Self-regulation, we think, is the evolutionary root of free will. Rational choice is the main rival for that claim. As we shall show, however, self-regulation and rational decision making draw on a common resource, which suggests these did not evolve as separate mechanisms. Rather, the common resource suggests that evolution created that resource for one of them, and human beings enjoy the second as a by-product (in biological terms, a spandrel). The question is therefore whether self-regulation or rational choice was the first to appear and was therefore the driving force.

We think self-regulation was more likely the first to appear and therefore deserves priority in the evolutionary analysis. We freely admit that this is mere educated guesswork, and we are willing to revise our assessment if contrary evidence (i.e., that rational choice preceded self-regulation) emerges. Let us, however, present the basis for our assumption.

Baumeister (2005) distinguishes social animals from cultural animals. It has become something of a truism in social psychology that human beings are social animals (Aronson, 1995). They are. But in no sense are they the only social animals—wolves, zebras, even ants are social animals. Humans may not even be the most social animals. We are, however, the most, and arguably the only, cultural animals. The evidence of culture in other species is limited, and in no sense is any other species as fundamentally cultural as we are. Frans de Waal (2001a, 2001b), one of the most passionate and persuasive advocates of culture in other species, readily concedes that no nonhuman culture remotely approaches the extent of human culture. Although other animals do qualify as cul-

tural in multiple respects, these reflect isolated adaptations that capitalize in a very limited manner on the powerful biological benefits that culture can offer. In contrast, humans are thoroughly cultural, to the extent that human life is almost unimaginable without culture. Put another way, other animals occasionally dabble in culture, whereas human beings rely indispensably on culture for all our survival and reproduction. Among the 6 billion humans alive today on the planet, hardly any survive and reproduce independently of their culture.

The distinction between social and cultural animals is therefore crucial. Being social involves coordinated action between conspecifics. Being cultural depends on use of meanings to organize collective action. Social hunters may swarm, working together to achieve what none could do alone, but cultural hunters employ division of labor to benefit from expertise and generate systemic benefits. Social animals copy each other, thereby benefiting from one another's adaptive actions, but cultural animals can transmit knowledge from one generation to another. Thus, a pack of wolves today, though undeniably social, lives largely the same as a pack of wolves did 10,000 years ago, with no accumulation of knowledge or progressive improvement of techniques and technology, let alone redefinition of gender roles or organizational structures. In contrast, human life has changed drastically and dramatically even just in the past century, and less than 1% of the human population lives like its ancestors of 10,000 years ago.

To our (admittedly speculative) view, self-regulation was already important for social animals, whereas rational choice is limited to cultural animals. Therefore, if one of those deserves priority in evolutionary analysis, it should be self-regulation. Self-regulation is beneficial for social life. The ability to override one response, to substitute a more adaptive alternative, would be helpful to merely social (i.e., not cultural) animals. As one example, if the alpha male dictates that certain mates or certain foods should be reserved for him alone, then other males would benefit by being able to inhibit their impulses to pursue those gratifications for themselves. Pursuing them would lead to severe physical punishment and possibly expulsion from the group (if not death). Rational analysis here is irrelevant. The rules that operate in social groups of the biological relatives of humankind require self-regulation but not rationality, because they depend on the immediate stimulus environment. If the alpha male is absent, his rules can be flouted: One can eat his favored food or perhaps even copulate with his favored mates. We think that nonhuman primates only follow rules when there is the prospect of immediate punishment. In contrast, human beings follow rules even in the absence of any visible enforcers. Such behavior would be unknown and incomprehensible to merely social animals, who mainly follow rules enforced by powerful others who are present and ready to enforce them immediately.

In contrast, rationality is reserved to cultural animals, who can use meaning and language and abstract reasoning to dictate the optimal course of action. Social animals without language cannot exploit the power of reasoning,

for the most part, because logical reasoning operates within the rules of meaning, which require language to understand and process. To be sure, logical reasoning may in some respects be even more powerfully adaptive than self-regulation, because choices can in principle be made on the most optimal and hence adaptive basis. Still, insofar as self-regulation arose earlier and is more basic than rationality, rational choice may have been a side effect (spandrel). The resource needed for both self-regulation and rationality was selected by evolution first to serve the need for self-regulation, and then it was applied to enable full rational choice.

If we abandon the absurd requirement proposed by some opponents of free will that free will should be for the sake of purely random action, and if we assume instead that free will evolved to promote adaptive actions and choices, then we can discern the themes that are in common between self-regulation and adaptation. Self-regulation is vital for social animals because it enables them to match their behavior to externally dictated standards, such as rules imposed by the alpha male. Rational choice entails that individuals can work out for themselves (by logical analysis) standards and rules, and so rational behavior enables people to alter their behavior to conform to standards that they themselves have constructed. It is thus a more advanced stage of free will, in the sense that conforming to one's own standards entails greater autonomy than conforming to someone else's rules. Psychologically, the same mechanism may be involved in self-regulation and rational choice, even if rational choice represents a philosophically more advanced purpose. But both are highly adaptive.

BASIC THEORETICAL ISSUES

Having explicated the theoretical context of self-regulation, we turn now to consider how it operates. We shall first survey several central or controversial theoretical issues and assumptions surrounding self-regulation. Then we turn to consider the three essential components of self-regulation, namely, commitment to standards, monitoring of relevant behavior, and the capacity for overriding responses and altering behavior.

Irresistible Impulses or Acquiescence?

In everyday life, people seem to have a ready explanation for failures at self-control: "I couldn't resist." The implication is that certain impulses are irresistible, and so they overwhelm the powers of the self. This view depicts self-control as a struggle between the strength of the impulse and the strength of the self, and whether the person resists temptation depends on the strength of the impulse. Somehow, apparently, neither nature nor nurture has provided people with strong enough powers to resist many of the temptations they encounter, or so they say.

While reviewing the research literature on self-regulation, Baumeister, Heatherton, and Tice (1994) became increasingly skeptical of the doctrine of irresistible impulses. To be sure, there are some truly irresistible im-

pulses. For example, the urge to go to sleep, stop standing up, or urinate can eventually become so overwhelming that no amount of self-regulatory power can restrain it. But these may be exceptions. When a shopper returns home and explains to a disgruntled mate that the lovely but overpriced sweater had to be purchased, wreaking havoc on the family budget, because "I just couldn't resist," the mate may justifiably think this irresistibility is not on a par with those unstoppable biological urges. Likewise, when jurors hear a defendant claim that he or she committed the crime because his or her anger created an irresistible urge to kill the victim, they are probably justified in thinking that the defendant ought to have been able to resist that violent impulse.

There are empirical signs that so-called irresistible impulses may be resistible after all. Peele (1989) noted that addiction, which is commonly understood to cause irresistible cravings, is much less compelling than often surmised. For example, many American soldiers became addicted to heroin during the Vietnam War but then seemingly easily gave up heroin when they returned home. Even more surprising, others were able to use heroin occasionally after returning to the United States without resuming their addiction, contradicting the common view that a recovered addict is in constant danger of resuming full addiction if he or she gets any small amount of the addictive substance. Many heroin addicts may experience their cravings as irresistible, but this is perhaps attributable to their own chronic weakness of will rather than anything in the nature of heroin itself.

Converging evidence comes from studies of people who suffer from obsessive–compulsive disorder (OCD). The public may assume that obsessive thoughts are somehow unstoppable, but interviews with these individuals tend to yield the pattern that they attribute their problems to weakness of self and will rather than to any overwhelming power of the thoughts (Reed, 1985). Indeed, successful treatment of OCD is barely conceivable without acknowledging the person's capacity to alter his or her thoughts.

A similar observation comes from a very different source, namely, violent criminals. Douglas (1996) rejected the view that serial killers and other brutal criminals are driven by unstoppable impulses to commit their crimes. He observed that he and his colleagues had investigated hundreds of such crimes by many different individuals, yet no such crime was ever committed in the presence of a police officer. Police officers are found in many places, perhaps especially in the sorts of locations criminals pass by, and so the odds are good that sometimes police officers would be present when a violent killer gets an irresistible impulse to commit violence. The fact that no crimes take place under such circumstances suggests that these impulses are somewhat resistible after all.

Cultures can certainly help individuals perceive some impulses as irresistible, but this may be more a matter of convention than of recognizing reality. One famous example of culturally sanctioned loss of control was the pattern of "running amok," observed in the Malay of the Indian Archipelago. According to the local customs, young men who felt they had been treated unfairly or offended might lose control and go on a violent rampage, doing damage to property and even to other people. These rampages were strongly rooted in the belief that under those circumstances people could not possibly restrain themselves. One consequence was that such rampages were not punished or only lightly punished, which seems reasonable given the assumption that the individual could not have stopped himself from the violent and destructive acts. However, when the British colonized that area, they took a dim view of running amok and began punishing men who did it. The practice diminished with surprising rapidity, indicating that it had been more controllable than people thought all along.

The "gun to the head" test was proposed by Baumeister and colleagues (1994; Baumeister & Heatherton, 1996) as a way of distinguishing the truly irresistible impulses from the more resistible ones. If an impulse is truly irresistible, then one will act on it even if someone with a gun were threatening to shoot one if one acts that way. The examples we listed earlier, such as sleep, sitting or lying down, and urinating, all pass this test: Eventually the person will perform those acts even if threatened with imminent death. But buying the expensive sweater or committing the crime would probably turn out to be resistible (see Pervin, 1996).

The implication is that most undesirable thoughts and actions are probably far more resistible than people are likely to admit. To understand failures at self-regulation, therefore, we cannot simply invoke the commonsense model of powerful urges overwhelming the self. Rather, the person may acquiesce in yielding to temptation. The shopper could resist the sweater but somehow opts not to do so.

What Is Controllable?

The previous section suggests that many impulses are more controllable than some people may admit. The human capacity for controlled processing is impressive, but it is certainly limited. Hence it becomes necessary to distinguish what is controllable from what is not.

In the 1970s and 1980s, psychology was heavily influenced by the distinction between automatic and controlled processes. This simple dichotomy has, however, evaporated with the accumulation of data (e.g., Bargh, 1994). Most relevant to the present analysis is the necessity to invoke a series of processes that might normally be automatic but that could potentially be controlled. These are thus ripe for self-regulation, whereas the hard-core uncontrollable processes are not.

Self-knowledge thus becomes an important resource for effective self-regulation (Higgins, 1996). It is helpful for people to know what they can versus cannot change about themselves. The more extensive and accurate that self-knowledge is, the more people can profitably alter the controllable responses and avoid wasting their time trying to change unchangeable things. Seligman (1994), for example, has written a book attempting to dispel myths about the controllability of some responses and the uncontrollability of others.

Much of self-regulation is often subsumed under the term "impulse control," but impulse control may be a misnomer. Most impulses are automatic responses and cannot be prevented from arising. Strictly speaking, a person with so-called good impulse control does not really control the impulse itself but rather the behavior that would follow from it. Priests who live up to the Catholic Christian ideal of celibacy, for example, do not genuinely prevent themselves from having sexual desire. Rather, they experience desire, but they refrain from acting on it and seeking sexual activity (see Sipe, 1995).

Emotion is an important category of largely uncontrollable responses. That is, people cannot generally create or terminate an emotional state by act of will. Effective affect regulation is possible, but mostly by means of indirect strategies. For example, an angry person may not able to exert control over the emotion directly, but by distracting oneself, or by reframing the issue to interpret the situation in less upsetting terms, or by exercising to the point that one grows tired and the arousal dissipates, the person can possibly help the anger to dissipate. Someone with a false belief in the controllability of the emotion itself would thus be less effective at escaping the anger than someone who correctly appreciated the need to focus on controllable things (such as how one thinks, or whether to undertake vigorous exercise) and hence used those to exert indirect influence over the emotion.

Lapse-Activated Patterns

Lapse-activated responses refer to a class of behaviors that come into play after an initial (possibly quite minor) failure of self-control. Marlatt (e.g., Marlatt & Gordon, 1985) is one of the most influential researchers into lapse-activated responses. He documented an abstinence violation effect among problem drinkers. Once such drinkers believe they have had any alcohol, they may become consumed with a sense of futility and lose their confidence that they can resist temptation. (Zero-tolerance doctrines support such a response by claiming that any alcohol will cause a problem drinker to lose control utterly.) Marlatt showed, moreover, that the abstinence violation effect is psychological rather than physiological, in the sense that it depends more on the beliefs, perceptions, and assumptions of the drinker than on any irresistible, physiological consequence of consuming alcohol. In some studies, drinkers who falsely believed they had consumed alcohol were prone to go on a binge, whereas drinkers who falsely believed they had not had alcohol maintained restraint.

Similar findings have been documented in the eating realm, under the rubric of counterregulation or, more colloquially, the "what the hell" effect (Cochran & Tesser, 1996; Herman & Mack, 1975). Dieters who believe their diet is blown for the day eat more than dieters whose diets are presumably intact. Moreover, these responses depend on the perception rather than the actual caloric consumption. In one classic set of studies, dieters who ate salads maintained control over their eating subsequently, whereas those who ate ice cream abandoned restraint and overate—even if the salad contained twice as many calories as the ice cream (Knight & Boland, 1989).

Thus, again, beliefs about the self and about the controllability of responses contribute to effective (or ineffective) self-regulation. Researchers who proposed that some recovering alcoholics can learn to use alcohol in controlled amounts have been vilified, because their recommendations go against the prevailing zero-tolerance doctrines (Sobell & Sobell, 1984). But it can be counterproductive for people to believe that any lapse will inevitably lead to a full-blown binge. In reality, preventing the first sip or first bite is probably easier than stopping after a couple, but people can also learn to stop after a limited indulgence.

Beliefs are of course not the only factor relevant to lapse-activated patterns. Alcohol abuse has been implicated in nearly every form or sphere of self-control failure (see Baumeister et al., 1994, for review), from sex and violence to overeating to just drinking all the more alcohol. Apparently alcohol has special powers to undermine self-regulation. In our view, this is most likely connected with the fact that alcohol undermines self-awareness, thereby making it difficult for the person to continue keeping track of behaviors. We return to this issue below, as we explore exactly how self-regulation operates. For now, it is sufficient to observe that alcoholic indulgence facilitates loss of control over a broad range of behaviors, enabling initial lapses to snowball into serious breakdowns.

Transcendence, and Delay or Gratification

The ability to regulate or inhibit behavior is not uniquely human. Most dog owners have been able to observe that dogs can follow simple rules, at least when the owner is present to enforce them. (Our experience is that when a dog owner tries to teach a dog not to get up on the couch, it mainly learns not to get up on the couch when the owner is present; when he or she comes home from the office, there are still dog hairs on the couch.) If the dog has learned the "stay" command, it will sit still and stare fixedly at the bacon biscuit until the owner says the word that permits the dog to come forward and eat it. If the treat is tempting enough, the dog's owner can even observe the inner struggle, as the dog's legs shake with incipient motions and the dog has to struggle to remain in place.

No doubt this capacity for restraint was something that natural selection favored during human evolution, producing perhaps increased willpower among humans. However, there appears to be one crucial aspect of self-regulation in which people differ seriously, perhaps categorically, from other species. Humans can respond to circumstances beyond the immediate stimulus environment. This is crucial for our success as cultural animals.

We favor the term "transcendence" to refer to the human capacity to process and respond to things or events that lie beyond the immediate stimulus environment. Transcendence thus does not imply any kind of spiritual or metaphysical experience (e.g., transcendentalism) but simply a psychological capacity to respond to something

that is not physically present. There is little evidence that any nonhuman animal can do this.

Much of self-regulation depends on transcendence. Indeed, perhaps the most common dilemma concerns a conflict between being tempted to enjoy something in the immediate stimulus environment versus being restrained according to some abstract rule or standard, which may be linked to something in the distant past or future. The Jewish practice of keeping kosher, for example, involves refusing to eat what most animals would regard as perfectly good food, on the basis of religious principles that were laid down centuries ago.

Such self-regulation is qualitatively different from the earlier example of the dog's regulatory efforts. The dog resists the tempting food but mainly because the master is in view and presents an imminent threat of physical punishment if the dog's self-control fails. In contrast, a Jew may refuse to eat a ham sandwich even if no one else is present and no one would ever know he ate it.

In the same way, self-regulation can be guided by distal future goals. A college student who passes up a tempting beer party in order to study at the library may be guided by concerns that have little force in the present, and indeed the immediate stimuli (such as beer-guzzling roommates) may all favor joining the party. The conscientious, good student may, however, transcend the party-favoring stimulus environment in favor of doing something that will contribute to goals that may lie weeks (the final exam) or even years (graduating with honors and going on to a better career) in the future.

Transcendence is thus instrumental for delay of gratification, and the capacity for delay has contributed both collectively and individually to human success. Farming is just one of the many activities that depend on the capacity to delay gratification and that also have provided immense benefits to human beings as a species. (We also noted that the capacity for delay produces immense benefits for individuals, including in modern society.) Getting an education is a fine illustration of the importance of pursuing delayed gratifications. Attending class, going to the library, reading, studying, taking examinations, and similar activities are not intrinsically enjoyable for either human or nonhuman animals, but humans are willing to perform them over and over, in part because they confer immense advantages in the very long run. Americans with college degrees earn tens, even hundreds of thousands of dollars more than those without such degrees, but these benefits are over a lifetime, and in the short run most people could earn more money and live more comfortably by dropping out of college and taking a job.

Research by Mischel and his colleagues (Mischel, Ebbesen, & Zeiss, 1972; Mischel, Shoda, & Peake, 1988; Mischel, Shoda, & Rodriguez, 1989; for a review, see Mischel & Ayduk, 2004) has underscored the importance of transcendence for effective self-regulation in delaying gratification. In his studies, children must resist the highly salient temptation to enjoy a cookie or marshmallow in order to garner greater pleasures and rewards in the (admittedly not-so-distant) future. Observations of children in these studies show them attempting to blot out the immediate stimulus environment, such as by shutting their eyes, turning away from the sight of the tempting stimulus, or distracting themselves via singing. To the initial surprise of the research team, seeing representations of the rewards (e.g., pictures of cookies) facilitated self-regulation, in the sense that children who looked at pictures (and not the actual cookies) were better able to delay gratification. The implication is that such representations can enable transcendence by helping the child to think of the large future reward and to disregard the most appealing properties of the immediate temptation, thereby bolstering the child's ability to delay gratification.

BASIC ELEMENTS OF SELF-REGULATION

Self-regulation depends on three main components, and in this section we discuss each in turn. The first is commitment to standards. The second is monitoring of the self and its behaviors. The third is what is needed to change the self's responses. All are necessary for effective self-regulation. Hence, a breakdown or problem with any one of them can produce failure at self-regulation.

Commitment to Standards

Goal-directed behavior is impossible without a goal. In the same way, self-regulation cannot proceed without a standard, insofar as self-regulation is the effortful attempt to alter one's behavior to meet a standard. Standards are concepts of possible, often desirable states. They include ideals, expectations, goals, values, and comparison targets (such as the status quo, or what other people have done). Self-regulation is essentially a matter of changing the self, but such change would be random or pointless without some conception of how the self ideally ought to be.

There is some evidence that problems with standards can contribute to self-regulation failure. In particular, vague, ambiguous, or conflicting standards can undermine self-regulation. For example, if the two parents disagree as to how the child should behave, or even if they disagree as to the desirability of some particular kind of behavior, children are far less likely to learn to behave properly. Conflicting standards is one important source of self-regulatory breakdown (Baumeister et al., 1994)

Probably the most important work on standards comes from Higgins (1987) and his colleagues (e.g., Higgins, Roney, Crowe, & Hymes, 1994; Shah, Higgins, & Friedman, 1998). They distinguish between "ideal" and "ought" standards. Ideals form the basis of positive strivings toward the way one would like to be. Oughts are also rooted in concepts of how one would like to be, but the focus is on what to avoid rather than what to pursue. Ideals and oughts can come from within the self (e.g., if a person embraces a particular goal or value system) or from other people (e.g., parents can communicate expectations about their offspring's behavior). Higgins (1987)

further proposed that different sets of emotions are associated with the different types of regulatory standards, though of course both types of regulatory failure produce bad (negative) emotions. That is, failures to self-regulate toward ideals produces dejected, low-arousal emotions, such as sadness and disappointment. In contrast, failures to self-regulate according to oughts lead to agitated, high-arousal emotions, such as anxiety and worry.

One possible way of accounting for these differential emotional responses is that violating ought standards is more troubling than violating ideal standards, and so failures in the "ought" domain produce arousal. Arousal, as generally understood, prepares the body for action and mobilizes physical responding. It is more important or at least more pressingly urgent not to break important rules, such as by performing immoral behavior, than it is to move toward one's ideals. This may reflect the broader principle that bad is stronger than good (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001). A single violation of moral rules, for example, can directly cause ejection from social groups or create other problems for a long period of time, whereas a failure to move toward one's ideals does not necessarily have such ramifications and might simply mean that the person will have to seek another opportunity. Put another way, failing to advance toward ideals does not generally do as much harm to one's social position as does violating one's "ought" requirements.

Thus far we have emphasized having and knowing standards, but there is a motivational aspect as well. Specifically, the person must have some inner wish or proclivity to conform to the standards. If people do not care about being good, they will not be so good! Variation in self-regulatory motivation is an important but understudied aspect of self-regulation. A dieter, for example, may mostly want to keep to the diet, but under some circumstances that motivation may wane, and the person may knowingly violate the diet simply because at that moment he or she ceases to care much about losing weight.

Perhaps the most active area of research on self-regulatory motivation is Higgins and colleagues' work on regulatory focus (see Higgins & Spiegel, 2004, for a review). Building on Higgins's work on ideal and ought standards, regulatory focus refers to a person's characteristic motivational orientation. Some people are typically promotion focused, meaning that they are mainly motivated to achieve desirable outcomes using eager, approach-oriented strategies (i.e., to pursue ideals and obtain nurturance). Other people are prevention focused, meaning that they are mainly motivated to achieve desirable outcomes using vigilant, avoidance-oriented strategies (i.e., to satisfy oughts and to obtain security). A burgeoning literature attests to both the subjective and objective consequences of these two regulatory foci. For example, people who are characteristically promotion focused experience greater success when a task is framed as a pursuit of ideals. Conversely, people who are prevention focused experience greater success when a task is framed as an opportunity to do what one ought to do (Shah et al., 1998). Thus, when people experience regulatory fit—when their preferred regulatory motivation is matched by characteristics of the task at hand (Higgins, 2000)—self-regulatory outcomes are improved. To return to the dieting example, the regulatory focus view would suggest that promotion-focused individuals would best adhere to a diet if they considered it as a way to obtain an ideal body image, whereas prevention-focused individuals would best maintain motivation if their diet was framed as a way to eat in a responsible manner.

Monitoring

One hotbed of research activity in the 1970s was the study of self-awareness (from Duval & Wicklund, 1972). A landmark event in the evolution of that line of work was Carver and Scheier's (1981) book, which was received as a book on self-awareness but presented itself (starting with its title) as a book on self-regulation. At the time, the term "self-regulation" was quaint and obscure, whereas self-awareness was a familiar term. But Carver and Scheier were proposing that the main purpose of self-awareness was to facilitate self-regulation.

Linking self-awareness to self-regulation was a key, relevant point: It is very hard to change or otherwise a behavior if one is not aware of it. Monitoring one's behavior is an indispensable component to regulating it.

The writings of Carver and Scheier (1981) persuaded most social psychologists that the feedback-loop theory, originally developed by cybernetics theorists to enable missiles to hit and destroy their targets more effectively, could be profitably adapted to the psychology of self-regulation of human behavior. The core concept of the feedback loop (see Powers, 1973) involved a sequence of steps, under the acronym TOTE (for test, operate, test, exit). The test phases consist of comparing the self's current status against the relevant standards. This fact alone represents a key insight from the self-awareness research of the 1970s, which is that attending to self is not a mere act of noticing the self but, rather, almost invariably involves comparing the perceived aspect of self against some standard.

The operate phase consists of attempts to alter discrepancies between the perceived aspect of self and the relevant standards. Thus, if the test phase determines that the self is not trying hard enough, the operate phase will most likely consist of efforts to reduce that discrepancy by trying harder. During or after these operations, the self may perform additional tests to see whether the discrepancy has been resolved.

The exit phase terminates the process (as the name implies). When the self has met the relevant standards, there is no need for further operations, and the self's attention can turn to other issues.

The feedback loop is a rather cognitive theory, but of course emotion is a powerful feedback system in the human psyche, and emotion does influence self-regulatory processes. One mechanism is that emotions serve to highlight discrepancies that arise in test phases. Thus, being below the standard is likely to yield aversive emotions, whereas positive emotions may arise when the test reveals the self to have surpassed the relevant standard.

A more nuanced theory was proposed by Higgins (1987), who distinguished different types of standards. As noted earlier, he proposed that different categories of emotion are linked to different kinds of discrepancies. Failing to live up to "ought" goals produces high-arousal, agitated emotions such as anxiety, whereas failing to live up to "ideal" standards causes low-arousal, dejected emotions such as sadness.

Another useful advance was contributed by Carver and Scheier (1990), who concluded that emotions do not simply react to whether a discrepancy exists or not. Emotions are better designed to register change, and so Carver and Scheier proposed that emotions react to the rate of progress toward the goal or standard. If the person is moving toward the goal on or ahead of schedule, positive emotions will be felt, whereas if progress is overly slow, negative emotions will be felt. The novel contribution is that emotion can be positive long before one reaches one's goal, simply because the person feels he or she is making satisfactory progress.

We noted previously that alcohol is implicated in a broad range of self-regulatory failures, probably because alcohol reduces self-awareness (Hull, 1981). Alcohol use thus leads to more alcohol consumption, partly because drinkers soon lose track of how much they have consumed. Based on the notion of external monitoring, we can suggest one policy that might work to reduce problem drinking: Prohibit bartenders and other liquor servers from removing the empty glasses until the patron leaves the establishment. That way, someone who has had six drinks cannot fool him- or herself into thinking it has been just a couple. The empty glasses will furnish a clear tally of the number of drinks consumed.

Improving monitoring is one good way to improve self-regulation. Success is more likely when people observe their own behavior, such as by noticing the types of situations that induce anxiety to anticipate them or even avoid them in the future. People may also rely on external monitoring, such as keeping a journal of when they exercise or how they spend money, or when they ask the bartender not to remove the empty beer bottles.

The Strength Model

The third component of self-regulation is the capacity to make changes. This corresponds to the "operate" phase in the TOTE model, and cybernetic and other theorists were often rather vague about just how self-regulatory "operations" operated. It was plausible that a wide assortment of behaviors could be invoked to resolve discrepancies and bring the self into line with standards. Few obvious generalizations about such operations were apparent.

Recently, however, some evidence has accumulated to suggest that self-regulatory operations tend to consume a limited resource that seems to operate like an energy or strength. The idea that self-regulation depends on such a resource has long been anticipated in various places, including folk wisdom, which has treated "willpower" as an important ingredient in self-control. The term "willpower" implies a strength or energy that the self uses to bring about the changes it seeks. In a more scientific source, the research literature on self-regulation reviewed by Baumeister and colleagues (1994; Baumeister & Heatherton, 1996) also led those authors to conclude that self-regulation seemed to operate as if it depended on a limited resource resembling strength or energy. This would provide a useful explanation for an assortment of empirical findings and informal or anecdotal observations pointing to the apparent pattern that after people exert self-control to regulate some behavior, they seem vulnerable to self-regulatory breakdowns in other and seemingly unrelated spheres. For example, most university personnel have observed that students seem around examination time to exhibit a multitude of signs of poor self-control (e.g., they may eat badly or irregularly, become irritable or rude, resume smoking or other bad habits, or neglect personal grooming). Saying that "stress" causes these consequences is too vague: Stress produces no one of those effects reliably. Instead, the common mechanism may be that when examinations loom, students need to use all their limited self-regulatory resources (i.e., their willpower) to manage their studying, including completing assignments that may be late or have been neglected, and trying to master a complex amount of material in a short period of time. Because the demands of studying and preparing for the examinations consumes a large share of their self-regulatory resources, they have less left over for other, more everyday concerns of self-regulation, such as eating properly, being polite, managing their feelings, and keeping their bad habits under control.

These observations led to the formation of a strength model of self-regulation. Its main ideas are as follows. First, acts of self-regulation consume the limited resource, so that after performing such an act, the person's stock of this resource is at least temporarily reduced. Second, when the resource has been somewhat depleted, the person will be less effective at other self-regulatory tasks. Third, the same resource is used for a wide assortment of self-regulatory activities. Fourth, as with strength, the resource can be restored via rest and possibly other mechanisms. Fifth, also as with strength, regular exercise can increase strength over the long term. Thus, although the immediate result of exercising self-control is to reduce the person's capacity for more self-control, the long-term effect is the opposite, namely, to increase the capacity. Sixth, the self may begin to alter its responses long before the resource is fully depleted. Like athletes conserving their muscle strength when the first part of the athletic contest has begun to produce some degree of fatigue, everyday self-regulators may seek to conserve what is left of their strength when some of it has been depleted.

A series of laboratory investigations sought to test the strength model against other plausible models, including the idea that self-regulatory processes are essentially knowledge modules and hence would operate along the lines of information-processing systems, and the view that self-regulation is a skill that is gradually acquired during development and socialization. Those theories make competing predictions as to what would happen if a person engages in one act of self-regulation and then,

soon thereafter, engages in another act of self-regulation, possibly in a very different sphere. The strength model is based on the notion that the same resource is used for a wide range of different self-regulatory efforts, and so once some of that resource has been expended, subsequent self-regulation will likely be impaired, even in seemingly unrelated areas.

In one study by Muraven, Tice, and Baumeister (1998), participants watched an emotionally distressing film clip under instructions to try to suppress their emotional reactions, to amplify and maximize these reactions, or to let their emotions go without trying to alter or manage them. The first two of those conditions consumed self-regulatory resources as people tried to alter their emotional state, but the third condition would not consume them. Then all participants were given a test of physical stamina in the form of a handgrip exerciser, which they were to squeeze as long as possible. The people who had tried to alter their emotional reactions subsequently performed poorer on the handgrip stamina task, as compared to participants who had not tried to regulate their emotions. Thus, apparently, the effort to regulate emotional responses consumed some of the resource, leaving the people with less to use in performing well on the handgrip task.

In other studies, people who first tried to control their thoughts by suppressing any thoughts about white bears (adapted from Wegner, Schneider, Carter, & White, 1987) subsequently gave up faster on unsolvable anagrams. Trying to suppress the thought of a white bear seems to have consumed some strength, leaving less available for making oneself keep trying and working on the anagram test (Muraven et al., 1998). Likewise, people who had to resist the temptation to eat chocolates and cookies to instead make themselves eat radishes subsequently gave up faster on difficult puzzles (Baumeister, Bratslavsky, Muraven, & Tice, 1998).

These and similar studies (e.g., with solvable puzzles and other dependent measures) yielded generally consistent support for the first crucial aspect of the strength model, namely, that self-regulation depends on a limited resource that becomes depleted when the person exerts self-regulation. Thus it became appropriate to speak of "regulatory depletion" (Muraven et al., 1998).

Other Executive Functions: Making Choices

Not only self-regulation but also acts of effortful choice and volition use the same resource. The stimulus for these studies was Baumeister's (1998) review of research on the self. Baumeister grouped self-regulation under the broader category of the self's executive function, which involves exerting control or choice in relation to the external world alongside efforts to exert control over the self. Baumeister wondered whether the self-regulatory resource would prove to be relevant to choice as well. A first study (Baumeister et al., 1998) borrowed the choice procedure from cognitive dissonance research: Participants were either assigned to make a counterattitudinal speech with no attempt to enlist their

concurrence or make them decide or else requested and subtly pressured into making it, despite the experimenter's insistence that "the decision is entirely up to you" (e.g., Linder, Cooper, & Jones, 1967). Afterward, all participants were given the task of solving unsolvable geometric puzzles, and their perseverance was measured as an index of self-regulatory powers. Apparently, making the choice depleted the resource, insofar as participants in the high-choice conditions quit significantly faster than participants in the various control conditions. This provided a first indication that making an effortful choice depleted the same resources that were needed for self-regulation. The concept of regulatory depletion therefore seemed too narrow. In homage to Freud, whose theory of the ego was one of the last and only energy models of the self, we adopted the term "ego depletion." In the new formulation, this limited resource was needed not only for self-regulation but also for all acts of volition, including choice and active responding (as opposed to passivity).

The link between ego depletion and choice has been made in a further series of studies by Vohs and colleagues (2005). Using a variety of procedures and measures, these authors repeatedly showed that making an effortful choice (or, more commonly, a series of choices) depletes some resource relevant to self-regulation, as reflected in poorer self-regulation afterward. Thus, in one study, participants who made a series of choices about which commercial products they would prefer to own (and one of which they were slated to receive) subsequently showed impaired self-regulation as compared to people who merely rated the same products on a variety of dimension without having to make choices among them. The self-regulation measure consisted of making oneself drink an ostensibly healthful but quite bad-tasting beverage. Ego-depleted participants drank less than one-third as much as those in the control conditions. In another study, participants who had made more choices while shopping gave up faster on a mathematical computation task.

Increasing Strength

Another aspect of the strength model is that self-regulation should improve with regular exercise. If self-control does resemble a muscle, then exercise should strengthen the muscle. Several studies have yielded findings consistent with this view, though each has found considerable noise in the data. A variety of self-control exercises have been used, such as using one's non-dominant hand for routine tasks such as opening doors and brushing teeth, modifying one's speech such as by using complete sentences and avoiding abbreviations or profanity, and cultivating good posture. Participants who perform these exercises regularly for some weeks have been found to perform better afterward in laboratory tests of self-regulation (Muraven, Baumeister, & Tice, 1999; Oaten & Cheng, 2004; Oaten, Cheng, & Baumeister, 2004; for review, see Baumeister, DeWall, Gailliot, & Oaten, in press).

Replenishing the Depleted Self

Perhaps the least well understood aspect of the strength model is how the resource is replenished. Rest appears to be the most common route to replenishment; for example, when people get a good sleep their self-control (even if previously depleted) is better. One sign of this is that self-control appears best after a good night's sleep, whereas it gets weaker as the day wears on, as indicated by the diurnal distribution of self-regulatory failures: Diets are broken in evenings more than mornings, drug or alcohol binges are rare in the morning, most impulsive crimes are committed after midnight, and so forth. These patterns suggest that a person gets up (at least after a good night's sleep) with a full complement of resources, which are then gradually expended as the day wears on. There are also some indications that sleep deprivation weakens self-regulation, though more systematic data would be desirable.

The hypothesis that rest replenishes the self also received some support in a dissertation by Smith (2002). After an initial exercise designed to deplete the self, participants performed a variety of tasks, after which their self-regulatory effectiveness was measured. Participants who had been guided through a brief meditation period prior to the final regulatory task performed much better on it than those who performed other tasks such as reading magazines. Even resting quietly did not work as well as meditation for restoring the self's powers.

Several studies have suggested that positive emotion may help restore the self's resources. In a series of laboratory studies, affect inductions have been interpolated between two self-regulatory tasks. Neutral and bad moods do little to reverse ego depletion, but positive moods (such as induced by watching a humorous video clip) seem to improve the individual's performance on subsequent self-regulation tasks (Tice, Muraven, Slessareva, & Baumeister, 2004). Similarly, asking people to think and write about what is truly important to them (a method of self-affirmation; see Steele, 1988) appears to offer some protection from ego depletion. Three experiments have found that self-affirmation prior to or immediately after initial self-regulatory acts prevents impaired performance on subsequent, target self-regulatory tasks (Schmeichel, Vohs, & Baumeister, 2005). In the main, however, finding strategies to prevent or counteract ego depletion remains an important avenue for further research.

Possible Mechanisms

Thus, self-regulation operates like a strength or energy. It becomes depleted when used, regular exercise appears to increase the resource, and rest may be crucial for replenishing it. What exactly this resource is remains a formidable challenge for future research. At present, there are some early signs that have begun to illuminate the inner processes that attend ego depletion.

The resource does not appear to be closely linked to emotion. Many ego depletion studies have administered emotion measures, and these typically show no effects or changes as a result of depletion manipulations (e.g., Baumeister et al., 1998; Muraven et al., 1998; Vohs & Schmeichel, 2003). The only exception is that sometimes depleted people rate themselves as more tired than other participants, but this effect has only been found in some studies and with some measures. It is possible that depletion is felt as tiredness only when it reaches a certain threshold, even though its effects on behavior appear well before that.

Recent studies addressing the interrelation between the restraint component of self-regulation and the impulse component suggest that depletion may affect how strongly an urge is felt. Vohs and Baumeister (2004) depleted participants by asking them to control their thoughts about a white bear, whereas other participants were given free rein to think about a white bear and thus were not taxed of self-regulatory resources. Subsequent to the mental control task, participants were shown an emotionally provocative video and then immediately afterwards described their feelings in response to the video. Participants who had earlier suppressed thoughts about a white bear reported stronger emotional reactions to the video. This study and others like it (see Vohs & Faber, 2005, who showed that people report more intense urges to buy impulsively when they are depleted) indicate that one consequence of initial self-regulation attempts is an amplification in the experience of impulses and urges, suggesting another route by which self-control fails under depletion.

Most of the work on the ego depletion model has tested new spheres in which self-control processes may be operating (e.g., interpersonal processes, intelligent thought, and addictions), boundary conditions, and specifications of the tenets of the model. A recent paper by Vohs and Schmeichel (2003) attempted to pin down a mechanism to account for the negative effects of resource depletion on subsequent self-control capacity. They identified time perception as one potential signal that people are in a state that precedes a lack of self-control. The experience of time as moving very slowly (i.e., that tasks seem to take more time than they do in actuality) as a mechanism was suggested by findings from the animal literature and from the time perception literature. Animals appear not to have a sense of the far future (Roberts, 2002). Rather, they experience reality as an "extended now" period in which impulses take precedence over anything resembling long- or midterm goals, such as accruing resources (e.g., acorns) to consume later (e.g., in the winter). From the time perception literature, it is known that being highly aware of time (e.g., asking oneself "how long has it been?") leads to perceptions of longer duration (Block & Zakay, 1997). Vohs and Schmeichel (2003) tested the idea that depletion leads to altered time perception and this reduces later self-regulatory ability. More precisely, the researchers found that ego depletion made people more likely to think that they had been continuing their activity longer than people who had not been regulating. Thus, in a sense, depleted people become like animals who are stuck in the

present and unable to orient their behaviors toward future outcomes.

Nonconscious Self-Regulation

The self-control strength model, and indeed the majority of research we have reviewed so far, has considered self-regulation to be a conscious, active process. However, evidence continues to mount for the existence of highly efficient, automatic self-regulatory mechanisms as well.

First, the automotive model of Bargh and colleagues (Bargh, 1990; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Troschel, 2001) explains that goal pursuit, from goal setting to goal completion, may proceed completely outside conscious awareness. Hence some goal-directed activity may occur in the absence of conscious intent and even without the person being aware that they are engaging in goal-directed behavior.

For example, one recent set of laboratory experiments demonstrated that social and behavioral goals could be activated outside conscious awareness and then pursued as if they were consciously intended (Bargh et al., 2001). Research participants who had been primed with the concepts "achievement" or "cooperation" went on to achieve better performance or to cooperate with a partner more readily on a task, respectively, compared to participants who had not been primed with those concepts.

Similarly, simply thinking about a significant other (such as a family member) can prime goals that one associates with that significant other (Fitzsimons & Bargh, 2003) and also prime goals that the other has for the self (Shah, 2003). Once those goals are activated, even though they may not occupy conscious awareness, behavior may conform to the activated goal. For example, among students who believed that their mothers would be pleased by their academic achievement, priming by stimuli related to their mothers led to improved performance on a verbal achievement task compared to participants who did not associate their mothers with academic achievement (Fitzsimons & Bargh, 2003).

Not just goal-directed behavior but also emotional states may also be regulated outside conscious awareness and without intentional intervention. For example, recent research found that shortly after an emotional experience, people spontaneously generated mood-incongruent thoughts presumably as a way to alter their mood state (Forgas & Ciarocchi, 2002). Similarly, Gilbert, Pinel, Wilson, Blumberg, and Wheatley (1998) have posited the existence of a "psychological immune system," which works to ameliorate negative feelings and protect the self from psychological threat. This system is thought to operate automatically, without purposeful self-direction. Again, then, it appears that some emotional states may be regulated nonconsciously.

What is the relationship between conscious and nonconscious self-regulation? Research by Gollwitzer and colleagues (Gollwitzer, 1999; Gollwitzer & Brandstätter, 1997) provides some insight into this relationship. In their study of *implementation intentions*, they have found that forming explicit behavioral plans (e.g., by designating a specific time and place in which goal-directed activity will be pursued) triggers subsequent goal-directed behavior efficiently and automatically, without continued self-intervention. For example, participants who formed explicit implementation intentions enacted goal-directed behavior immediately upon discerning the appropriate conditions to do so (Gollwitzer & Brandstätter, 1997). Similarly, establishing an implementation intention to ignore a target person's gender led to less gender-stereotypical responding after exposure to the target than having no such intention (Gollwitzer, Achtziger, Schaal, & Hammelbeck, 2002). Thus, it appears that conscious intentions can facilitate nonconscious or automatic self-regulation.

According to the self-regulatory strength model outlined earlier in this chapter (see also Baumeister, Muraven, & Tice, 2000; Schmeichel & Baumeister, 2004), conscious and active self-regulation leads to ego depletion, which results in poorer subsequent self-regulation. It is likely that nonconscious self-regulation, because it does not entail purposeful intervention by the self, would not deplete self-regulatory resources. Furthermore, ego depletion should not interfere with efficient nonconscious self-regulation because this type of regulation does not require active involvement by the self.

One recent pair of studies touched on the links between conscious and nonconscious self-regulation by combining implementation intentions and ego depletion (Webb & Sheeran, 2002). These researchers had participants perform the Stroop color–word interference task, a classic cognitive task that depletes self-regulatory strength. The Stroop task requires the performer to inhibit the natural tendency to read a word and to replace the reading tendency with the naming of the ink color in which the word is printed. Webb and Sheeran found that participants who performed the Stroop task gave up more quickly on an ensuing test of self-regulation, consistent with the self-regulatory strength model. However, if participants had formed explicit implementation intentions about Stroop task performance (i.e., to ignore the meaning of the words and to name the colors), they did not show the ego depletion effect. Apparently, forming an implementation intention made Stroop task performance less dependent on conscious self-regulation, and so it became less taxing of self-regulatory strength.

Similar findings come from recent research showing that nonconscious priming can overcome the damaging effects of ego depletion. Weiland, Lassiter, Daniels, and Fisher (2004) depleted some participants using a task involving complex rules about when to cross out "e's" in a page of printed text, whereas other participants were given simple rules to follow. Prior to the editing task, some participants encountered supraliminal but nonconscious achievement-related primes in a scrambled sentence task whereas others encountered neutral primes. On a subsequent task involving unsolvable puzzles, depleted participants who received the neutral prime quit significantly sooner than depleted participants who received the achievement prime. These findings suggest that at least some self-regulatory resource depletion effects involve motivation more than ability, in-

sofar as goals activated outside awareness can overcome deficits in self-control ability due to depletion.

Thus, there is some evidence that depleting self-regulatory tasks may be made less depleting by an act of conscious will. Forming an implementation intention can make later self-regulation more automatic, and therefore less reliant on regulatory strength. Moreover, one study found that nonconscious achievement-goal primes can protect self-regulated behavior from weakened regulatory strength. These encouraging notions deserve further research attention, as does the broader issue of the relationship between conscious and nonconscious forms of self-regulation.

APPLICATIONS

In this section, we focus on applications of the self-regulatory strength model. These show some of the phenomena that are involved in self-regulation.

Cognitive Processing

Intelligent thought, especially logical reasoning, is a supreme achievement of the human intellect that sets humans substantially apart from what other animals are capable of doing. Logical reasoning is hardly natural, in that sense, and thus it may well require guidance by the self's executive function. Simply put, self-regulation of thought may be necessary for successful thinking. Ego depletion may therefore impair the self's ability to think effectively.

A series of studies by Schmeichel, Vohs, and Baumeister (2003) confirmed that some (but not all) forms of intelligent thought are impaired by ego depletion. Specifically, the relatively simple forms of thinking that may proceed automatically were largely unimpaired, whereas thought that required active work to transform one set of ideas into another often showed substantial impairments. To use the terms favored by (some) intelligence researchers, depletion should affect *fluid intelligence* (the capacity to reason, manipulate abstractions, and discern logical relationships), but it should have relatively little effect on crystallized intelligence (involving the retrieval of knowledge acquired via learning and experience; Cattell, 1987; Garlick, 2002). Sure enough, depletion impaired people's performance on fluid intelligence tasks, such as extrapolating from a known to an unknown quantity, or logical reasoning, or being able to take newly acquired information and apply it to novel problems or questions. In contrast, depletion had no substantial effect on such crystallized intelligence tasks as rote memory for nonsense syllables and responding to a test of general knowledge. Thus, apparently, some forms of thought involve self-regulation and depend on the self's precious resource, whereas other (simpler and more automatic) processes do not use this resource. Put another way, ego depletion makes people stupid in complex ways but leaves them intelligent in simple ways (see also Schmeichel, Demaree, Robinson, & Pu, 2006).

Recent work has begun to explore the links between self-regulation and memory. Schmeichel, Gailliot, and Baumeister (2005) set out to show that active self-involvement in memory processes can improve memory, and that resource depletion may undermine this benefit. Previous work had suggested a *self-choice effect* on memory, which is to say that people remember stimuli they have chosen better than stimuli chosen by others or not chosen (e.g., Kuhl & Kazén, 1994). Schmeichel et al. found signs that ego depletion destroys the self-choice effect. The most likely explanation is that when the self has been depleted, people choose in a less effortful and possibly more arbitrary manner, with the result that such choices leave a weaker memory trace. Thus, in one study subjects were instructed by the experimenter to remember some items and forget others; depleted subjects performed just as well as nondepleted ones on this task. However, when subjects were permitted to choose which items to remember and which to forget, depleted subjects performed more poorly (as measured by total recall of both "remember" and "forget" items) than nondepleted subjects. In another study, subjects were given a list of possible tasks. They chose some for themselves to perform, and others were chosen for them by the experimenter (and others were unchosen). Nondepleted subjects remembered the ones they chose best, followed by the ones the experimenter chose, and followed by the unchosen ones—which would seemingly be an adaptive pattern of memory. Depleted subjects, however, recalled the unchosen ones just as well as the self-chosen ones, with the experimenter-chosen ones recalled even worse. The memory for self-chosen items was significantly worse among depleted than among nondepleted subjects.

The idea that the self regulates thought processes is not new, and indeed many of Freud's theories can be read as support for this view. For example, defense mechanisms involve the attempt to think certain thoughts and especially to avoid other thoughts. To be sure, Freudian defense mechanisms did not typically invoke conscious, deliberate control, but presumably some forms of self-regulation can become automated, resulting in nonconscious self-regulation (e.g., Bargh, 1990; Higgins et al., 1994; see the section "Nonconscious Self-Regulation").

A systematic treatment of (often motivated) self-regulation of cognitive processes was provided by Baumeister and Newman (1994). Insofar as self-regulation involves seeking to bring inner responses and processes into line with standards, it was necessary to distinguish two main types of standards that guide thought. One is the truth, in which case self-regulation may seek to guide processing toward the most accurate conclusions, whatever they may be. The other is an a priori preferred view, in which case self-regulation tries to guide thought toward the preordained conclusion. Baumeister and Newman used the metaphors of intuitive scientist for the first (insofar as scientists ideally seek the truth, whatever it may be) and intuitive lawyer for the second (because lawyers try to make the best case for their client, such as to establish his innocence even if he is guilty).

The two self-regulatory goals then yield different prescriptions for regulating thought across four stages of

cognitive processing. The first stage involves gathering evidence. The intuitive scientist seeks to get the most thorough and reliable evidence available, whereas the intuitive lawyer seeks evidence that fits the desired conclusion, such as by means of selective attention and confirmation bias. The second step involves recognition of the implications of various pieces of evidence. This step is mostly automatic, and as such it offers relatively little opportunity for self-regulation.

The third step is reassessment of implications. The automatic conclusions arising from the second step can be scrutinized. The intuitive lawyer may search for sources of bias or distortion that might render some evidence less conclusive than they first seemed. If so, the person might adjust or recompute the tentative conclusions. Meanwhile, the intuitive lawyer may reject some tentative conclusions or implications insofar as they conflict with the preferred conclusion. Selective criticism of unwelcome evidence is one strategy that can be used, such as when people subject hostile evidence to critical scrutiny while accepting congenial evidence uncritically (Lord, Ross, & Lepper, 1979).

The fourth step is to combine and integrate the various implications to formulate a general conclusion. The intuitive scientist can try extra hard to make sure that all viewpoints are considered and that the decision criteria are maximally fair. In contrast, the intuitive lawyer can channel the decision process toward the desired conclusion by selecting decision rules or criteria that favor it.

Probably most people are capable at both intuitive lawyer and intuitive scientist modes of regulating cognitive processes. The intuitive lawyer is useful for self-deception, for supporting a party line or boss's dictates, and generally for helping people sustain the comfortable views they prefer. The intuitive scientist is useful for careful decision making and, more generally, for making decisions in which the person does not have a vested interest in a particular outcome.

Another influential line of research on the self-regulation of thought processes has considered some of the unintended consequences of intentional thought control. Wegner and colleagues (Wegner, 1989, 1994; Wegner et al., 1987) have created an elegant model of the ironic processes of mental control. According to their model, attempts to suppress or inhibit particular thoughts often have as a consequence an increase in those very thoughts. For example, when people try not to think about a white bear, they may succeed temporarily only to experience an abundance of white-bear thoughts a short while later (Wegner et al., 1987). This rebound in the unwanted thought is a result of two related mental processes—a monitoring system, which operates automatically to scan the mental landscape for the occurrence of the forbidden thought, and the conscious operator system, which attempts to occupy the mind with anything but the unwanted thought. The operator requires mental resources to function, and thus successful thought suppression depends on the workings of the conscious operator system. When mental resources are diverted, however, the conscious operator fails to fulfill its function while the automatic monitor continues to operate normally. This combination of events leads to increases in awareness of the unwanted thought.

Earlier we considered the relationship between conscious and nonconscious self-regulation, and Wegner's work on thought control is also relevant in this connection. The automatic monitor in Wegner's model, responsible for detecting occurrences of an unwanted thought, is a nonconscious aspect of self-regulation initiated by the person's conscious attempt at thought control. However, the nonconscious monitor may actually work against the conscious goal of thought suppression by making the forbidden thought more available to conscious awareness. When the conscious operating system is distracted or depleted (which is apparently all too often), the automatic monitor continues to report occurrences of the forbidden thought to conscious awareness, and thus the nonconscious monitor may work against the conscious attempt at mental control.

Interpersonal Processes

We have suggested that self-control abilities probably played a vital role in the social and cultural lives of our forebears, such that those who could suppress or hide their selfish desires acquired advantages that their more uninhibited counterparts did not. For example, keeping one's unpopular views to oneself may have made it easier to get along with others in the group, and it would also have reduced the likelihood of being socially sanctioned or even cast out of the tribe. Recent research has begun to support these speculations by indicating the role of self-regulatory strength in interpersonal functioning.

Self-Presentation

Projecting to others just the right impression of oneself is a tricky task, but one that is crucially important for meeting social goals. Meeting prospective in-laws, negotiating a higher salary, and trying to talk one's way out of extra housework are only a few instances in which self-presentation demands are high. Research by Vohs, Baumeister, and Ciarocco (2005) demonstrated that self-presentation and self-regulation are linked, with each process affecting the other. In a series of studies, Vohs and colleagues found that self-presentation demands influenced self-regulatory ability.

In one study, presenting oneself very positively to a friend or modestly to a stranger—patterns that run counter to typical self-presentational patterns of being modest with friends and enhancing with strangers (see Tice, Butler, Muraven, & Stillwell, 1995)—led to decreased persistence on a series of math problems as compared to acting in ways that are consonant with typical self-presentational patterns. In four additional studies, Vohs and colleagues (2005) found that diminished self-regulatory resources led to less successful self-presentation. In one study participants were asked to ignore (by carefully controlling attention) a series of irrelevant words scrolling across the bottom of the screen of a videotaped interview, whereas other participants saw the same tape (with the same extraneous words) but were not

instructed to ignore the words. The former group would presumably use more regulatory resources than the latter, and therefore be poorer at effortful self-presentation afterward. The self-presentation task in that experiment was self-descriptions of oneself on a narcissism scale. Sure enough, the depleted group was more egotistic, as evidenced by higher scores on the narcissism scale.

The ability to control one's expressive behaviors to make a very specific impression on other people is crucial to interpersonal success. It is, however, costly in the sense that such effortful self-presentation depletes the self of resources that could otherwise be used for a broad variety of goal-related and self-regulatory endeavors. These findings may perhaps explain why people are sometimes unable (or unwilling) to manage their self-presentations optimally to be seen in a socially acceptable light: They are using those limited resources to achieve other regulatory goals.

Close Relationship Maintenance

Interpersonal interactions presumably require some degree of self-regulation, but interacting with people who are demanding or difficult likely requires even more self-regulation. This hypothesis was tested by Finkel, Campbell, Brunell, and Burke (2004), who parsed interactions into two kinds: high and low maintenance. High-maintenance (HM) relationships take effort and are relatively inefficient at accomplishing complex tasks, whereas low-maintenance (LM) relationships are relatively effortless and efficient. Finkel and colleagues found that participants assigned to interact with a HM (vs. LM) confederate later exhibited passivity by choosing to work on easy anagrams as opposed to more challenging anagrams. Indeed, 85% of HM-condition participants chose to work on easy rather than challenging anagrams, whereas only 38% of LM participants preferred the easy task.

In order for romantic relationships to survive, persons in such relationships must be able to cope constructively with negativity on the part of one's partner. Finkel and Cambell (2001) found that the ability to be accommodative and constructive in the presence of a partner's bad behavior relies on self-regulatory resources. In one study, participants in the depletion condition were assigned to control their emotional responses during an emotionally evocative film, whereas participants in the no-depletion condition watched the same film but did not have to control their emotions. Later, all participants were asked how they would respond to a series of potentially destructive partner behaviors (e.g., being 2 hours late for a date), a measure that tapped participants' accommodative tendencies. Finkel and Campbell found that tendencies to be accommodative were lower among participants who had earlier controlled their emotions than among those who had been allowed freely to express their emotional reactions. Hence, people whose regulatory resources had been drained by previous instances of self-control were at risk for responding destructively to others' bad behaviors.

The tendency for people to credit successes to their own internal, stable abilities but to blame others or the situation for failures is called the self-serving bias and it is one of the most reliable attribution effects in psychology (see Campbell & Sedikides, 1999). In the context of a romantic relationship, shared credit for success and taking responsibility for one's role in failure would seem to be beneficial to the health of the relationship. Two persons in a couple who both behaved that way would have a very nice relationship indeed ("Without your help we never would have made it this far" or "I am sorry that I made a mistake"). Initial findings suggest that having more self-regulatory strength allows one to think and speak in these unselfish ways (Vohs & Baumeister, 2004).

Remaining committed to a romantic partner not only means turning a blind eye to the partner's potentially destructive behavior (see Finkel & Campbell, 2001) but also to the positive, attractive aspects of alternative potential partners. Miller (1997) demonstrated that eye-gaze length is an indication of attraction toward another: Short gazes reflect superficial consideration and long gazes reflect deeper processing of the person's attributes. Moreover, Miller's research showed that length of time spent looking at pictures of attractive persons predicted relationship dissolution 2 months later.

Vohs and Baumeister (2004) hypothesized that ego depletion would cause people to look longer at such tempting alternative partners. Just as dieters must turn their attentions away from tempting but forbidden snacks in order to remain faithful to their diet, would-be faithful relationship partners must turn their attentions away from the temptations of new partners. To deplete participants of their regulatory strength, participants were asked to read aloud dull historical biographies under instructions to exaggerate their emotional and facial expressions. In the no-depletion condition, participants read aloud the same biographies but were not given explicit instructions on how to do so. The former was presumed to require more behavioral control and thus tax self-regulatory resources more than the latter. Subsequently participants were told to page through a booklet of scantily clad male and female models, a task they performed while being secretly videotaped. Time spent paging through the booklet was the dependent measure. Consistent with predictions, depleted participants spent more time looking through the book of attractive, near-naked models than did nondepleted participants. Moreover, their slowness did not reflect mere passivity, because the effect was stronger for pictures of opposite-sex models than for same-sex models, as would be consistent with an attraction to alternates hypothesis. When people are low in regulatory strength, they may not have the willpower to turn their eyes away from attractive alternate partners.

Resisting Persuasion

Getting people to do what one wants often entails having to wear down their resistance, which suggests that depleting people's regulatory strength is one route to increasing persuasion. A series of studies by Knowles and colleagues tested this hypothesis by predicting that initial persuasion attempts will be rebuffed more easily than

later persuasion attempts because strength will have been drained combating the earlier attempts, leaving people unable to resist later on. In one study, Knowles, Brennan, and Linn (2004) gauged people's reactions to political advertisements. The results showed participants were indeed most skeptical (i.e., most resistant) when rating the first ad as compared to the last ad, indicating that regulatory resources and thus resistance were worn away with each need to be critical.

Moreover, before rating the last ad, participants were given the task of either listing all the potential problems of going on a Fiji vacation or listing which activities they would like to do on a Fiji vacation. (Participants had earlier watched a video of Fiji vacations so they had the information necessary to complete this task.) The hypothesis was that by having a break from being skeptical and describing the Fijian activities they thought they would enjoy, participants' regulatory strength would be able to rebuild and therefore ratings of the final ad would be more skeptical than if they had to spend the extra time continuing to be skeptical (this time of tropical vacations). The results confirmed this hypothesis in showing that when participants got a pleasant rest from having to be critical, they were then able to be more skeptical when evaluating the last political ad relative to participants who had listed problems with Fiji. However, the finding emerged only for participants who reported being fairly accepting of political ads, suggesting perhaps a practice effect or an individual difference that moderates these effects. In sum, people who battled repeated persuasive attempts became less able to defend against those attempts and consequently became accepting of advertising messages. After a pleasant break, though, the strength rebounded and enabled people to be resistant again. These results also converge with other findings that positive feelings help restore depleted regulatory strength (see Tice et al., 2004).

Controlling Prejudice

Interacting with people of an ethnicity other than one's own may also represent a self-regulatory challenge. Stereotypes and expectations about outgroup members appear to spring automatically to mind in interracial interactions (e.g., Devine, 1989), and so attempting to keep these thoughts at bay may deplete self-regulatory strength. Richeson and Shelton (2003) found support for this view in their study of interracial interactions. When prejudiced White participants interacted with a Black person, they went on to perform more poorly on the Stroop task (a classic measure of cognitive control) compared to when they had just interacted with a White person.

Apparently, face-to-face interaction with an outgroup member is not the only context in which stereotype suppression may deplete self-regulatory strength. Research by Gordijn, Hindriks, Koomen, Dijksterhuis, and Van Knippenberg (2004) found that suppressing stereotypes while writing a short narrative about an outgroup member also led to ego depletion effects. Reduced self-control strength after stereotype suppression was most pronounced among people low in internal motivation to suppress stereotypes (see Plant & Devine, 1998). Moreover, when people with low internal suppression motivation had to suppress stereotypical thoughts, they subsequently showed an increased reliance on stereotypes in general, even stereotypes unrelated to the ones that had initially been suppressed. Presumably, suppressing stereotypes depleted self-control strength so that all manner of stereotypical thoughts increased in salience subsequently.

If suppressing stereotypes is depleting, can exercises aimed at increasing self-regulatory strength enable people to resist stereotypes more easily (i.e., without the detrimental effects)? A series of studies by Gailliot, Plant, Butz, and Baumeister (2004) suggested a positive answer.

The emergence of perceptions and attendant stereotypical associations of stigmatized individuals are affected by people's current self-control strength, according to the results of recent research. Participants who performed a strength-reducing version of the Stroop color-naming task were more likely to mistakenly identify a gun (when it was a tool) after the presentation of a Black (vs. White) face (Govorun & Payne, 2004). This effect occurred only among participants who possessed a strong automatic race bias. Thus, prejudicial tendencies to associate Black faces with dangerous weapons were more likely to emerge and affect behavior when people's regulatory strength was weakened.

Having a stigmatized social identity will likely affect self-regulatory resources in contexts that contain threats related to the social identity. Research by Inzlicht, McKay, and Aronson (2003) supports this contention. The Stroop color-naming task was used to threaten participants, who in the threat condition were told that the task was an intellectual test, which acts as a threat to Black more so than White participants. Time spent completing the Stroop task was used as the dependent measure of self-control, and this measure showed that Blacks who thought the task was diagnostic of intellectual ability performed the task more slowly than Blacks who were not told of the task's purported diagnosticity and more slowly than Whites in the threat condition. A second study with men and women showed effects on a second self-control task, such that women who thought an initial math task was related to gender differences performed worse on a handgrip task than women who did not believe the task was related to gender. They also performed worse than men in the gender differences condition. Thus, self-identification as a stigmatized person can render one vulnerable to ego depletion when faced with a task that accentuates perceived deficits of that identity.

Rejection and Ostracism

Given the supreme importance other people play in our lives and the fundamental nature of the human need to belong, working actively against belongingness needs by ostracizing another person probably requires self-control. Thus, purposefully ostracizing another person may cause ego depletion. In one set of studies, participants who actively ignored another person subsequently

showed poorer self-regulatory performance in terms of physical stamina and persistence in the face of failure (Ciarocco, Sommer, & Baumeister, 2001). Actively ostracizing another person also led to worse mood in the ostracizer, but poorer mood did not account for the poorer self-control. These studies suggest that although self-regulatory abilities probably exist to increase belongingness and interpersonal bonds, they may also be used to ostracize others and prevent bonds from being formed (see Vohs & Ciarocco, 2004).

Rejection can also be bad for self-regulation among the people who are rejected. A series of studies by Baumeister, DeWall, Ciarocco, and Twenge (2005) showed that people who had been rejected by a group or told that their future lives would be lonely performed worse on a variety of self-regulation tasks, including making themselves drink a bad-tasting beverage, restraining their consumption of snack foods, persisting on a frustrating task, and attention control (dichotic listening). Further studies indicated that rejected people were able to self-regulate if there was a compelling, self-interested reason, such as a cash incentive. Thus, apparently, rejection does not render people unable to self-regulate but merely unwilling.

The impact of rejection brings us back to the importance of self-regulation for social connection. Self-regulation enables people to get along with each other, but some of this occurs at a cost to the self, insofar as self-regulation functions to stifle selfish and self-interested impulses in order to do what is best for others (or for the relationship). Humans are social and cultural animals, and so in general the rewards of belongingness are sufficient to justify the sacrifices required for self-regulation. However, when people are socially excluded, they act as if they no longer find it worthwhile to regulate themselves. In that sense, self-regulation is part of an implicit bargain between the individual and society, such that the individual makes the effortful sacrifices in exchange for the benefits of belonging to the group. The bargain can break down on either side. Individuals who fail to self-regulate sufficiently are often rejected by others, such as in divorce, peer ostracism, and even imprisonment. Conversely, when society withholds belongingness (e.g., by rejecting the person), the individual responds with a significantly decreased willingness to self-regulate—except for explicitly selfish rewards.

Individual Differences

Undoubtedly some people are better at self-regulation than others. As noted earlier, a trait measure of self-control was recently published by Tangney and colleagues (2004). It appears to be an effective manner of differentiating people who are good self-regulators from those who are not, although undoubtedly some people may claim better self-control than they actually have. Tangney and colleagues went to great lengths to include many different spheres of self-control in their measure to be able to advance self-regulation theory by establishing a clear factor structure (which would be reflected in the subscales of their measure). However, the factor structure did not replicate well, and all the subscales essen-

tially performed as weaker measures of the full scale. The implication is that self-control is a fairly unidimensional construct, and people who are good at some aspects of self-regulation tend to be good at most of them. This too fits the view of self-regulation as depending on a single, common resource or strength.

Another question that individual difference measures can illuminate is whether there is such a thing as too much self-control. Popular wisdom and anecdotal evidence suggests that people with too much self-control might alienate friends (e.g., by lacking spontaneity) or perform less well in work (e.g., by obsessing about details and failing to make progress). Tangney and colleagues (2004) included a broad set of outcome measures including adjustment, school performance, mental health, and relationship quality, and they aggressively conducted statistical tests for nonlinearity in order to find any downturn in outcomes at the high end of self-control scores. None of these tests yielded any results supporting the notion that a person may have too much self-control ability. Thus, at least to the extent that self-report measures are valid, there is no sign that high levels of self-control produce bad outcomes. The better the self-control, the better the person's other outcomes.

The investigation of individual differences in self-control has also yielded an interesting twist. Self-control trait scores were significantly correlated (at around .5) with scores on a social desirability questionnaire (Crowne & Marlowe, 1960). Social desirability scales are often used as "lie scales" in research, on the assumption that they assess people's willingness to distort the truth to make themselves look good. By this reasoning, it might be assumed that self-control scores are tainted by deceptive self-presentations. On the other hand, we have proposed that self-regulation functions primarily to enable people to overcome selfish impulses so they can behave in ways that are better for interpersonal relations, which means that having self-control should actually and honestly make people perform more socially desirable acts. Tangney and colleagues (2004) found that the effects of self-regulation remained significant and nearly unchanged when they controlled for social desirability, whereas the effects of social desirability on the dependent measures dropped below significance when they controlled for self-control. Thus, it appears that self-control (rather than social desirability) is the more fundamental predictor of positive outcomes, and indeed self-control is probably responsible for many socially desirable acts.

Apart from individual differences in self-control, other individual differences may affect self-regulatory performance. Any given challenge may require self-regulation for one person but not another. Consider alcohol consumption. Restraining alcohol intake probably requires only very little self-control for a person who does not normally drink or who does not particularly care for alcohol. However, some people drink alcohol regularly and may even be addicted to it. Thus, only frequent drinkers should become depleted by restricting alcohol intake. Furthermore, ego depletion should only interfere with alcohol restraint among those who must actively self-regulate their drinking impulses. This view was sup-

ported in experiments reported by Muraven, Collins, and Nienhaus (2002). They found that people who reported a high level of preoccupation with alcohol drank more beer after a depleting thought-control task than similarly preoccupied people who had not done the thought-control task. Beer consumption among people only modestly interested in alcohol was not substantially affected by prior ego depletion.

People may also differ with regard to their social orientation, such that some people are more sensitive to society's demands (i.e., are "other-oriented") than others. As we suggested earlier, self-control abilities probably developed in order to facilitate social interaction and the development of culture. Therefore, people who are preoccupied with smooth social interaction and who prioritize the needs of the group over the needs of the individual should be well practiced at self-control. In support of this view, Seeley and Gardner (2003) found that people high in other orientation were more resistant to ego depletion than people low in other orientation, consistent with the view that other orientation is linked with frequent self-control and therefore greater self-control strength.

Framing a given task in a manner that is concordant or discordant with one's preferred regulatory style is also likely to affect how depleting the task will be. Work by Grant and Park (2003) and Johnson and Shah (2004) indicated that situational demands interact with people's chronic regulatory focus (Higgins, 1997) to affect depletion levels.

In Grant and Park's (2003) studies, students from America and Korea completed two consecutive tasks in which each task was framed as either a promotion task or a prevention task. The researchers took advantage of the finding that Americans are typically promotion focused and Asians are typically prevention focused and hypothesized that it would be less depleting to perform consecutive tasks with a shared regulatory focus (i.e., either two promotion tasks or two prevention tasks), especially when the tasks matched the chronic style of the performer. Their findings were generally supportive of this expectation, such that American students persisted longest on an anagram task (a measure of self-control capacity) when the task and a typing task that preceded it were framed as promotion tasks. Thus, for American students who are mainly promotion focused, two promotion tasks in a row were less depleting than two prevention tasks or either set of mixed-focus tasks. Korean students, conversely, showed the most depletion in the promotion–promotion condition. For Korean students, performing a task that contained at least one prevention-focused aspect buffered against depletion.

Johnson and Shah (2004) took a more evolutionary approach to the study of regulatory focus and self-regulatory strength. They surmised that accomplishing promotion-related tasks would be dependent on the availability of self-regulatory strength, whereas accomplishment of promotion-related tasks would be independent of regulatory strength. In one study they found evidence for this pattern in showing that participants who were depleted by having to use a rule that became more complex between a practice task and the test task were

more likely to solve difficult anagrams under a prevention frame than a promotion frame. In a second study, Johnson and Shah tested for positive emotional states that would suggest a fit between regulatory focus and the situation (e.g., Higgins, Idson, Freitas, Spiegel, & Molden, 2003) and found that depleted participants felt less dejected and more satisfied if they had performed a task under prevention, compared to promotion, instructions, suggesting that prevention was a better fit under depletion than promotion.

These two sets of studies give a hint of what is to come for self-regulatory strength research: the integration of different theories of self-regulation to see where and how they converge. Research by Grant and Park (2003) and Johnson and Shah (2003) illustrates the important role that chronic regulatory focus plays in determining how taxing a given self-control task will be.

The specificity of ego depletion effects among individuals preoccupied with alcohol and people with different social and regulatory orientations highlights the role that chronic differences play in the fluctuation of self-regulatory strength. Surely other individual differences play a role in making some self-regulation particularly depleting for some people but hardly depleting for others. For example, some people are more emotionally expressive than others, and so suppressing emotional reactions should be more depleting for the highly expressive people. Continued study of individual differences and how those differences relate to self-regulatory strength promises to increase understanding of when, and why, some self-regulatory behaviors are particularly taxing.

Affect Regulation

The control of emotional states is a self-regulatory problem that probably touches the lives of every person. Only some people must regulate their alcohol intake or gambling behavior, whereas all people feel emotions and must occasionally strive to manage them. We mentioned earlier some findings suggesting that emotional states may be regulated outside conscious awareness, but much more work has considered the purposeful and active regulation of emotion.

The process model of emotion regulation (see Gross, 1998, 2001) distinguishes between emotion regulation that occurs before the onset of an emotional experience (antecedent-focused emotion regulation) and regulatory effort initiated during or after an emotional experience (response-focused regulation). The best understood antecedent-focused strategy is reappraisal, which entails anticipating an emotional event and resolving not to react to the event by reinterpreting its meaning. For example, a person might remind oneself prior to viewing a scary movie that the events to be depicted are fictional, and that the people in the movie are not actually being tormented by a knife-wielding psychopath. Contrast this reappraisal strategy to the one that requires the active stifling of fear and disgust while watching the movie. By the time the fear hits, one may be too involved in the movie to think rationally about its fictional nature.

Richards and Gross (1999, 2000) studied the cognitive consequences of these two forms of emotion regulation. They found that when people suppressed their emotional reactions while watching a gory slide show, they later had poorer memory for the information presented with the slides than participants who had reappraised what they were seeing. Apparently, stifling emotional reactions interfered with the cognitive processing of the nonemotional information. Reappraisers, by contrast, successfully limited their emotional responses and also showed good memory for the presented information.

The pattern of findings reported by Richards and Gross (1999) suggests that response-focused emotion regulation taxes self-regulatory strength, whereas antecedent-focused regulation may not. Research by Vohs and Schmeichel (2003) confirmed this view. They had research participants suppress, exaggerate, or reappraise their reactions while watching an emotional film clip. Only the response-focused regulators (i.e., the suppressors and the exaggerators) showed reduced self-regulatory strength, while the reappraisers showed no evidence of reduced strength. Thus, consistent with the work of Gross and colleagues, only response-focused emotion regulation reduced self-control strength.

Finally, some evidence suggests that low self-control strength impairs emotion regulation ability. In one study, one group of participants purposefully suppressed a forbidden thought while the other group was free to think whatever they wanted. Later, all participants watched a funny film clip and were instructed to limit their laughter. Those who had suppressed thoughts were relatively unable to prevent themselves from laughing subsequently (Muraven et al., 1998). Thus, ego depletion due to mental control disrupted later response-focused emotion regulation. Whether depletion influences emotion reappraisal or other antecedent-focused regulation strategies is still an open question.

Dieting and Addiction

In everyday life, people most often decide for themselves whether an object is "off limits" or should be denied. That is, individuals frequently create their own regulatory guides (Higgins, 1996). Individual differences in chronic inhibitions are examples of rules or guides individuals undertake to reach their goals. Chronic inhibitions have been studied in terms of their influence on self-regulation under tempting conditions. Externality theory (Schachter, 1968) proposed that obese individuals—who presumably are trying to inhibit food intake—are guided more by external cues than by their internal states. Research by Schachter and his colleagues demonstrated that one consequence of external responsiveness is diminished ability to resist temptation. For instance, Herman, Olmsted, and Polivy (1983) found that obese diners were more likely to order dessert after being given a luscious description of it, relative to when they were simply told that dessert was available. In addition, chronic dieters consumed significantly more snack foods in the presence of salient food cues relative to neutral cues, but this effect did not pertain among

nondieters (Collins, 1978). Thus, dieters appear especially vulnerable to food cues, perhaps because the presence of such cues is more tempting for them than it is for nondieters.

Studies of addictive and compulsive behaviors provide additional evidence for the idea that chronically resisting temptation can lead to deleterious effects, especially with respect to self-regulatory processes (Polivy, 1998). For instance, research on consumer buying habits demonstrates that when consumers resist the temptation to purchase a product, they experience a dramatic increase in desire for the product, which apparently is due to feelings of deprivation (Hoch & Loewenstein, 1991). This effect is exacerbated by the presence of environmental cues that encourage buying, such as free trial periods or free samples. Many people who are dependent on an addictive substance actively try to minimize use of that substance, and research has shown that addicts are especially vulnerable to cues relevant to their particular addiction. For example, smokers exposed to smoking-related cues have been found to exhibit shorter latencies to begin smoking, smoke more cigarettes, report stronger urges to smoke, and show changes in heart rate and blood pressure (Herman, 1974; Rickard-Figueroa & Zeichner, 1985). Likewise, alcoholics who have been exposed to salient alcohol cues report stronger urges to drink (for a review, see Niaura et al., 1988). These studies suggest that individuals set themselves up for failure when they engage in chronic inhibition.

Research on chronic dieting directly tested this suggestion within the framework of the self-regulatory strength model. Vohs and Heatherton (2000) exposed chronic dieters and nondieters to tempting foods that were said to be either available for eating or that were not allowed to be touched (as they were there, supposedly, for a future experiment). The researchers reasoned that dieters but not nondieters would have to actively exert control over their desire to eat the available candies by virtue of their ongoing restriction of off-limit foods. (Indeed, although several of the nondieters dipped in and ate the candies, only one of the dieters did so.) Later, dieters and nondieters were asked to sample three flavors of ice cream ostensibly for the purposes of completing a perceptual ratings task. Dieters who had been tempted by the freely available snack food ate considerably more ice cream than did their counterparts who were told "please, don't touch" the snacks. Nondieters' eating (tested only among those did not partake in the snacks earlier) was unaffected by these manipulations, presumably because they did not have to expend self-regulatory strength in order to not eat the snacks in the earlier phase. Two additional studies confirmed the globality of this effect in showing that a food temptation led dieters to give up sooner on a task involving persistence, and also that an emotional regulation task caused dieters to consume significantly more ice cream consequently.

More recently, work by dieting researchers showed that interpersonal demands in the form of conforming to the group can have significant effects on consumption among people who chronically inhibit their eating. Kahan, Polivy, and Herman (2003) used an Asch-type

conformity task in which dieters and nondieters responded to a visual task either while in a room alone or in a room with confederates who uniformly gave the wrong answers to certain target stimuli. Under the pressure of having to conform, the researchers reasoned, dieters would use up regulatory strength that would otherwise help them not to overconsume food, a prediction that was supported by the increased eating among dieters who were in the conformity condition. Nondieters' eating was unaffected by conformity pressures, not because conformity did not deplete their resources but, rather, because they normally do not put their resources toward curbing caloric intake.

Hence, work on chronic dieters shows how habitual goals interact with situational demands to affect regulated behavior. Whether through emotion control, resisting temptation, or a need to conform, even chronic self-regulation goals can be undermined when momentary pressures deplete precious self-regulatory strength.

CONCLUDING REMARKS

Social scientists have been fascinated by questions of self and identity for many years, but only the past two decades has there been widespread recognition that self-regulation is a centrally important process. Not only does it hold important keys to self theory, but it also has extensive pragmatic applications. Indeed, the majority of personal and social problems faced by modern Western citizens—addiction, violence and crime, debt, sexually transmitted diseases, underachievement, unwanted pregnancy, obesity, failure to exercise, gambling, failure to save money, and others—are rooted in failures of self-regulation.

Self-regulation is one of the key adaptations of the human psyche to enable it to live in cultural groups. It allows people to change their behavior to conform to the expectations of others and, as culture develops, to the abstract rules of the group such as morals and laws. It is an important root of free will in the sense that it enables people to override their first impulses and it furnishes people more complex and flexible ways of deciding and behaving.

This chapter has emphasized a strength model of self-regulation. Altering the self's responses consumes a limited resource that can be conserved, replenished, and even strengthened via exercise. This model is compatible with other contributions to self-regulation theory, such as Higgins's (1987, 1996) self-guide model and Carver and Scheier's (1981) feedback-loop model.

Decades ago, Freud (1930/1961) proposed that most animals could not easily live together in a cultural civilization, and he suggested that some of the psyche's energy had to be rechanneled into the superego in order to make the human being capable of such collective life. Although the march of progress in psychology has moved beyond many of Freud's ideas, in retrospect there does seem to have been something correct about the view that an energy-based capacity for self-regulation is vital for the success of human culture, at both the individual and the collective level. Further research on self-regulation promises to shed light on one of the key aspects to human nature.

ACKNOWLEDGMENTS

REFERENCES

Aronson, E. (1995). *The social animal*. New York: Worth/Freeman.

Bargh, J. A. (1990). Auto-motives: Preconscious determinants of social interaction. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 93–130). New York: Guilford Press.

Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trotschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, *81*, 1014–1027.

Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York: McGraw-Hill.

Baumeister, R. F. (2005). *The cultural animal: Human nature, meaning, and social life*. New York: Oxford University Press.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.

Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Twenge, J. M. (2005). Social exclusion impairs self-regulation. *Journal of Personality and Social Psychology*, *88*, 589–604.

Baumeister, R. F., DeWall, C. N., Gailliot, M., & Oaten, M. (in press). Self-regulation and personality: Strength-boosting interventions and trait moderators of ego depletion. *Journal of Personality*.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*, 1–15.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego, CA: Academic Press.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Baumeister, R. F., Muraven, M., & Tice, D. M. (2000). Ego depletion: A resource model of volition, self-regulation, and controlled processing. *Social Cognition*, *18*, 130–150.

Baumeister, R. F., & Newman, L. S. (1994). Self-regulation of cognitive inference and decision processes. *Personality and Social Psychology Bulletin*, *20*, 3–19.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin and Review*, *4*, 184–197.

Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, *3*, 23–43.

Carver, C. S., & Scheier, M. E. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.

Carver, C. S., & Scheier, M. E. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: North Holland.

Ciarocco, N. J., Sommer, K. L., & Baumeister, R. F. (2001). Ostracism and ego depletion: The strains of silence. *Personality and Social Psychology Bulletin*, *27*, 1156–1163.

Cochran, W., & Tesser, A. (1996). The "what the hell" effect: Some effects of goal proximity and goal framing on performance. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self-regulation* (pp. 99–120). Mahwah, NJ: Erlbaum.

Collins, J. E. (1978). Effects of restraint, monitoring, and stimulus salience on eating behavior. *Addictive Behaviors*, *3*, 197–204.

Cox, S. P. (2000). *Leader character: A model of personality and moral development*. Doctoral dissertation, University of Tulsa.

Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology*, *24*, 349–354.

de Waal, F. B. M. (2001a). *The ape and the sushi master*. New York: Basic Books.

de Waal, F. B. M. (2001b). *Tree of origin: What primate behavior can tell us about human social evolution*. Cambridge, MA: Harvard University Press.

Dennett, D. C. (2003). *Freedom evolves*. New York: Penguin.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Douglas, J. E. (1996). *Mindhunter: Inside the FBI's elite serial crime unit*. New York: Pocket Books.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.

Finkel, E. J., & Campbell, W. K. (2001). Self-control and accommodation in close relationships: An interdependence analysis. *Journal of Personality and Social Psychology*, *81*, 263–277.

Finkel, E. J., Campbell, W. K., Brunell, A. B., & Burke, B. J. (2004). *High-maintenance interaction: Poor interpersonal coordination impairs self-regulatory effectiveness*. Unpublished manuscript, Northwestern University, Evanston, IL.

Fitzsimons, G. M., & Bargh, J. A. (2003). Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology*, *84*, 148–164.

Forgas, J. P., & Ciarrochi, J. V. (2002). On managing moods: Evidence for the role of homeostatic cognitive strategies in affect regulation. *Personality and Social Psychology Bulletin*, *28*, 336–345.

Freud, S. (1961). *Civilization and its discontents* (J. Riviere, Trans.). London: Hogarth Press. (Original work published 1930)

Gailliot, M. T., Plant, E. A., Butz, D. A., & Baumeister, R. F. (in press). Increasing self-regulatory strength can reduce the depleting effect of suppressing stereotypes. *Personality and Social Psychology Bulletin*.

Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review*, *109*, 116–136.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (1998). *Cognitive neuroscience: The biology of the mind*. New York: Norton.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*, 617–638.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Gollwitzer, P. M., Achtziger, A., Schaal, B., & Hammelbeck, J. P. (2002). *Intentional control of stereotypical beliefs and prejudicial feelings*. Unpublished manuscript, University of Konstanz, Germany.

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, *73*, 186–199.

Gordijn, E. H., Hindriks, I., Koomen, W., Dijksterhuis, A., & Van Knippenberg, A. (2004). Consequences of stereotype suppression and internal suppression motivation: A self-regulation approach. *Personality and Social Psychology Bulletin*, *30*, 212–224.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.

Govorun, O., & Payne, B. K. (2004). *Ego depletion and prejudice: Separating automatic and controlled components*. Unpublished manuscript, Ohio State University.

Grant, S. J., & Park, J. W. (2003, October). *The effect of goal orientation on self-regulatory depletion*. Paper presented at the annual meeting of the Association for Consumer Research, Toronto.

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*, 224–237.

Gross, J. J. (2001). Emotion regulation in adulthood: Timing is everything. *Current Directions in Psychological Science*, *10*, 214–219.

Herman, C. P. (1974). External and internal cues as determinants of the smoking behavior of light and heavy smokers. *Journal of Personality and Social Psychology*, *30*, 664–672.

Herman, C. P., & Mack, D. (1975). Restrained and unrestrained eating. *Journal of Personality*, *43*, 647–660.

Herman, C. P., Olmsted, M. P., & Polivy, J. (1983). Obesity, externality, and susceptibility to social influence: An integrated analysis. *Journal of Personality and Social Psychology*, *45*, 926–934.

Higgins, E. T. (1987). Self-discrepancies: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1996). The "self digest": Self-knowledge serving self-regulatory functions. *Journal of Personality and Social Psychology*, *71*, 1062–1083.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*, 1140–1153.

Higgins, E. T., Roney, C. J., Crowe, E., & Hymes, C. (1994). Ideal versus ought predilections for approach and avoidance distinct self-regulatory systems. *Journal of Personality and Social Psychology*, *66*, 276–286.

Higgins, E. T., & Spiegel, S. (2004). Promotion and prevention strategies for self-regulation: A motivated cognition perspective. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 171–187). New York: Guilford Press.

Hoch, S. J., & Loewenstein, G. F. (1991). Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research*, *17*, 492–507.

Hull, J. G. (1981). A self-awareness model of the causes and effects of alcohol consumption. *Journal of Abnormal Psychology*, *90*, 586–600.

Inzlicht, M., McKay, L., & Aronson, J. (2003, February). *Losing control: Stereotype threat as ego depletion*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, Los Angeles.

Johnson, C. S., & Shah, J. (2004, January). *Running on empty: The effects of self-regulatory depletion on goal pursuit*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Kahan, D., Polivy, J., & Herman, C. P. (2003). Conformity and dietary disinhibition: A test of the ego-strength model of self-regulation. *International Journal of Eating Disorders*, *32*, 165–171.

Knight, L., & F. Boland (1989). Restrained eating: An experimental

disentanglement of the disinhibiting variables of calories and food type. *Journal of Abnormal Psychology*, *98*, 412–420.

Knowles, E. S., Brennan, M., & Linn, J. A. (2004). *Consuming resistance to political ads*. Manuscript in preparation.

Kuhl, J., & Kazén, M. (1994). Self-discrimination and memory: State orientation and false self-ascription of assigned activities. *Journal of Personality and Social Psychology*, *66*, 1103–1115.

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, *6*, 245–254.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *38*, 257–269.

Marlatt, G. A., & Gordon, J. R. (Eds.). (1985). *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. New York: Guilford Press.

Miller, R. S. (1997). Inattentive and contented: Relationship commitment and attention to alternatives. *Journal of Personality and Social Psychology*, *73*, 758–766.

Mischel, W., & Ayduk, O. (2004). Willpower in a cognitive–affective processing system: The dynamics of delay of gratification. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 99–129). New York: Guilford Press.

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*, 204–218.

Mischel, W., Shoda, Y., & Peake, P. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*, 687–696.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933–938.

Muraven, M., Baumeister, R. F., & Tice, D. M. (1999). Longitudinal improvement of self-regulation through practice: Building self-control through repeated exercise. *Journal of Social Psychology*, *139*, 446–457.

Muraven, M., Collins, R. L., & Nienhaus, K. (2002). Self-control and alcohol restraint. An initial application of the self-control strength model. *Psychology of Addictive Behaviors*, *16*, 113–120.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*, 774–789.

Niaura, R. S., Rohsenow, D., Binkoff, J. A., Monti, P. M., Pedraza, M., & Adams, D. B. (1988). Relevance of cue reactivity to understanding alcohol and smoking relapse. *Journal of Abnormal Psychology*, *97*, 133–152.

Oaten, M., & Cheng, K. (2004, February). *Longitudinal gains in self-control*. Poster session presented at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Oaten, M., Cheng, K., & Baumeister, R. F. (2004). *Strengthening the regulatory muscle: The longitudinal benefits of exercising self-control*. Manuscript in preparation.

Peele, S. (1989). *The diseasing of America*. Boston: Houghton Mifflin.

Pervin, L. A. (1996). Does it take a gun to the head to assess problems of volition? *Psychological Inquiry*, *7*, 72–74.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.

Polivy, J. (1998). The effects of behavioral inhibition: Integrating internal cues, cognition, behavior, and affect. *Psychological Inquiry*, *9*, 181–204.

Powers, W. T. (1973). *Behavior: The control of perception*. Chicago: Aldine.

Reed, G. F. (1985). *Obsessional experience and compulsive behaviour: A cognitive–structural approach*. Orlando, FL: Academic Press.

Richards, J. M., & Gross, J. J. (1999). Composure at any cost? The cognitive consequences of emotion suppression. *Personality and Social Psychology Bulletin*, *25*, 1033–1044.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and mem-

ory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*, 410–424.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287–290.

Rickard-Figueroa, K., & Zeichner, A. (1985). Assessment of smoking urge and its concomitants under an environmental smoking cue manipulation. *Addictive Behaviors*, *10*, 249–256.

Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin*, *128*, 473–489.

Rothbaum, E, Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. *Journal of Personality and Social Psychology*, *42*, 5–37.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*, 296–320.

Schachter, S. (1968). Obesity and eating. *Science*, *161*, 751–756.

Schmeichel, B. J., & Baumeister, R. F. (2004). Self-regulatory strength. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications*. (pp. 84–98). New York: Guilford Press.

Schmeichel, B. J., Demaree, H. A., Robinson, J. L., & Pu, J. (2006). Ego depletion by response exaggeration. *Journal of Experimental Social Psychology*, *42*, 95–102.

Schmeichel, B. J., Gailliot, M. T., & Baumeister, R. F. (2005). *Ego depletion undermines the benefits of the active self to memory*. Manuscript submitted for publication.

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology*, *85*, 33–46.

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2005). *Self-affirmation and the self's executive function*. Manuscript in preparation.

Searle, J. R. (2001). *Rationality in action*. Cambridge, MA: MIT Press.

Seeley, E. A., & Gardner, W. L. (2003). The "selfless" and self-regulation: The role of chronic other-orientation in averting self-regulatory depletion. *Self and Identity*, *2*, 103–118.

Seligman, M. E. P. (1994). *What you can change and what you can't*. New York: Knopf.

Shah, J. (2003). Automatic for the people: How representations of significant others implicitly affect goal pursuit. *Journal of Personality and Social Psychology*, *94*, 661–681.

Shah, J., Higgins, E. T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology*, *74*, 285–293.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford, UK: Erlbaum.

Sipe, A. W. R. (1995). *Sex, priests, and power: Anatomy of a crisis*. New York: Brunner/Mazel.

Smith, R. W. (2002). Effects of relaxation on self-regulatory depletion. *Dissertation Abstracts International*, *63*(5-B), 2605.

Sobell, M. B., & Sobell, L. C. (1984). The aftermath of heresy: A response to Pendery et al.'s (1982) critique of "Individualized behavior Therapy for Alcoholics." *Behavior Research and Therapy*, *22*, 413–440.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York: Academic Press.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*, 271–322.

Tice, D. M., Butler, J. L., Muraven, M. B., & Stillwell, A. M. (1995). When modesty prevails: Differential favorability of self-presentation to friends and strangers. *Journal of Personality and Social Psychology*, *69*, 1120–1138.

Tice, D. M., Muraven, M., Slessareva, L., & Baumeister, R. F.

(2004). *Replenishing the self: Effects of positive affect on performance and persistence following ego depletion*. Manuscript submitted for publication.

Tomasello, M., & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.

Vohs, K. D., & Baumeister, R.F. (2004). *Depletion of self-regulatory resources makes people selfish*. Unpublished manuscript, University of British Columbia, Vancouver, BC, Canada.

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology*, *88*, 632–657.

Vohs, K. D., Baumeister, R. F., Twenge, J. M., Schmeichel, B. J., Tice, D. M., & Crocker, J. (2004). *Decision fatigue: Making multiple decisions depletes the self*. Manuscript in preparation.

Vohs, K. D., & Ciarocco, N. J. (2004). Interpersonal functioning requires self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 392–408). New York: Guilford Press.

Vohs, K. D., & Faber, R. J. (in press). Spent resources: Self-regulatory resource ability affects impulse buying. *Journal of Consumer Research*.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249–254.

Vohs, K. D., & Schmeichel, B. J. (2003). Self-regulation and the extended now: Controlling the self alters the subjective experience of time. *Journal of Personality and Social Psychology*, *85*, 217–230.

Webb, T. L., & Sheeran, P. (2002). Can implementation intentions help to overcome ego-depletion? *Journal of Experimental Social Psychology*, *39*, 279–286.

Wegner, D. M. (1989). *White bears and other unwanted thoughts.* New York: Vintage.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*, 5–13.

Weiland, P. E., Lassiter, G. D., Daniels, L., & Fisher, A. (2004, January). *Can nonconscious goals moderate self-regulatory failure?* Paper presented at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Self-Interest and Beyond

Basic Principles of Social Interaction

PAUL A. M. VAN LANGE
DAVID DE CREMER
ERIC VAN DIJK
MARK VAN VUGT

What interpersonal orientations drive social interactions? Does selfishness underlie most of our behavior? Are we also inclined to benefit others? Are we naturally committed to sharing and pursuing equality? Do we tend to compete with others, even if we suffer from it by doing so? When and why do we aggress? Such questions are among the most fundamental to understanding interpersonal relations and group processes, which may explain why these topics have attracted the attention of so many scientists from so many fields and disciplines for so long. A complementary reason may be that the questions raised above touch on the long-standing scientific debate about "human nature": Are people by nature good or bad? Thomas Hobbes is often acknowledged as being one of the first to explicitly address this basic question. In *Leviathan* (1651/1996) he raised the interesting problem of why societies and collectivities are able to function at all, if—so he believed—humankind is basically self-interested. The puzzle, which later was termed "the Hobbesian paradox," is central to much theory developed in the social and behavioral sciences. It deals with relationships between the individual and the society at large, but also to smaller scale issues, such as the relationships between individuals in dyads or small groups, and to relationships between groups. How have the social and behavioral sciences sought to solve the Hobbesian paradox?

THE ASSUMPTION OF THE BENEFICENT INVISIBLE HAND

Over a century after Hobbes' writings, Adam Smith (1776) sought to solve the Hobbesian problem by his famous notion of the beneficent *invisible hand*, assuming that private and collective interests tend to correspond rather than conflict. Indeed, in his *Wealth of Nations*, Adam Smith assumed that, for the most part, groups and societies are well-functioning *because* individuals pursue their self-interest. The underlying assumption is that the pursuit of self-interest often has the unintended consequence of enhancing collective interest.

It is now widely acknowledged that Adam Smith's notion of the beneficent "invisible hand" is too limited—and perhaps too simple to be true. In fact, not long after his writings, many scientists came to subscribe to the Hobbesian paradox, assuming that self-interest is often, and in important ways, incompatible with collective interests. As such, the paradox gave rise to two interrelated questions. First, is human behavior primarily or exclusively guided by self-interest? And second, if the costs of selfishness outweigh its benefits, how then can we control selfishness? In the latter question, the costs often refer to collective costs (shared by all involved), whereas the benefits often refer to the gains for the individual.

It took a long time before these questions were studied empirically. In fact, it is only five to six decades ago that some influential books were written that systematically addressed such issues from a formal, mathematical perspective (Luce & Raiffa, 1957; Von Neumann & Morgenstern, 1944) and from a psychological perspective (Thibaut & Kelley, 1959). These books, and especially the empirical research that they inspired, have exerted an enormous influence on the science of interpersonal orientations. First, by systematically analyzing situations, scientists informed each other about the various situations that may (or may not) exist in everyday life. For example, some scientists were able to logically deduce around 96 situations from 2×2 matrices (which represented two persons each having two behavioral options). This work has led to an understanding that there are many kinds conflicts in everyday life—conflicts between self-interest and collective interest, conflicts between self-interest and equality, conflicts between equality and collective interest, and so on.

Second, by actively examining a wide variety of situations, in the laboratory or the field, it became increasingly clear that many situations represent a conflict between self-interest and collective interest. Such situations are omnipresent in our close relationships (e.g., whether to preemptively do the dishes), in relationships with colleagues (e.g., whether or not to prepare very well for a meeting, when it takes costly time to do so), and in our links with organizations or the society at large (e.g., whether or not to engage in citizenship or volunteering activities to help others). Clearly, a relationship is unlikely to be healthy or even to persist if people would not engage in costly acts that benefit the partner. A collaboration between colleagues is unlikely to be fruitful if either or both partners are often ill prepared for a meeting. And a society is unlikely to function well if most people, for example, pollute the environment, never intervene in emergency situations, or volunteer for the greater good of all.

In fact, conflicts between self-interest and collective interests are so pervasive in everyday life that one can go so far as to claim that the most challenging task that governments, groups and organizations, as well as friends and close partners, face is to successfully manage conflicts between self-interest and collective interest. This may explain why many various scientific disciplines have such a long-standing interest in themes that are directly relevant to understanding conflicts between self-interest and collective interest, or *social dilemmas* (e.g., Dawes, 1980; Komorita & Parks, 1995). Beyond the scope of empirical research on social dilemmas, there has been a strong interest in social psychology for cooperation and competition, prosocial behavior, altruism, aggression, trust, reciprocity, and many more. These topics are primarily studied from an interpersonal or small-group perspective, but it should be clear that they have also been studied from an intergroup perspective or from a large societal perspective. Thus, the broad scientific and societal relevance of social dilemmas is beyond dispute.

We are discussing conflicts between self-interest and collective interest in so much detail because it is precisely this domain of situation that is relevant to all the topics discussed previously. If the social world was not social (e.g., the world of Robinson Crusoe before Friday came), or if the world was much like Adam Smith initially imagined (i.e., almost no conflict of interest), many of the specific themes just described would be irrelevant. Cooperation and competition would not be called for, and one cannot communicate or develop trust if there are no conflicts between self-interest and collective interest. This would be a world in which "good and bad" do not seem to matter. But, of course, Robinson Crusoe started to face many opportunities and constraints after Friday's arrival. He and Friday could share food in an equal manner, overbenefit themselves a little every now and then, reciprocate favors over time, cooperate on building shared goods, or compete for scarce resources. Because they became interdependent in terms of fulfilling their basic needs, each of them developed orientations toward each other, which are essential to adapting to they various situations that they face. For example, they could develop orientations toward cooperation, equality, altruism, individualism, competition, or aggression.

In this chapter, interpersonal orientation is broadly defined as the set of cognitions, affect, and motivation that underlie interpersonal behavior and social interaction. We deliberately use a broad definition to reveal its relevance to many interpersonal topics, from affiliation to attachment, and from altruism to aggression. The conceptual basis for interpersonal orientation is derived from Kelley and Thibaut's (1978) interdependence theory assuming that people may *transform* interpersonal situations into new situations that guide their behavior and interactions. Also, we should note that in illustrating principles and mechanisms, we focus on research on social value orientation, which is formally defined as preferences for distributions of outcomes for self and other. The concept of social value orientation deals with prosocial, individualistic, and competitive orientations and often has been examined and conceptualized as an individual difference variable. In this chapter, we use the concept of interpersonal orientations to explicitly acknowledge the assumption that such orientations can be influenced by the person, the situation, or the interaction partner, as we discuss later.

SOCIAL INTERACTION IS A FUNCTION OF PERSONS A AND B AND THE SITUATION

The illustration about Robinsoe Crusoe already illustrates the power of the situation—after Friday's arrival, Robinson's life changed dramatically. Indeed, the essence of a social psychological "way of thinking" is often described in terms of the power of the situation. A classic case in point is, of course, the Lewinian equation $B = f(P, E)$, which assumes that that behavior (B) is shaped not only by properties of the person (P) but also by features of the situation, or social environment (E) (Lewin, 1935). The essence of a social psychological analysis can be even more fully expressed by construing our goals in terms of the relationships *between* two (or more) persons. To develop a truly *social* psychology, we may wish to expand

our formulation, noting that an interaction (I) between persons A and B can be conceptualized in terms of the persons' needs, thoughts, and motives in relation to one another (A and B) in the context of the specific social situation (S) in which their interaction transpires (Holmes, 2002; Kelley et al., 2003; Van Lange, Otten, De Bruin, & Joireman, 1997). Expressed in an equation, $I = f(S, A, B)$.

To illustrate the utility of an interaction-based analysis, imagine two scenarios for John and Mary, who are deciding where to spend their summer vacation. In one scenario their interests conflict in that John wants to go to a beach resort whereas Mary wants to go to Paris. In this type of situation, each person will seek to communicate the basis for his or her preference ("I need the excitement of Paris"), and each will engage in cognitive activity oriented toward understanding the other's needs ("Does John want to relax because he had a stressful year?"). The situation makes it possible for each person to display his or her goals and motives (e.g., selfish vs. prosocial). Communication and information seeking will center on each person's needs, goals, and motives in relation to those of the partner ("Whose needs are more pressing?"; "Will Mary be responsive to my needs?"). The two may rely on fairness norms to resolve their problem ("It's my turn"; "You deserve a break"). Thus, situations involving conflicting interests are interpersonally rich, affording psychological processes such as self-presentation and attributional activity, and activating morality- and benevolence-relevant motives and norms.

In a second scenario John's and Mary's interests correspond, in that both want to vacation in Paris. Neither person is likely to be particularly concerned with information seeking, self-presentation, or attribution in that there is no problem and "nothing to think about." It is not possible for either person to display benevolent motives in that the course of action that would benefit John simultaneously benefits Mary. Interaction is a coordination problem—the two must agree on a date for their vacation, and one person must arrange for travel and lodging. Thus, in comparison to situations with conflicting interests, situations with corresponding interests are relatively simple in that they are less likely to inspire activities such as information seeking or self-presentation and are unlikely to give rise to moral dilemmas or questions of benevolence.

These scenarios very simply illustrate an important point: To understand social interaction we must consider the person (*the Self*), the interaction partner (*the Partner*), and *the Situation*. Likewise, social interaction experiences can be shaped by any of these three components, independently or in combination. For example, a person may be likely to yield noncooperative, selfish interactions because of person influences (e.g., the person does not tend to trust others' cooperativeness), partner influences (e.g., the partner holds in fact a competitive orientation), or situation influences (e.g., the two people often are faced with zero-sum-like situations, with very little opportunity for fruitful exchange through cooperation).

Several theories tend to assume such influences, although often focusing on one of these influences. Models or theories that focus on self-fulfilling prophecies tend to focus more strongly on influences of the Self. For example, individuals with competitive orientations are likely to elicit noncooperative behavior from others, because they expect noncooperation from others, they behave noncooperatively toward others, through which they elicit noncooperative behavior from others—thereby supporting their initial belief that "everybody is selfish" (cf. Kelley & Stahelski, 1970). There are many models that suggest strong partner influences. For example, traditional formulations of attachment theory suggest that early social interaction experiences tend to underlie the development (or not) of secure attachment, and that "partner influences" are strong (Bowlby, 1969). In particular, when the primary caregiver (usually the mother) acts in a cold, unloving, and untrusting manner, the child is unlikely to develop secure attachment—which is more likely to be developed when the primary caregiver is highly responsive to the primary needs of the child, communicating trust and love. Finally, there are some classic theories or models that emphasize the important role of situation. Perhaps the most illustrative example is the Robber's Cave experiment, revealing that the presence of conflicting goals among groups of children undermined friendly behavior and turned it into hostility, distrust, and overt aggression between the two groups (Sherif, Harvey, White, Hood, & Sherif, 1961/ 1988).

Social interactions are important in their own right (i.e., as a topic of study), but we suggest that an interaction-based analysis has strong theoretical benefits. First, it is true, almost by definition, that interaction is a function of the situation and the persons involved. This analysis forces us to analyze situations in terms of what orientations they afford (what orientations they call for, or may activate). Interdependency theory has advanced a taxonomy of situations. The degree to which individuals' interests correspond versus conflict (i.e., covariation in interests), discussed and illustrated earlier, is only one of the six dimensions that contemporary formulations of interdependence theory incorporate (Kelley et al., 2003). The other dimensions capture (1) degree of dependence (how strongly are outcomes determined by the partner's actions or the partner's actions in combination with one's own actions); (2) mutuality of dependence; (3) basis of dependence (whether dependence derives solely from the partner's behavior [partner control], or from partner's behavior in combination with one's own behavior [behavior control]); (4) information availability (e.g., the degree to which we have information about the partner's preferences); and (5) extended situations (e.g., the degree to which interaction situation extend over time and/or the degree to which diverse behavioral options are available). It is beyond the scope of this chapter to fully discuss and illustrate these dimensions (for a detailed overview, see Kelley et al., 2003; Rusbult & Van Lange, 2003). We do wish to note, however, that the dimension of corresponding versus conflicting interest is among the most essential to understanding interpersonal orientations.

Second, the concept of interaction is essential to observation. We never directly see people's motivations or in-

tentions displayed, but we do see two (or more) people reacting to each other, thereby usually producing good or not so good outcomes for each other. Thus we can speak of cooperative interactions (when two people behave cooperatively toward each others), noncooperative interactions (when two people ignore one another's interests), or aggressive interactions (when to people seek to produce bad outcomes for each other). Because observation is essential to social learning and modeling, it is likely that the observation of social interaction, along with the (causal) analysis of it, is an important determinant of our beliefs regarding the orientations that other people may have as well as our beliefs regarding the norms for appropriate conduct. For example, when watching a fighting couple, people may strengthen their belief that most people are not to be trusted and perhaps come to believe that even small forms of verbal abuse are violating norms of decency and respect.

Third, perhaps even more essential than observation, the most direct experiences we have with our social environment are derived from own social interactions. Given that social interactions can be chronically influenced by some situational factors (e.g., the degree to which we needed to share important resources with our siblings), or by an essential interaction partner (e.g., "the primary caregiver"), people may acquire different social interaction experiences. These social interaction experiences are likely to shape the relatively stable interpersonal orientations that people may rely on and use with particular partners (e.g., a prosocial orientation toward one's caring father) or across multiple interaction partners (e.g., a prosocial orientation across most [nonclose] interaction partners). Of course, any interpersonal orientation is subject to continuity *and* change.

Fourth, and finally, psychological processes such as cognition and affect are often both determinants of social interaction and consequences of social interaction. Cognition, motivation, and affect in many ways guide our behavior, reactions, and ultimately interactions. Indeed, much of our thinking and affect is oriented toward making sense of interaction situations and the partner(s) that is so essential to interaction. Automatic or more controlled forms of impression formation are obvious examples—and it is certainly true that much of our thinking and feeling are "for doing" (Fiske, 1992; cf. Jones & Thibaut, 1958). At the same time, during and after social interactions, people are likely to evaluate and summarize their interaction outcomes—for example, cognitions may help us understand the partner's actions, motivation may provide the frame for interpretation while emotions may signal satisfaction or dissatisfaction with the outcomes (along with potential emotions such as anger, disappointment, happiness, etc.). The important point is that, in many ways the concept of social interaction is key to understanding the functions of cognition, motivation, and affect.

To conclude, an analysis that focuses on social interaction has the theoretical benefits of understanding "the Situation," understanding social learning through observation, understanding social development (continuity and change) of interpersonal orientations, as well as cog-

nition, motivation, and affect as determinants and consequences of social interaction.

BASIC PRINCIPLES OF INTERPERSONAL ORIENTATIONS

Which interpersonal orientations help us understand interpersonal behavior and social interaction phenomena? What types of interpersonal orientations, other than selfishness or individualism, should be meaningfully distinguished? Briefly, we suggest the importance of three prosocial orientations (cooperation, equality, and altruism), two proself orientations (individualism and competition), and one antisocial orientation (aggression). The theoretical basis for these orientations is largely derived from interdependence theory (Kelley & Thibaut, 1978) and early research and theory of social value orientation (MacCrimmon & Messick, 1976; McClintock, 1972; Messick & McClintock, 1968). It is interesting to note that this early research and theory by Messick, McClintock, and their colleagues has inspired the transition of a model of social exchange, which largely departed from the assumption of rational self-interest (Thibaut & Kelley, 1959), to the theory of interdependence, which assumes that individuals may "transform" a given situation according to broader orientations, such as cooperation, equality, or competition (Kelley & Thibaut, 1978).

Interdependence theory describes these four non-individualistic orientations in terms of outcome transformations, delineating enhancement of joint outcomes (MaxJoint), minimizing differences between own and other's outcomes (MinDiff), enhancing outcomes for other (MaxOther), enhancing relative advantage over others (MaxRel), and reducing other's outcomes (MinOther). Specifically, the theory argues that given settings of interdependence (i.e., the given matrix) may be transformed according to these orientations to yield a reconceptualized scheme (i.e., the effective matrix), which is more strongly predictive of behavior and social interaction. The given matrix is typically a function of basic, but nonsocial, preferences, such as whether a person prefers to watch movie X or movie Y. When two partners differ in their preferences but want to go to the theater together, they may take into account broader preferences. Such broader preferences are inherently social, because the individual takes into account the partner's preferences, which then yields a reconceptualization of the given matrix. That is, through transforming the given matrix by orientations such as cooperation, equality, altruism, or competition, the individual constructs an effective matrix, which may account for how the individual seeks to solve this interdependence problem (e.g., whether to give in, whether to persist in his or her initial preferences) as well as how the two partners eventually reach a solution (which movie they attend).

The broader considerations, or transformations, may be the product of systematic information processing, shallow or heuristic processing, or even virtually no processing at all (automaticity; Bargh, 1996). In fact, because

TABLE 23.1. An Overview of Basic Propositions of Interpersonal Orientations

Proposition 1
   Most people pursue good outcomes for self, either in the short term, the long term, or both (individualism), but this is often not the sole orientation that people adopt to interaction situations.

Proposition 2
   Interpersonal orientations reflect not only individualism (enhancement of own outcomes) but also cooperation (enhancement of joint outcomes), equality (enhancement of equality in outcomes), altruism (enhancement of other's outcomes), competition (enhancement of relative advantage over others), and aggression (minimization of other's outcomes).

Proposition 3
   The prosocial orientations of cooperation and equality frequently operate in a concerted or interactive manner. That is, these orientations tend to go hand in hand, and it is the interplay of both "prosocial" orientations that best accounts for behavior and interaction in settings of interdependence.

Proposition 4
   Interpersonal orientations are partially shaped by social interactions—therefore, shaped by the self, the interaction partner, and/or the situation.

Proposition 5
   Interpersonal orientations represent different probabilities with which one or more decision rules (e.g., outcome transformations such as MaxJoint, MinDiff) are activated and used.

we encounter several types of interdependence situations quite regularly, often with the same or similar partners, it is plausible that such transformations frequently take place in a habituated, automatic manner. For example, parents may fairly automatically respond to the basic needs and preferences of their children, friends may fairly automatically help each other without a lot of thought, and the desire "to compete" with others may sometimes come into being without any deliberation.

We advance five basic propositions relevant to interpersonal orientations. The term "proposition" is a deliberate choice, as we believe that alternative concepts are either too broad and too remote from the empirical world (e.g., assumptions) or too specific and too closely linked to direct empirical tests (e.g., hypotheses). The empirical literature relevant to these propositions focuses on basic work in social psychology and related fields. Table 23.1 presents an overview of the propositions advanced in this chapter.[1]

INTERPERSONAL ORIENTATIONS AS DECISION RULES

Proposition 1 states that "most people pursue good outcomes for self, either in the short term, the long term, or both, but this is often not the sole orientation that people adopt to interaction situations."

As noted earlier, Thomas Hobbes, and many of his contemporaries, assumed that humankind is basically self-interested, suggesting that humankind involves little (if any) motivation to enhance the well-being of others, to enhance the well-being of the collective, or to enhance equality in outcomes. While many philosophers since Hobbes (and before) held similar views (though less explicitly so), it is perhaps more surprising that this view continued to be influential for a long time. More recently, the notion of self-interest, later extended and termed the "assumption of *rational self-interest*," has dominated much of the traditional theories relevant to interpersonal and intergroup behavior, including early formulations of game theory (Luce & Raiffa, 1957; Von Neuman & Morgenstern, 1944) and of social exchange theory (Blau, 1964; Homans, 1961; Thibaut & Kelley, 1959). This seems especially true for economic theory. As Gordon Tullock (1976), an influential economist and theorist on public goods, once said: "the average human being is about 95 percent selfish in the narrow sense of the term" (cited in Mansbridge, 1990, p. 12).

But within psychology too, the assumption of rational self-interest is embedded in several key constructs, such as reinforcement, the pursuit of pleasure, and utility maximization, as developed in the context of behavioristic theory (including social learning theory), psychoanalytic theory, and theories of social decision making. Moreover, many of the "self-enhancement" phenomena documented in social psychology tend to assume that people seek out material or esteem-related outcomes for the self, often neglecting the power of considerations aimed at benefiting others. Although there is little doubt that people seek to construct realities in ways that serve to maintain or enhance a positive self-imagine (i.e., self-enhancement), it is also likely that similar tendencies are at work in describing close partners, friends, and members considered to belong to the own group (e.g., Murray & Holmes, 1993).

In the current article we do not wish to discard self-interest as a powerful motivation. We do, however, maintain that self-interest tells only part of the story, not all of it. Also, we suggest that Tullock's 95% should be regarded as an overestimation. But why are we so confident that self-interest tells only part of the story? First, several researchers have addressed the fundamental issue of whether people may be willing to make a cooperative choice, in the absence of several (although not all) self-serving goals such as reputational, self-presentational, or reciprocal concerns. Specifically, researchers have designed prisoner's dilemma situations in which participants are strangers who made a single and anonymous choice for relatively large amounts of money and interaction among participants was prevented before and after the experiment. These studies have revealed that under such conditions, a substantial number of people make a cooperative choice (for a review, see Caporeal, Dawes, Orbell, & Van de Kragt, 1989).

Second, in a different program of research, it has been demonstrated that feelings of empathy provide a power-

ful motivation to make a cooperative choice in single-trial prisoner's dilemmas, even if the other had just made a noncooperative choice (Batson & Ahmad, 2001). That is, people who are informed about the misfortune of another person (e.g., partner has ended a relationship) and instructed to put themselves in their position (empathy instruction) tend to act in ways that cannot be understood in terms of self-interest (for an overview of earlier evidence, see Batson, 1998).

Third, the long-standing research on justice and fairness reveals that (at least some) people are often inclined to favor fair outcomes over self-enriching outcomes that represent unequality. A more recent phenomenon is the notion of altruistic punishment, the well-supported tendency for people to punish others (at a cost to themselves) who fail to cooperate and thereby undermine the "cooperative atmosphere" in a small group (Fehr & Gächter, 2002). This phenomenon too clearly shows that people are strongly motivated to pursue equality and to "do justice" to those who tend to exploit others.

Fourth, what is impressive about the lines of research just described is that considerations other than selfishness can be observed with relative strangers, with whom they interact in a fairly abstract social dilemma task, often under completely anonymous conditions. Clearly, in the context of ongoing relationships, people should be quite prepared to engage in self-sacrificial acts, to "nurture," or to accommodate in an attempt to promote the well-being of family members, close partners, and friends (see Rusbult & Van Lange, 2003). Although such tendencies are not easy to isolate from long-term selfish interest in ongoing relationships (because there is a history and future to the relationship), research on communal relationships suggests that prosocial behavior often may occur in the absence of "recordkeeping" or reciprocity in favors. That is, people tend to respond to variation in the other's needs, and less so (or not at all) to whether the partner has engaged in similar acts in the past (Clark & Mills, 1993). And the fact that people harbor exceedingly favorable views of close others is certainly consistent with the notion that the partner's ego is quite important to themselves as well (Murray & Holmes, 1993).

Last but not least, the long-standing program of research on social value orientation, to be discussed later, is strongly at odds with the view of self-interest. In fact, this program of research was initiated in part because early research on the prisoner's dilemma and the like revealed pronounced intraindividual consistency in tendencies toward cooperation or selfishness.

Thus, various lines of research provide support for the notion that selfishness is not the only orientation that people adopt in interaction situations with others—close others, or even complete strangers. In this respect, we agree with recent insights that suggest that the importance of self-interest may be overstated. Miller and Ratner (1998; see also Ratner & Miller, 2001), for example, demonstrated that participants overestimate the impact of financial rewards on their peers' willingness to donate blood, as well as the power of social rewards (as assessed by group membership) on their peers' attitudes. Also, research has revealed that people tend to assume that most others adopt an individualistic orientation to a prisoner's dilemma, believing that most others are simply seeking to enhance their own outcomes with no or very little regard for other's outcomes (Iedema & Poppe, 1994; Maki & McClintock, 1983).

There may be several mechanisms that support the "myth of self-interest." For example, people are more likely to reciprocate noncooperation than to reciprocate cooperation. The implication is that a belief in the selfishness of others is more easily confirmed than a belief in the cooperative nature of others (Kelley & Stahelski, 1970). There are several specific mechanisms as well that support selfishness rather than cooperativeness. One example is the strong tendency for people to assign greater weight and attention to negative behaviors than to positive behaviors (e.g., Fiske, 1980; Skowronski & Carlston, 1989). Another mechanism derives from the availability of information. Often in the context of groups, what we can observe (noncooperative interaction) may actually be due to a few or even only one person, in that the cooperative intentions are (often) not visible. In other words, observable noncooperative behavior in groups may be due to noncooperative intentions of only a few group members. Finally, at the societal level, the myth of self-interest tends to be supported in the media, which tends to focus more on the bad parts of human nature than the good parts.

To conclude, we suggest that self-interest is a powerful motivation, but one that is often overestimated in strength. Such overestimation often is accompanied by a neglect of other important interpersonal orientations, to which we direction our attention next.

Proposition 2 states that "interpersonal orientations reflect not only individualism (enhancement of own outcomes) but also cooperation (enhancement of joint outcomes), equality (enhancement of equality in outcomes), altruism (enhancement of other's outcomes), competition (enhancement of relative advantage over others), and aggression (minimization of other's outcomes)."

Cooperation

There is a fair amount of research showing that the enhancement of joint outcomes, or cooperation, is an important consideration. People have a pronounced tendency to consider not only outcomes for themselves but also outcomes for others. The enhancement of joint outcomes may sometimes take the form of self-interest and assigning positive weight to other's outcomes (or doing no harm to others). But perhaps just as often, or more often, the enhancement of joint outcomes takes the form of enhancing outcomes for the group as a whole (a tendency sometimes referred to as collectivism, see Batson, 1994). In terms of decision rules, in both cases, individuals tend to enhance joint outcomes (even though they may assign greater weight to outcomes for self than to outcomes for other).

Psychologically, the two types of cooperation are substantially different. The tendency to assign some positive weight to other's outcomes may be accompanied by a va-

riety of mechanisms, such as want to act in line with the "no harm" principle (Batson, 1994), adopting a norm of social responsibility, which dictates helping. The tendency to enhance group outcomes may readily be activated (e.g., at the very beginning of group formation), and it is powerfully activated by identification with the group (e.g., Brewer & Kramer, 1986; Kramer & Brewer, 1984). To the extent that a person feels more strongly part of the group and valued by the group, or the extent to which a person derives self-definition and esteem from the group, individuals are more likely to behave cooperatively. A classic case in point is research by Brewer and Kramer (1986), in which participants were categorized as psychology students (i.e., the actual participants, hence strong group identity) or economics students (i.e., weak group identity). Using a specific resource dilemma, Brewer and Kramer showed that under conditions of strong identity, individuals were more likely to behave cooperatively when it was essential to the group (i.e., when the resources were near depletion). Such cooperative efforts were not observed when group identity was low. It has been suggested that under conditions of strong identity, there may be a blurring of the distinction between personal outcomes and collective outcomes— that is, me and mine becomes we and ours, just as we and ours becomes me and mine (e.g., De Cremer & Van Vugt, 1999).

Egalitarianism

The existence of egalitarianism or equality may be derived from various lines of research. To begin with, several experiments have been conducted within the realm of resource-sharing tasks to examine the factors that may determine different "rules of fairness." In these tasks, a group of people shares a resource and the problem that these decision makers are confronted with is how to optimally using the resource without overusing it. Research by Allison and Messick (1990) provided a powerful demonstration of what happens in such situations. That is, their results showed that when participants (in a group of six people) are asked to harvest first from the common resource, people almost without exception use the equal division rule. Individuals tend to favor equality in outcomes (rather than more complicated rules of fairness; for related evidence, see Van Dijk & Wilke, 2000). Allison and Messick (1990) suggested that equality represents a decision heuristic that has the advantages of being simple, efficient, and fair. Equality has great potential to promote the quality and effectiveness of interpersonal relationships, and therefore it can be considered a "decision rule" that is deeply rooted in people's orientations toward others (see also Deutsch, 1975, Grzelak, 1982; Knight & Dubro, 1984).

Another powerful illustration of equality in interdependence situations is when people have to negotiate allocations (e.g., how to allocate monetary outcomes). This problem is often addressed in research on ultimatum games, an exceedingly popular paradigm in experimental economics (see Güth, Schmittberger, & Schwarze,

1982). In this negotiation setting, two players have to decide on how to distribute a certain amount of money. One of the players, the allocator, offers a proportion of the money to the other player, the recipient. If the recipient accepts, the money will be distributed in agreement with the allocator's offer. If the recipient rejects the offer, both players get nothing. Some of the first studies using this research paradigm demonstrated that allocators generally proposed an equal distribution (i.e., a 50–50 split) of the money (for an overview, see Camerer & Thaler, 1995). Subsequent studies, however, wondered whether this was true fairness and that allocators may have acted out of fear that recipients would reject their offer. Recent evidence suggests that at least some people do persist in employing the equality rule in ultimatum games, even when recipients can be cheated on or when recipients hardly have any power over the decision to reject the offer or not (see Van Dijk, De Cremer, & Handgraaf, 2004). Again, equality seems to be an orientation that people carry with them when engaging in social interactions.

Although equality is in the eye of many the prime example of fairness, we already noted that fairness might also take different forms, independent of outcomes. More precisely, allocating outcomes is always accompanied by procedures guiding allocation decisions (Thibaut & Walker, 1975). People also wonder about how fair these procedures are and these perceptions in turn also have strong effects on people's behaviors and experiences in social relationships (De Cremer & Tyler, 2005). The focus on procedural fairness was further inspired by research showing that when people are asked to talk about their personal experiences of injustice they usually talk primarily about procedural issues, in particular about being treated with a lack of dignity and politeness when dealing with others (e.g., Messick, Bloom, Boldizar, & Samuelson, 1985; Mikula, Petri, & Tanzer, 1990).

Moreover, there is research revealing that the opportunity for "voice" (e.g., being asked your opinion) may convey strong surplus value in that people feel more strongly valued and respected. Voice also means that people are given an opportunity to express their values (i.e., "value-expressive" worth). For example, some research shows that people still rated a procedure to be fairer if they had voice than if they lacked voice, even if they estimated that what they said had little or no influence on the decisions made and on the outcomes that one would receive (Tyler, Rasinski, & Spodick, 1985).

An important field study by Tyler and Degoey (1995) examined people's perceptions of the fairness of the legal authorities in California and their sense of identification with their state. At the time of their study, California was plagued by a severe drought and people had to try to maintain water resources—a situation that resembles a social dilemma. Results revealed that perceptions of procedural fairness (i.e., how accurate, ethical, neutral, consistent, and participative they perceived the procedures enacted by the authority) significantly influenced people's willingness to save and maintain water resources. Especially when they exhibited a strong sense of identification with the community. High identifiers particularly

cared about the fairness of the procedures because this indicated to them that they were valued society members and thus should be treated with respect (Tyler & Lind, 1992). More recently, De Cremer and Van Vugt (2002) experimentally demonstrated the powerful effects of procedural fairness on cooperation behavior in a public good dilemma by showing that a procedurally fair leader (i.e., a leader allowing voice to group members in the decision on how to allocate the public good) promoted prosocial behavior, but particularly among those who identified strongly with the group toward the group (i.e., high group identifiers). These results thus indicate that procedural fairness, independent of outcomes, guides people's actions in social relationships, and especially when the focus is on the common group. More recent research supports the notion that procedural fairness (examining by the availability of voice or not) often is used as a cue or heuristic as to whether "the authority" is to be trusted. In fact, Lind (2001) notes that "people use overall impressions of fair treatment as a surrogate for interpersonal trust" (p. 65) (for empirical evidence, see Van den Bos, Wilke, & Lind, 1998).

To conclude, egalitarianism has received attention in distinct literatures, often supporting the notion that equality in outcomes and treatment is deeply rooted in our system and often serves as the norm as well as a heuristic for own actions and expectations regarding other's actions.

Altruism

The claim that altruism should be considered an interpersonal orientation is rather controversial. Indeed, as most readers know, there has been a fair amount of debate about the existence of altruism both within and beyond psychology. Much of the controversy, however, deals with definitions of altruism, ranging from behavioral definitions (i.e., acts of costly helping are considered altruistic) to definitions that seek to exclude any possible mechanism that may be activated in some way by self-interest. If we limit our discussion, for parsimony's sake, to research on cooperation and competition, and to allocation measures, then we see that altruism is not very prominent. For example, in assessments of interpersonal orientations in a specific resource allocation task, the percentage of people who should be classified as altruistic (i.e., assigning no weight to their own outcomes while assigning substantial weight to other's outcomes) is close to zero (Liebrand & Van Run, 1985). Similarly, when people playing a single-choice prisoner observe that the other makes a noncooperative choice, the percentage of cooperation drops to 5% or less (Van Lange, 1999).

But this evidence should not be interpreted as if altruism does not exist. In fact, what is more likely is that it does not exist under the (interpersonal) circumstances that are common in this tradition of research. People usually face a decision-making task, be it a social dilemma task, a resource allocation task, or a negotiation task, in which they are interdependent with a "relative stranger" in that there is no history of social interaction or other

form of relationship. Accordingly, there is no basis for feelings of interpersonal attachment, sympathy, or relational commitment. We suggest that when such feelings are activated, altruism may very well exist.

As alluded to earlier, recent research by Batson and Ahmad (2001) provides convincing evidence. Specifically, they had participants play a single-trial prisoner's dilemma in which the other made the first choice. Before the social dilemma task, the other shared some personal information that her partner had ended the relationship with her, and that she finds a it hard to think about anything else. Batson and Ahmad compared three conditions, one of which was a high-empathy condition in which participants were asked to imagine and adopt the other person's perspective. The other conditions were either a low-empathy condition, in which participants were instructed to take an objective perspective on the information shared by the other, or a condition in which no personal information was shared.

After these instructions, participants were informed that the other made a noncooperative choice. Batson and Ahmad found that nearly half of the participants (45%) in the high-empathy condition made a cooperative choice, while the percentages in the other low-empathy and control conditions were very low, as shown in earlier research (less than 5%, as in Van Lange, 1999). Hence, this study provides a powerful demonstration of the power of empathy in activating choices that can be understood in terms of altruism, in that high-empathy participants presumably assigned substantial weight to the outcomes for the other at the expense of their own outcomes.

Also, the existence of altruism was also supported by earlier research that was designed to test the hypothesis that feelings of empathy could promote choices that benefit one particular individual in a group rather than cooperation that benefits the entire group (Batson et al., 1995). Specifically, participants could choose to benefit themselves, the group, or other group members as individuals, which extends the dichotomy of self versus collective-as-a-group that is so common in social dilemma research. Using experimental manipulations of empathy (study 1) and naturally occurring variation in empathy (study 2), Batson et al. found that feelings of empathy created or enhanced the desire to benefit one particular other person in the group (i.e., the one for whom strong empathy was felt), thereby reducing tendencies toward benefiting the collective. This study indicates that just as tendencies toward individualism may form a threat to collective well-being, so may tendencies toward benefiting specific others, or altruism, form a threat to collective well-being. That is, feelings of empathy may lead one to provide tremendous support to one particular person, thereby neglecting the well-being of the collective. For example, as noted by Batson and colleagues (1995), an executive may retain an ineffective employee for whom he or she feels compassion to the detriment of the organization. We suggest that such tendencies toward altruism are likely to be observed when individuals deal with others with whom they have developed attachment, closeness, or sympathy.

Competition

There is also strong evidence in support of competition as an orientation quite distinct from self-interest. As noted earlier, the work by Messick and McClintock (1968) has inspired considerable research that reveals that not only cooperative orientations but also competitive orientations may underlie social interactions. For example, Kuhlman and Marshello (1975) have demonstrated that individuals with cooperative orientations do not tend to exploit others who exhibit cooperation at every interaction situation, irrespective of the individual's own behavior. They also showed that individuals with competitive orientations do not exhibit cooperation, even if cooperative behavior, rather than noncooperative behavior, best serves their own personal outcomes. For example, when interacting with a partner who pursues Tit-for-Tat (Axelrod, 1984), which begins with a cooperative choice and subsequently makes the same the choice as the other did in the previous interaction situation, it make sense to cooperate if one is selfishly oriented. The reason is that cooperative choices yield mutual cooperation (good outcomes), whereas noncooperative choices yield mutual noncooperation (less good outcomes). Interestingly, unlike individualists who do respond cooperatively, competitors do not tend to behave cooperatively in response to a Tit-for-Tat strategy. The plausible reason is that competitors do not seek to enhance their own outcomes in an absolute sense—they seek to maximize the gain (or minimize the losses) *relative* to the other person.

The importance of competition is even more directly shown in research on a decision-making task that represents a conflict between on the one hand cooperation and individualism (option A) and on the other hand competition (option B). Hence, the only consideration to choose option B is to receive better outcomes (or less worse outcomes) than the other, even though one could do better for oneself by choosing option A. Research using this so-called maximizing difference game has revealed that quite a few people choose the competitive alternative; it is also of some interest to note that among some (young) age groups competitive tendencies tend to be even more pronounced (McClintock & Moskowitz, 1976). Specifically, among very young children (3 years old) individualistic orientation dominates, after which competition becomes more pronounced (4–5 years), which is then followed by cooperative orientation (6–7 years).

Finally, one might wonder whether it is the aversion of "getting behind" or the temptation of "getting ahead" that underlies such competition. In a very nice study by Messick and Thorngate (1967), it was shown that the former tendency (aversive competition) is much more pronounced than the latter tendency (appetitive competition)—in other words, not losing seems a strong motivation than winning. This early research was later extended, and generalized, by Kahneman and Tversky's (1979) gain and loss frames in their prospect theory, and by Higgins's (1998) distinction between prevention and promotion focus as two distinct self-regulatory systems. Recent research has also revealed that under condi-

tions of uncertainty, competition may be especially pronounced, presumably because people really want to make sure that they do not get less than the other (Poppe & Valkenberg, 2003). Thus, there is little doubt that competition is an important orientation that needs to be carefully distinguished from self-interest.

Aggression

The orientation of aggression has received very little attention in research on social dilemmas. It is interesting to note that, especially in comparison to the orientation of altruism, much research on aggression focuses on genetic and biological factors. Examples are not only twin studies but also studies focusing on associations of aggression with hormonal activity, such as variations in levels of testosterone. Generally, this body of research supports the view that aggressiveness, examined by self-report methodology, is substantially "influenced" by genetic factors and biological makeup. For example, research shows that manipulations of levels of testosterone, varied as part of a treatment for sexual transformations, influence the proclivity to anger. There is an increase in the tendencies toward anger among individuals who transform from woman to man, and a decrease in such tendencies among individuals who transform from man to woman (Van Goozen, Frijda, & Van de Poll, 1995).

Importantly, the correlation between aggressiveness and testosterone is especially pronounced for scale items assessing aggressiveness-in-response-to-provocation (Olweus, 1979), suggesting that aggression needs to be considered in terms of anger that is interpersonally activated. Indeed, the methods typically used to study aggression consist of examining aggressiveness in response to provocation by another person. Hence, anger and aggressiveness should be easily aroused by others who fail to exhibit cooperative behavior. Indeed, the fact that there is not much systematic research on aggression in social dilemmas is not to imply that aggression is not an important orientation or motivation in the context of social dilemmas. We suspect that many or most of the readers who have conducted social dilemma experiments will immediately recognize not only the involvement but also the hostility described by Dawes, McTavish, and Shaklee (1977):

> One of the most significant aspects of this study, however, did not show up in the data analysis. It is the extreme seriousness with which subjects take the problems. Comments such as, "If you defect on the rest of us, you're going to live with it the rest of your life," were not at all uncommon. Nor was it unusual for people to wish to leave the experimental building by the back door, to claim that they did not wish to see the "son of bitches" who double-crossed them, to become extremely angry at other subjects, or to become tearful. (p. 7)

Because it is unlikely that aggression is a self-activated phenomenon in social dilemmas, people are unlikely to approach one another aggressively, with the primary goal in mind to reduce the outcomes for other(s). As noted earlier, aggression may be activated when others

fail to cooperate. This interpersonal basis of aggression is important, and suggests several interesting phenomena. For example, it may well be that tendencies toward aggression are most pronounced among those who do not expect others to behave selfishly. As a point in case, Kelley and Stahelski (1970) provide some evidence for what they referred to as *overassimilation*, the tendency for cooperative individuals (at least, some cooperative individuals) to behave eventually even more noncooperatively than the fairly noncooperative partner with whom one interacts (see also Liebrand, Jansen, Rijken, & Suhre, 1986).

But why might people respond so aggressively to noncooperative behavior by others? Is it only because the other's noncooperative behavior provides one with much less good outcomes than the other's cooperative behavior? We think not. In fact, it may well be strongly linked to a violation in equality of outcomes that often is created (and often perceived as intentionally created) by the other's noncooperative behavior. But then the question becomes, "Why would people respond so aggressively to a violation of equality in outcomes?" Speculatively, three reasons seem especially noteworthy.

First, a violation of equality is generally easily observed. When comparing two outcome situations, it seems easier to compare both situations in terms of equality in outcomes than it is to compare them in terms of quality of joint outcomes (cf. Allison & Messick, 1990). Second, people often use social standards for evaluating the quality of their own outcomes (cf. comparison level; Kelley & Thibaut, 1978). In the context of a social dilemma, the social standard (or social comparison) is also salient (1) because typically people can "explain" any given outcome directly in terms of the other's behavior, and to some degree, the other's intentions, and (2) because individuals' own behavior, at least in part, may be guided by expectations regarding other's behavior (e.g., Kelley & Stahelski, 1970). Third, people are generally aversive to receiving fewer good outcomes than others. One is reminded here of classic research by Messick and Thorngate (1967), revealing that aversive tendencies toward ensuring that the other does not attain greater outcomes than oneself are stronger than "appetitive" tendencies toward attaining greater outcomes for oneself. In most situations, a violation of equality, caused by others' noncooperative behavior, may not only hinder or frustrate one's interaction goals, but also negatively influence a person's pride, honor, or self-esteem (i.e., two consequences that are likely to instigate anger, see Averill, 1982).

It is interesting that responses to aggressive acts (specifically, offenses) have recently received greater attention in studies on interpersonal forgiveness. In support of the notion that (aggressive) offenses often are violations of justice, it has been shown that forgiving is effectively promoted by a compensatory act or an apology by the offender (McCullough, Worthington, & Rachel, 1997). If such restorations are not made, forgiving is less likely to happen, especially when justice concerns remain prominent. Such may lead to an inability to forgive, which in turn may challenge quality of relationships and

undermine psychological well-being (e.g., Karremans, Van Lange, Ouwerker, & Kluwer, 2003). Aggression is, of course, by no means confined to dyads or small groups. Also, in large-scale social dilemmas, aggression, or at least subtle forms of aggression, may account for patterns of reactance, resistance, protest, and so on. Such aggression is often evoked by the behavior of specific group members, managers, or local and global authorities. Much research on large-scale social dilemmas has focused on individuals' willingness to contribute or cooperate, which may be regarded as a line of research that would benefit from greater attention for the opposite side of the coin (i.e., examining the psychological aspects of individuals' readiness to aggress in subtle or more explicit ways). Also, the topic of forgiveness is, of course, of great relevance to resolving conflict between large groups. To conclude, it is surprising that aggression has received so little attention in social dilemmas, because—unless research suggest otherwise—aggression seems an important orientation in social dilemmas, albeit one that seems activated primarily by the behavior of others.

Proposition 3 states that "the prosocial orientations of cooperation and equality frequently operate in a concerted or interactive manner. That is, these orientations tend to go hand in hand, and it is the interplay of both "prosocial" orientations that best accounts for behavior and interaction in settings of interdependence."

Thus far, we distinguished among six orientations, which, in decreasing order of benevolence, are (1) altruism, (2) cooperation, (3) egalitarianism, (4) individualism, (5) competition, and (6) aggression. As noted earlier, it is unlikely that each of these orientations operates in a completely independent manner. We argue that two or more orientations may well activate each other in some way, and thus may over time become "psychologically interrelated" orientations. As illustrated in Table 23.2, we suggest a model of interpersonal orientations that focuses on five relatively distinct interpersonal orientations, whereby "prosocial orientation" is the broader term representing both cooperation and egalitarianism as two interrelated orientations.

There is good theoretical and empirical reason to believe that at least two "prosocial orientations" (i.e., coop-

---

TABLE 23.2 An Overview of Five Orientations

1. *Altruism*
   Enhancement of outcomes for other

2. *Prosocial orientation*
   Enhancement of joint outcomes (cooperation)
   Enhancement of equality in outcomes (egalitarianism)

3. *Individualism*
   Enhancement of outcomes for self

4. *Competition*
   Enhancement of relative outcomes in favor of self

5. *Aggression*
   Reduction of outcomes for other

eration, and egalitarianism) tend to go hand in hand, at least in social dilemmas. How so? To begin with, one very robust phenomenon observed in the two-person prisoner's dilemma is the phenomenon of *behavioral assimilation* (Kelley & Stahelski, 1970). This phenomenon, which may also be referred to as reciprocity, holds that individuals with a prosocial orientation cooperate with others who also cooperate but turn to noncooperation when others do not cooperate (i.e., they become behaviorally similar to noncooperative others). The phenomenon of behavioral assimilation has been observed and supported in the most intensely studied prisoner's dilemma. Importantly, one could theoretically infer that if an individual is merely concerned with enhancing joint outcomes, one should behave cooperatively *irrespective* of the other's behavior. Thus, individuals with prosocial orientations should do more than simply enhancing joint outcomes. In fact, a model in which prosocial orientation is understood in terms of (1) egalitarianism alone or (2) cooperation and egalitarianism together (an "integrative model") is able to account for behavioral assimilation.

In past research, the phenomenon of behavioral assimilation has been supported only in research on iterated social dilemmas. In such repeated choice situations, reciprocity could be guided by a multitude of specific considerations, following from an interplay of other's past choices (or past interactions) and individuals' long-term interaction goals (e.g., the perceived feasibility of attaining particular interaction goals). For example, a partner's past actions may to some degree influence considerations relevant to long-term interaction goals, because the partner's past actions (e.g., noncooperative choices) might bring about beliefs regarding the feasibility of attaining particular long-term interaction goals (e.g., diminished confidence in the feasibility of establishing patterns of mutual cooperation). Thus, because considerations regarding the past, present, and future are inextricably linked to patterns of choice in iterated prisoner's dilemmas, it is difficult to understand the specific considerations and motivations that underlie patterns of reciprocity (but Gallucci & Perugini, 2003; Parks & Rumble, 2001; Sheldon, 1999).

Such accounts are irrelevant to a single-trial social dilemma, in which participants make only one choice. In such contexts, the only basis for choice follows from the present (the immediate present), and not from the past or the future. In one such study, participants made a choice after the other had made a choice (Van Lange, 1999). As noted earlier, this study manipulated information about the other's choice, having participants believe that the other gave away one chip, two chips, or three chips from a total of four chips, which were more valuable to the self than to the other. The participant him- or herself also possessed four chips, which were more valuable to the other than to the self. This situation represents a prisoner's dilemma because giving away chips is costly, but both would be better off to the degree that they exchanged a greater number of chips. Prior to the social dilemma, we assessed participants' social value orientations using a nine-item decomposed game technique (i.e., the triple-dominance measure of social value orien-

tation), to examine whether tendencies toward reciprocity would be more pronounced among prosocials than among individualists and competitors.

The analysis focused on reciprocity choices, giving away exactly the same number of chips as the other had given away. Across the three conditions, prosocials exhibited greater reciprocity (64%) than did individualists (33%) or competitors (17%). In another study, we examined reciprocity in the context of a single-trial social dilemma in which the participant and the other made their choices simultaneously (Van Lange, 1999). Reciprocity choices were operationalized as giving away exactly the same number of chips as they expected the other to give away. In this study, too, prosocials (79.6%) exhibited greater reciprocity than did individualists (58.4%) and competitors (45.4%).

The covariation between cooperation and egalitarianism is also supported in some other research. For example, relative to individualists and competitors, prosocials use and recall decision-making heuristics that focus on enhancement of joint outcomes (e.g., "take a problem-solving approach") and enhancement of equality of outcomes (e.g., "play fair" or "share and share alike"; De Dreu & Boles, 1998). Such findings are also interesting because they indicate that individuals may fairly automatically (i.e., without a lot of thought) attach different meanings to the same situation (cf. Liebrand et al., 1986; Sattler & Kerr, 1991; Van Lange & Kuhlman, 1994).

That prosocials are concerned with enhancing both collective outcomes and equality in outcomes is also demonstrated by recent research on ultimatum bargaining. As noted earlier, it has been argued that in ultimatum bargaining offers tend to be "strategic" in that allocators may offer an equal split of the money to the recipient simply to avoid the recipient's rejection of an unequal offer. Van Dijk and Vermunt (2000), for example, designed an ultimatum game in which bargainers had to divide 100 chips that were worth twice as much to the allocator than to the recipient. In the symmetric information condition the allocators were led to believe the recipient too was informed about this differential value, whereas in the asymmetric information condition allocators were led to believe that the recipient was not informed about differential value. Allocators in the symmetric information condition tended to give the recipient more than half of the chips in order to compensate for the differential value. But allocators in the asymmetric information condition made substantially lower offers, suggesting that allocators exhibit a tendency of self-servingly using informational advantage. That is, because the recipient does not know about the differential value, the allocator can offer to split the number of chips equally—a seemingly fair offer to the recipient—without much fear that the recipient is going to reject the offer. Such tendencies have been interpreted in terms of the strategic use of fairness (e.g., Kagel, Kim, & Moser, 1996; Pillutla & Murnighan, 1995).

Interestingly, a recent study by Van Dijk and colleagues (2004) revealed that only individuals with proself orientations (individualists and competitors) used fairness in a strategic, self-serving manner. In contrast, indi-

viduals with prosocial orientation revealed a "true" preference for an equal distribution of the outcomes. That is, in agreement with the notion that prosocials assign great weight to minimizing differences in outcomes, their findings indicated that even when prosocial allocators thought that the recipient was not aware that chips were worth twice as much to the allocator, they did compensate for the differential value of the chips by offering twice as many chips to the recipient as to themselves. Such findings are not only in line with the integrative model of interpersonal orientations but also suggest that prosocials are genuinely concerned with equality in outcomes.

The integrative model of interpersonal orientation is also supported in research on coalition formation, an area of research that has not been conceptualized in terms of egalitarianism. Specifically, Van Beest, Wilke, and Van Dijk (2003) compared bargaining behavior of prosocials and proselfs in a three-person negotiation setting. In this setting, group members could form two-person coalitions by excluding a third party from the coalition—the excluded party would then yield substantially lower outcomes because it does not benefit from the coalition. Alternatively, group members could form a grand coalition of all three parties, yielding a somewhat lower outcome for each party than in a two-party coalition but yielding equality in outcomes. The results indicated that prosocials were more than proselfs reluctant to excluding another party from a coalition. This tendency to not exclude, and to include all members in distributing the bargaining payoff, once again suggests that prosocials are strongly motivated to obtain equality in outcomes.

Similar conclusions can be reached on the basis of research on social dilemmas. For example, Samuelson (1993) investigated in a resource dilemma how prosocials and proselfs reacted to collective inefficiency and inequality. He investigated people's preferences for structural change when they observed that the common resource became depleted (as compared to efficient use of the resource), and when they observed that some members harvested more than others (as compared to a more equal distribution of harvests). Both dimensions—collective inefficiency and inequality—appeared to be more important to prosocials than to proselfs.

A recent study by Stouten, de Cremer, and Van Dijk (2005) provided further evidence for Samuelson's findings by examining emotional reactions to violations of equality. In this study, participants learned that they were a member of a four-person group, and that their group could obtain a monetary bonus if the combined contributions of the group members would surpass a certain threshold. After participants had decided on their contribution they received (bogus) feedback: They were informed that their group had not been successful, and that the total contribution fell below the threshold needed for provision because one member had violated the equality rule by contributing less than an equal share. After this feedback, however, Stouten and colleagues introduced a manipulation of outcome feedback by informing half of the participants that even though the

contributions fell short, the public good would be provided after all. Thus, for these participants what seemed like failure turned out to be a success after all. For the other half of the participants the negative outcome was not altered.

Interestingly, Stouten and colleagues (2005) found that the emotional reactions of proselfs were less negative and more positive if they learned that the public good was provided after all than if they learned that the outcome remained unchanged. In contrast, information that the public good would be provided after all was not enough for enhancing mood in prosocials. That is, even if the public good was provided after all, prosocials remained angry and unhappy. These findings suggest that prosocials' emotional reactions are deeply affected by violations in equality—the anger and frustration caused by one of the members who did not contribute (and received much greater outcomes than the others) was not resolved by yielding a good result for all four.

Taken together, there is good support for the link between cooperation and egalitarianism. Enhancement of joint outcomes and enhancement of equality tend to go together and are characteristic of how prosocials tend to approach social dilemmas and related situations of interdependence. One might further speculate about the *relative* importance of cooperation and equality. There is some initial evidence suggesting that enhancement of equality is "stronger" than enhancement of joint outcomes (e.g., Eek & Gärling, 2000; Gärling, 1999). For example, Gärling (1999) found that relative to individualists and competitors, prosocials exhibited greater levels of universalism, an attitude closely related to equality and fairness, but no greater levels of benevolence, an attitude closely related to altruism in the model discussed earlier. As noted earlier, it is plausible that in the context of prisoner's dilemmas and related structures, the violation of equality is so strong that mutual noncooperation is preferred to even weak forms of unilateral cooperation (or weak forms of altruism) whereby one behaves—or expects to behave—somewhat more cooperatively than the other. That is, prosocials may behave cooperatively up to the point that it violates equality in outcomes too strongly. Future research could examine how, more precisely, these two orientations work in concert, and whether some of the other orientations may in some ways activate each other.

DETERMINANTS OF INTERPERSONAL ORIENTATIONS

Proposition 4 states that "interpersonal orientations are partially shaped by social interactions—therefore, shaped by the self, the interaction partner, and/or situation."

To most social psychologists this proposition should not come as a surprise in that it adds very little (if anything at all) to what most of us already assume. So, why is the proposition stated at all? The reason is that we want to illustrate "the power of the situation" (the situational view) as well as seek to clarify some issues relevant to "influences" of personal dispositions (the dispositional

view) and the partner's observable tendencies ("the partner view"). These goals are all the more important because, empirically, interpersonal orientations are primarily addressed from the dispositional standpoint (i.e., known as social value orientations). We begin our discussion with the situational view.

In their review of interdependence processes, Rusbult and Van Lange (1996) advance three sources of interpersonal orientation, arguing that interpersonal orientations are manifested in at least three general forms: (1) *interpersonal dispositions*, or person-specific inclinations to respond to particular patterns of interdependence in a specific manner across numerous interaction partners; (2) *relationship-specific motives*, or partner-specific inclinations to respond to particular patterns in a specific manner within the context of a given relationship; and (3) *social norms*, or rule-based inclinations to respond to particular patterns of interdependence in a specific manner, either across numerous interaction partners (e.g., never be the first to "defect") or within the context of a given relationship (e.g., never betray your best friend). Clearly, relationship-specific motives and social norms form an important situational basis of interpersonal orientations. For example, a relationship-specific motive may be derived from commitment to a partner, embodying feelings of attachment, intent to persist, and long-term orientation (Rusbult, Verette, Whitney, Slovik, & Lipkus, 1991; see also Agnew, Van Lange, Rusbult, & Langston, 1998). Commitment is ultimately a product of previous social interaction experiences, and broadly shaped by satisfaction with a relationships, alternatives to a relationship, and investments to a relationship. Importantly, relative to partners to whom we feel not very committed, partners to whom we feel strongly committed are more likely to elicit or activate prosocial orientations (Rusbult et al., 1991; Van Lange, Rusbult, et al., 1997). Similarly, the degree to which social norms activate prosocial versus proself orientations is powerfully linked to differences in situations. In some situations, such norms are very strong and often habituated, whereas in other situations such norms may be less salient or more ambiguous (i.e., when two or more social norms tend to conflict). For example, the "equality norm" is a powerful norm in informal, communal situations, whereas other norms, such as equity (Adams, 1965), might be more important in formal, business-like situations. In yet other situations, it may be a norm to compete, often meaning doing the best one can, as in many games or sports where only one can win.

There is even good reason to believe that prosocial orientations (or proself orientations) are fairly easily activated by relatively subtle situational differences. Slight variations in the degree to which another person is perceived as likable or unlikable, close or not so close, similar or dissimilar might exert considerable influence on the activation of prosocial versus proself orientations (e.g., De Bruin & Van Lange, 2000). Similarly, slight variations in the degree to which some norms are made salient in a given situation might exert considerable influences on the activation of prosocial versus proself orientations. For example, Hertel and Fiedler (1994) found higher levels of

cooperation after a morality prime than after a power prime. Several studies have replicated these findings, while at the same time showing that the effects of priming morality may be especially pronounced for individuals who do not tend to have a stable social value orientation (e.g., Hertel & Fiedler, 1998; Smeesters, Warlop, Van Avermaet, Corneille, & Yzerbyt, 2003). And finally, there is evidence indicating that priming people with "intelligence" may strengthen prosocials' tendencies to cooperate, and—more significantly—strengthen competitors' tendency to take advantage of a partner's cooperation (Utz, Ouwerkerk, & Van Lange, 2004). Thus, there is little doubt that the situation (even subtle situational differences, we believe) may exert powerful influences on the activation of prosocial versus proself orientations.

At the same time, decades of early research on the prisoner's dilemma and related situations revealed a remarkable consistency in individuals' orientations. That is, across various situational manipulations, some individuals tended to behave in a prosocial manner, whereas other individuals tended to behave in a proself manner. These observations inspired several researchers to examine individual differences in interpersonal orientations. Indeed, the important line of research on social value orientation (e.g., Messick & McClintock, 1968) provided the methodological tools for assessing prosocial versus proself orientations. Subsequent research has demonstrated that even brief measures involving allocational choices (such as the nine-item decomposed-game instrument, see Appendix 23.1) are predictive of cooperative and noncooperative behavior in various settings, including two-person prisoners' dilemmas, social dilemmas, resource dilemmas, and actual forms of helping behavior. An example of the latter is that individuals with prosocial orientations are more likely to donate time to the university than do individualists and competitors (McClintock & Allison, 1989). There is also evidence that these differences are linked to motivations for willingness to sacrifice in ongoing relationships (Van Lange, Agnew, Harinck, & Steemers, 1997), and to various forms of prosocial behavior in the context of large communities (e.g., donation to noble causes; Van Lange, Van Vugt, Bekkers, Schuyt, & Schippers, 2005).

Some researchers and theorists might believe that the situational view is inconsistent with the dispositional view, thinking that it is an "either–or" matter. We regard both views as perfectly consistent as well as perfectly complementary, and we believe that theoretical analyses would benefit from taking into account *both* views rather than focusing on either point of view. How so? First, it is the situation that *affords* interpersonal orientations. That is, it is the situation that dictates the relevance of a particular interpersonal orientation and determines which orientations are in conflict with one another. For example, the prisoner's dilemma, especially the single-trial prisoner's dilemma, affords cooperative orientations versus self-interested orientations. A coordination situation, on the other hand, affords none of the orientations outlined in this chapter. Thus, first and foremost, it is important to analyze and define situations in terms of "affordances": What is it that the situation calls for?

Second, within a domain of situations that afford cooperative versus noncooperative orientations (e.g., the so-called mixed-motive situations), the distinction between "strong" and "weak" situations, advanced by Snyder and Ickes (1985; see also Mischel, 1977), becomes important. Strong situations are ones that "provide salient cues to guide behavior and have a fairly high degree of structure and definition, whereas weak situations do not tend to have salient cues to guide behavior and are relatively unstructured and ambiguous" (Snyder & Ickes, 1985, p. 904). Strong situations are the ones in which situational influences should be large, whereas weak situations are the ones in which dispositional influences should be large. Applying these concepts to the domain of mixed-motive situations, it is important to note that by its very structure, mixed-motive situations are almost by definition ambiguous. Indeed, they often represent "dilemmas." Thus, the structure itself, by its affordances, is weak and therefore suggests the importance of dispositional influences. This may explain why the "remarkable consistency in individuals' orientations" in mixed-motive situations should in fact not be all that remarkable. However, even mixed-motive situations have the capacity to become strong. In particular, they may become strong because of relationship-specific motives (e.g., commitment) or because of social norms (which, as suggested earlier, may even be activated through some subtle priming procedures). Under such circumstances, the dispositional influences should be substantially weaker.

The implication for research is that when one compares strong with weak situations, one should obtain statistical interactions of disposition and situation (Magnusson & Endler, 1977), because the influences of dispositions should be greater in weak rather than strong situations. For example, preexisting differences in social value orientation do predict willingness to sacrifice in close relationships when one's commitment to the relationship is relatively weak but fail to predict willingness to sacrifice in close relationships when one's commitment to the relationship is strong (Van Lange, Agnew, et al., 1997). Research by Kramer, McClintock, and Messick (1986) provides another illustration revealing that the effects of social value orientation are especially pronounced when the dilemma reaches the point at which the resources are close to being depleted (or circumstances of scarcity and urgency), and much weaker when the resources seem abundant.

Third, the dispositional view does not hold that there is always a perfect correspondence between orientation and behavior. This applies even to very "weak" situations that by their structures afford cooperative behavior versus noncooperative behavior (i.e., when there is in fact a perfect match between orientations and the situational features). For example, the correspondence between prosocial (vs. proself) orientation and cooperative (vs. noncooperative) behavior need not be perfect, even in a single-trial prisoner's dilemma. Why not? First, it is unlikely that one particular orientation is completely independent of some other orientation. Indeed, Proposition 3 suggests that there is a correspondence between enhancement of joint outcomes and enhancement of equality in outcomes. Second, and perhaps more important, the more accurate characterization of the dispositional view is that people differ in the *probability* with which one or more of the interpersonal orientations will be activated—which we discuss as the next proposition.

Proposition 5 states that interpersonal orientations represent different probabilities with which one or more decision rules (e.g., outcome transformations such as MaxJoint and MinDiff) are activated and used.

It is not uncommon for scientists and laypeople alike to assume (often implicitly, we believe) that a disposition or orientation must translate directly into behavior. Perhaps due to the human need for predictability and control, we parsimoniously tend to believe that "prosocial people behave (almost) always prosocially" just as "competitive people behave (almost) always competitively." Rather than taking a deterministic perspective, a more accurate characterization of the dispositional view is probabilistic, based on the assumption that people differ in the *probability* with which one or more of the interpersonal orientations will be activated. As a metaphor, we prefer to frame this in terms of the *slot-machine model of interpersonal orientations*. But what does it mean—more precisely? We suggest that for relatively stable orientations (as dispositions or as partner-specific orientations) people differ in terms of the percentages of slots that represent prosocial, individualistic, and competitive preferences—just as slot machines represent different frequencies of bananas, lemons, and oranges (so we assume). For example, a prosocial person is a person with a relatively high percentage of prosocial slots (let's say, 70%), and relatively low percentages of individualistic and competitive slots (let's say, 20% and 10%). Similarly, a person with strong attachment for his sibling may have the same distribution of slots when facing dilemmas with his or her sibling. The reverse pattern is likely to be for a competitive person, while an individualistic person may take an intermediate position (with 60% individualistic slots, 20% prosocial slots, and 20% competitive slots).

The slot-machine metaphor of interpersonal orientation is reasonable because people behave in a variety of different interaction situations, even with the same partner. Experience accumulates across interaction situations, which is likely to shape a "probability distribution of interpersonal orientations." Indeed, it would appear to be dysfunctional or maladaptive if people relied on only a single orientation in their interactions with others, even if the situational features are the same. The slot-machine model of interpersonal orientation is also plausible (1) because there is variation in the external (and impersonal) circumstances to which individuals may respond in some way (e.g., the weather, noise), and (2) because there is a fair amount of variation within an individual even on a day-to-day basis, which may also exert influences on the activation of a particular orientation (e.g., differences in mood states, or differences in energy levels on a particular day).

The slot-machine metaphor has important implications. One implication is that the metaphor assumes flexibility and adaptation. If a person were to repeatedly

(and rigidly) adopt the same orientation (irrespective of whether it is prosocial, individualistic, or competitive orientation) across multiple partners, or even to one and the same important partner, the person would be unlikely to adapt to small but important changes in the situation or to small but important changes in the partner's behavior. Indeed, rigidity would probably imply that one does not even notice certain changes in the situations (e.g., new possibilities for effective communication) or changes in the partner's behavior (e.g., increased tendency toward cooperation, increased tendency toward "cheating"). Hence, interpersonal orientations require flexibility to be adaptive—and indeed, if we were to be the slave of a particular orientation, our adaptive quality, and hence survival opportunities, would be very slim.

A second implication of the slot-machine metaphor is that people will have experience with prosocial, individualistic, and competitive "states." This is important because it suggests that people should be able to change perspectives when called for. For example, it has been shown that prosocials are more likely than individualists and competitors to evaluate other's cooperative and noncooperative actions in terms of "good versus bad" associating cooperation with goodness and noncooperation with badness—they adopt readily a morality perspective. Conversely, individualists and competitors are more likely than prosocials to evaluate other's actions in terms of strength and weakness, associating cooperation with weakness and noncooperation with strength—they adopt readily a "might" perspective (Liebrand et al., 1986; Van Lange & Kuhlman, 1994). According to the slot-machine metaphor, people should not find it hard to change perspectives: Prosocials should not find it difficult to adopt a perspective whereby competing is construed as a sign of strength, while competitors should not find it difficult to see that cooperation is often the right (or good) thing to do. People should also adapt by changing perspectives when dealing with their close partner from the perspective they have when dealing with a secondhand car salesman (or at least the stereotype thereof). While it may be seen as immoral to misinform your close partner, it may be seen as fairly "smart" to do so when buying (or selling) a secondhand car.

But is there empirical evidence for the slot-machine model of interpersonal orientation? Although the evidence is very indirect, we think of three complementary sources of empirical support. First, as discussed earlier, relatively subtle cues or associations seem to be able to activate one orientation rather than another. Priming morality, fairness, competence, power, and competition have all been shown to affect behavior in prisoner's dilemmas. Moreover, merely describing a situation as a business transaction may be enough to evoke more self-interested behavior (Batson & Moran, 1999; see also Elliott, Hayward, & Canon, 1998). Interestingly, recent research on social dilemmas has suggested that whether a situation is perceived as a business transaction depends not only on the actual words used to describe the situation but perhaps even on other situational characteristics. For example, Tenbrunsel and Messick (1999) demonstrated that perceptions of social dilemma situations

are strongly affected by the introduction of sanctions on selfish behavior. After introduction of sanctions on selfish behavior, the perception of a social dilemma may shift so that people are more likely to regard their decision as a business-like decision rather than an ethical decision. Hence, sanctions of selfishness may activate individualistic or competitive orientations (see also Gneezy & Rustichini, 2000; Mulder, Van Dijk, De Cremer, & Wilke, 2006). Perhaps, the use of explicit interventions may undermine a more natural tendency ("intrinsic motivation"; Deci & Ryan, 2000) to exhibit cooperation among prosocials—those likely to do so under other circumstances.

A second source of indirect support is that the temporal stability of social value orientation is good but far from excellent. As noted earlier, there are often high levels of intrapersonal stability (and interpersonal variability) within various types of social dilemmas that are partially accounted for by measures of social value orientation. At the same time, while the test–retest reliability of social value orientation (i.e., the nine-item, triple-dominance measure) is generally good, it is not excellent. In a study involving a small sample size, it appeared that 18 of 24 classifiable participants (75%) at time 1 expressed the same interpersonal orientation at time 2 (*Kappa* = .60; Van Lange & Semin-Goossens, 1998). In another study (Van Lange, 1999, study 1), the sample was large, fairly representative of the Dutch adult population, and the time lag between measurement sessions was 19 months. Despite some differences in instructions and procedures between the two measurements, it appeared that 342 of 581 participants (58.8%) expressed the same interpersonal orientation at time 1 and time 2 (*Kappa* = .19). Clearly, the stability of interpersonal orientation is somewhat lower than one would expect from a "stable dispositional" point of view, yet comparable to those found for other individual-difference variables (e.g., adult attachment styles; Shaver & Brennan, 1992). We suggest that temporal states that may be accounted for by variability in day-to-day mood, prior experiences with situations resembling social dilemmas, or other "subtle influences" (e.g., media influences) may determine whether prosocial, individualistic, or competitive orientations are more easily activated.

Third, within the context of specific partners, we tend to see considerable variation in the interpersonal orientations we adopt. Clearly, some key relational constructs, such as commitment and trust, are able to predict various propartner behaviors that align with altruism, cooperation, and fairness (cf. Holmes, 2002; Rusbult & Van Lange, 2003). Yet even when the relational circumstances are ideal, we witness behavior that resembles individualism or even competition. Conversely, even when the relational circumstances are bad, we may witness inherent forms of propartner behavior. And despite decades of research on social dilemmas and the like, no empirical overview can point at one variable that is most certainly going to direct all (or even most) people into making cooperative or noncooperative choices. The only exception may be, as noted earlier, when the partner repeatedly engages in noncooperative behavior (cf. Kelley & Stahelski, 1970).

From a more conceptual perspective, it may be instructive to relate the slot-machine metaphor to two complementary theoretical constructs. First, the metaphor is congruent with the notion of "accessibility." For example, in forming impressions of others, for some people some trait concepts tend to be chronically accessible (e.g., appearance) whereas for other people different trait concepts tend to be chronically accessible (e.g., politeness; Higgins, King, & Maven, 1982). Such accessibility differences may influence impressions of others without people being aware of it. Such effects have also been demonstrated for attitudes, for attributes relevant to people self-definitions, and the like. Similarly, a person with a prosocial orientation is more likely to use and rely on decision rules that dictate "equality in outcomes" or "collective outcomes," and so the person may—often without being aware of it—attend to information regarding equality in outcomes and collective outcomes.

The other concept that is related to the slot-machine metaphor is the notion of *goal activation* and habits. As an interesting case in point, Aarts and Dijksterhuis (2000) have shown that habitual bicycle riders rapidly responded to the word "bicycle" when they had been primed with the goal of traveling to nearby locations (e.g., the university); in contrast, nonhabitual bicycle riders did not. Similarly, competitors' tendencies to compete may be activated by relatively small "cues" in social dilemmas and the like whereas such tendencies may not at all present among prosocials. As this literature of automatic goal activation suggests, such tendencies may be automatic—and beyond any awareness. Hence, some interaction goals ("slots") may be activated in an automatic manner (as well as in a more controlled manner).

Taken together, although the evidence is indirect, they do seem to converge on the point that consistent differences in interpersonal orientations represent consistent differences in the probability with which a particular interpersonal orientation may be activated. The same seems to be true for situational effects, which tend to take strong probabilistic (rather than deterministic) forms. The best oxymoron to capture such effects is perhaps "consistent contingency" whereby probability is influenced by numerous other variables that determine contingency, including randomness.

SUMMARY AND CONCLUDING REMARKS

This chapter discusses five propositions that are relevant to understanding the interaction goals with which people adapt to various situations in which self-interest and collective interest are conflicting. Interpersonal orientations are broadly defined as the set of cognitions, affect, and motivation that underlie interpersonal behavior and social interaction. They reflect *interaction goals* by which people seek to enhance the outcomes for themselves (individualism) as well as enhance the outcomes for other (altruism), enhance joint outcomes (cooperation), enhance equality in outcomes (egalitarianism), enhance relative advantage over others (competition), or minimize outcomes of another person (aggression). We suggest that interpersonal orientations are of broad relevance to diverse social psychological phenomena in that the construct is relevant to the internal processes (cognition and affect) that prepares one for interaction as well as to the internal processes that in many ways summarize the interaction—and that prepares one for the next interaction, either with the same partner or with a third, unrelated person. In the following paragraphs we discuss some further theoretical and evaluative issues relevant to these propositions that we have advanced in this chapter.

We begin by noting that the psychology of interpersonal orientations, while inherently social psychological, cuts across several shifts in the dominant theoretical paradigms in the past as well as integrates several fields of psychology—which is arguably important for any scientific topic to grow, bloom, and progress to yield cumulative knowledge (e.g., Kruglanski, 2006; Mischel, 2004). It is closely connected with almost any interpersonal process that is relevant social interaction. The list is endlessly long and is illustrated by (but by no means limited to) concepts such as altruism, generosity, fairness, equality, cooperation, forgiveness, sacrifice, trust, conflict, aggression, hostility, reactance, competition, suspicion, retaliation, and so on. Most of these topics are essential to understanding relationship processes underlying interactions among kin, friends, close partners, or colleagues, as well as group processes underlying interactions among members of teams, work units, interest groups, and even nations. Also, most of these topics have been studied not only by social psychologists but also by personality psychologists, developmental psychologists, health psychologists, cognitive psychologists, and so on. For example, the topic of forgiveness was originally studied by clinical and health psychologists and recently has become an exceptionally productive area of research within social psychology—and for good reasons, in that forgiveness is a response to an interpersonal offense, with strong implications for future interactions between two persons or two groups. Within social psychology, interpersonal orientations are at the heart of interpersonal and group processes, even though each of the phenomena described above applies (perhaps often with even greater societal relevance) to processes that operate within and between large groups (e.g., within and between nations; e.g., Bornstein, 1992). Many processes that receive considerable attention in contemporary social psychology, such as affect regulation, promotion and prevention foci, and stereotyping, are intimately—and importantly—linked to interpersonal orientations and the situations in which they are relevant. For example, those who cooperate, while discovering that the others did not, need to regulate their anger and frustration in some form; social dilemmas can easily evoke a prevention focus of seeking to reduce loss or a promotion focus of reaching mutual cooperation; and even subtle cues regarding another person's (stereotyped) characteristics can affect cooperation (e.g., De Bruin & Van Lange, 2000; De Dreu, Yzerbyt, & Leyens, 1995).

More generally, by being so closely related to social interaction, the literature of interpersonal orientations may serve as a bridge between (1) micro (or molecular)

approaches, with a strong focus on principles and mechanisms that may account for why individuals function as they do, and (2) macro (or molair) approaches, with a strong focus on principles and mechanisms that may account for why large groups and societies function as they do (for a related argument, see Kelley, 2000; see also Penner, Dovidio, Piliavin, & Schroeder, 2005; Van Lange, 2006). We discuss each in turn.

In support of the *micro* side of interpersonal orientations, we suggest that most if not all of the propositions can be supported by neuropsychological and related psychophysiological research. For example, there is research on the neurological basis of empathy, revealing for humans and monkeys that observing someone else's actions automatically activates neural systems underlying the production of our own actions (for a review, see Blakemore & Decety, 2001). Recently, this line of research has been extended to consider our ability to understand the feelings and sensations of others (i.e., our ability to *empathize*). Such research may well illuminate a neural basis for altruism (Singer & Frith, 2006).

Moreover, recent functional magnetic resonance imaging research reveals that attaining patterns of reciprocal cooperation in social dilemmas activates areas of the brain that are associated with the processing or rewards (Rilling et al., 2002). One may argue that the activation of this neural network helps individuals to resist the temptation to take advantage of the partner's cooperation—but instead to develop patterns of cooperation. Such research is consistent with Propositions 1 and 2, but also with Proposition 3, arguing that enhancing collective outcomes and equality in outcomes tends to go together. It would be interesting to examine whether the activation of cooperative goals tends to enhance the goal of equality and vice versa. On the basis of Proposition 3, and the empirical evidence in support of it, it seems likely that these goals are closely associated, and this association may well be demonstrated in a variety of automatic and controlled cognitive mechanisms (e.g., in implication associations), neuropsychological methods, and the like.

In support of the *macro* side of the interpersonal processes, interpersonal orientations are of great relevance to large groups, even nations. In fact, there is good reason to believe that the distribution of social value orientation is the way it is because of "functionality" at the societal level. Specifically, across a variety of different countries (although mostly "Western" countries), the distribution of prosocials, individualists, and competitors is around a 4:2:1 ratio (see Van Lange, Otten, et al., 1997). According to *frequency dependent selection* explanations, often used in evolutionary theory, such a ratio is unlikely to be random. Rather, the success of one strategy depends on the relative frequency of other strategies in the same population (Maynard Smith, 1982). For example, it is plausible that a stable, high-frequency group of prosocials invites individualists and competitors to develop and grow. In particular, a small percentage of competitors may always be there (and never become extinct) because of the presence of prosocials, who allow a little bit of "exploitation" (e.g., in single interactions), when there is no opportunity for behavioral assimilation (cf.

Kelley & Stahelski, 1970). It is of interest to note that the distribution of 4:2:1 has also been observed in computer simulations (e.g., Lomborg, 1996), which provides further evidence in support of the claim that distributions of social value orientation are as they are because of the functioning at the societal level (see also Ketelaar, 2004).

We also suggest that interpersonal orientations, more broadly, reflect how people interact not only with other individuals in the context of dyads but also with individuals as part of medium-size and large groups. There is a fair amount of evidence that social value orientation is also predictive of cooperative behavior in groups and to large-scale social dilemmas, such as environmental dilemmas (e.g., Parks, 1994; Van Vugt, Van Lange, & Meertens, 1996). Also, relative to individualists and competitors, prosocials are more likely to exhibit anonymous forms of prosocial behavior, for the good of their own society or even other societies (e.g., donations to the third world; Van Lange et al., 2005). Furthermore, although dispositions, such as social value orientation, have received little attention in the context of intergroup relations, there is strong evidence that intergroup interactions tend to be more strongly guided by individualistic and competitive goals than interindividual interactions (Insko & Schopler, 1998).

As alluded to earlier, the psychology of interpersonal orientations is closely connected to evolutionary theorizing. Why are people willing to cooperate at all? Why are they willing to engage in costly acts to benefit others or the group? Why do we do so even with complete strangers with whom there is no future of interdependence? Such issues are relevant to the evolution of cooperation. Interestingly, reciprocity is considered to be the key mechanism through which social interactions evolve, both as a direct mechanism (Axelrod, 1984) and as an indirect mechanism, accounting for cooperative behavior among strangers (Nowak & Sigmund, 1998). Reciprocity in social dilemmas seems to be accounted for by enhancement of *both* joint outcomes and equality (see Proposition 3), and thus it is important to examine the evolution of both mechanisms. Similarly, there is good reason to believe that some forms of altruism and generosity ultimately account for the evolution of cooperation. For example, when there is some uncertainty regarding other's actions (e.g., when there is "noise" so that social signals cannot always be detected), it is arguably adaptive to give the other the benefit of the doubt (and not immediately reciprocate) and behave more cooperatively than the other was believed to do (Van Lange, Ouwerkerk, & Tazelaar, 2002).

We also suggest that Propositions 4 and 5 can make important contributions to the evolutionary theory. First, Proposition 4 is relevant to a taxonomy of interaction situations, and it is clear that the scientific discussion about the evolution of cooperation would benefit from an analysis of situations—after all, social life is not limited to social dilemmas, or even to exchange situations. And even within such situational domains, there are differences in the size of groups, in the availability of information regarding the intentions underlying another's actions, and

the degree to which the situation extends in time. A *taxonomic* approach, such as the one advanced by Kelley and colleagues (2003), is essential to understanding the interpersonal circumstances under which various prosocial orientations versus selfish orientations are adaptive. Proposition 5 emphasizes probability in the ways in which interpersonal orientations are activated—by the self, the partner, or the situation. We suggest that the debate about the evolution of cooperation benefits from the notion that decision rules (or transformations) are activated in a probabilistic manner. After all, such a probabilistic approach provides people with the *flexibility* that is needed to adapt to changes in the partner's actions and needed to adapt to changes in situations.

We wish to close by noting that interpersonal orientations are strongly guided by cognitions and affect—a topic that has not yet received much empirical attention. The theorizing regarding interpersonal orientations is most directly rooted in Kelley and Thibaut's (1978) *transformational analysis*—which assumes that individuals may, depending on their orientations, transform a given situation into "an effective situation" that guides behavior and interactions. Part of such transformation processes are the cognitions and emotions that may help individuals "to make sense" of situations—often in a goal-oriented (yet not necessarily conscious) manner. Social dilemmas, in particular, afford multiple and conflicting cognitions (for many, it is a dilemma), and emotions that may guide behavior, and that summarize interaction outcomes (the reader is reminded of the spontaneous comments by participants, reported by Dawes, 1980). For example, people may interpret social dilemmas in terms of classic dimensions of judgment and impression formation, perceiving it in terms of moral evaluation, strength and weakness, intelligence, and the like (Osgood, Suci, & Tannenbaum, 1957; Rosenberg & Sedlak, 1972). Also, the anticipation of experiencing guilt may prompt prosocials to behave cooperatively, to avoid taking advantage of the other's cooperation, or to avoid being accused of such tendencies (Frank, 1988). It goes without saying that feelings of anger, disappointment, and regret may be experienced when the individual discovers that he or she is the only one who cooperated. Conversely, feelings of interpersonal liking, enjoyment, and gratification, may be experienced when individuals have developed stable patterns of mutual cooperation. And some pleasure (e.g., *schadenfreude*) may be derived from punishing free riders, as well as from observing others engaging in such punitive activities (e.g., Price, Tooby, & Cosmides, 2001; cf. Fehr & Gächter, 2002). Frequently, cognitions, and certainly emotions, are guided by strong norms, which often serve to counteract tempting tendencies toward cheating, deception, and otherwise hurtful forms of "rational" self-interest (Ketelaar, 2004; Mealey, 1995). After having criticized Adam Smith (1757/2000) for his (narrow) view regarding situations, we should give him considerable credit for bringing to our attention the important role of hot cognitions and emotions, in particular "moral sentiments" that help us overcome or resolve social dilemmas that threaten interpersonal, intergroup, and international relations.

APPENDIX 23.1. AN INSTRUMENT TO MEASURE INTERPERSONAL ORIENTATION

In this task we ask you to imagine that you have been randomly paired with another person, whom we will refer to simply as the "other." This other person is someone you do not know and that you will not knowingly meet in the future. Both you and the "other" person will be making choices by circling the letter A, B, or C. Your own choices will produce points for both yourself and the "other" person. Likewise, the other's choice will produce points for him or her and for you. Every point has value: The more points you receive, the better for you, and the more points the "other" receives, the better for him or her.

Here's an example of how this task works:

	A	B	C
You get	500	500	550
Other gets	100	500	300

In this example, if you choose A you would receive 500 points and the other would receive 100 points; if you chose B, you would receive 500 points and the other 500; and if you chose C, you would receive 550 points and the other 300. So, you see that your choice influences both the number of points you receive and the number of points the other receives. Before you begin making choices, please keep in mind that there are no right or wrong answers—choose the option that you, for whatever reason, prefer most. Also, remember that the points have value: The more of them you accumulate the better for you. Likewise, from the "other's" point of view, the more points he or she accumulates, the better for him or her.

For each of the nine choice situations, circle A, B, or C, depending on which column you prefer most:

	A	B	C			A	B	C
(1) You get	480	540	480		(6) You get	500	500	570
Other gets	80	280	480		Other gets	500	500	300
(2) You get	560	500	500		(7) You get	510	560	510
Other gets	300	500	100		Other gets	510	300	110
(3) You get	520	520	580		(8) You get	550	500	500
Other gets	520	120	320		Other gets	300	100	500
(4) You get	500	560	490		(9) You get	480	490	540
Other gets	100	300	490		Other gets	100	490	300
(5) You get	560	500	490					
Other gets	300	500	90					

Note: Participants are classified when they make six or more consistent choices. Prosocial choices are: 1c 2b 3a 4c 5b 6a 7a 8c 9b; individualistic choices are: 1b 2a 3c 4b 5a 6c 7b 8a 9c; and competitive choices are: 1a 2c 3b 4a 5c 6b 7c 8b 9a.

ACKNOWLEDGMENT

NOTE

1. There are several theoretically important issues that are important to the five propositions. For example, one key issue concerns whether we should include orientations other than outcome-distribution orientations in the analysis of interpersonal orientations (e.g., dominance vs. submission). Another key issue concerns the "explanations" for prosocial orientation and prosocial behavior: (1) whether prosocial orientation and behavior may to some degree be mediated by concrete self-rewards (e.g., enhancement of mood and the desire to uphold a moral principle), and (2) whether prosocial orientation and behavior to some degree may reflect the goal of enhancing long-term personal outcomes (i.e., the issue of "time"). Generally, we believe that interpersonal behavior and social interactions are guided by other broad orientations (such as dominance vs. submission), the automatic or controlled pursuit of specific self-benefits, and the broader considerations of long-term self-interest. These are very important theoretical issues but considered beyond the scope of this chapter.

REFERENCES

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures. Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, *78*, 53–63.

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York: Academic Press.

Agnew, C. R., Van Lange, P. A. M., Rusbult, C. E., & Langston, C. A. (1998). Cognitive interdependence: Commitment and the cognitive representation of close relationships. *Journal of Personality and Social Psychology*, *74*, 939–954.

Allison, S. T., & Messick, D. M. (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, *3*, 23–42.

Averill, J. R. (1982). *Anger and aggression: An essay on emotion*. New York: Springer.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169–183). New York: Guilford Press.

Batson, C. D. (1994). Why act for the public good. *Personality and Social Psychology Bulletin*, *20*, 603–610.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 282–316). New York: McGraw-Hill.

Batson, C. D., & Ahmad, N. (2001). Empathy-induced altruism in a prisoner's dilemma. II: What if the target of empathy has defected? *European Journal of Social Psychology*, *31*, 25–36.

Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., & Aldeguer, C. M. R. (1995). Empathy and collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology*, *68*, 619–631.

Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology*, *29*, 909–924.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *National Review of Neuroscience*, *2*, 561–567.

Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.

Bornstein, G. (1992). The free-rider problem in intergroup conflicts over step-level and continuous public goods. *Journal of Personality and Social Psychology*, *62*, 597–606.

Bowlby, J. (1969). *Attachment and loss: Volume 1. Attachment*. New York: Basic Books.

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology*, *50*, 543–549.

Camerer, C., & Thaler, R. H. (1995). Anomalies: Ultimatums, dictators and manners. *Journal of Economic Perspectives*, *9*, 209–219.

Caporael, L. R., Dawes, R. M., Orbell, J. M., & Van De Kragt, A. J. C. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, *12*, 683–739.

Clark, M. S., & Mills, J. (1993). The difference between communal and exchange relationships: What it is and is not. *Personality and Social Psychology Bulletin*, *19*, 684–691.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193.

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, *35*, 1–11.

De Bruin, E. N. M., & Van Lange, P. A. M. (2000). What people look for in others: Influences of the perceiver and the perceived on information selection. *Personality and Social Psychology Bulletin*, *26*, 206–219.

Deci, E. L., & Ryan, R. M. (20000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

De Cremer, D., & Tyler, T. R. (2005). Managing group behavior: The interplay between procedural justice, sense of self, and cooperation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 151–218). San Diego, CA: Academic Press.

De Cremer, D., & Van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology*, *29*, 871–893.

De Cremer, D., & Van Vugt, M. (2002). Intergroup and intragroup aspects of leadership in social dilemmas: A relational model of cooperation. *Journal of Experimental Social Psychology*, *38*, 126–136.

De Dreu, C. K. W., & Boles, T. L. (1998). Share and share alike or winner take all? The influence of social value orientation upon choice and recall of negotiation heuristics. *Organizational Behavior and Human Decision Processes*, *76*, 253–276.

De Dreu, C. K. W., Yzerbyt, V. Y., & Leyens, J. P. (1995). Dillution of stereotype-based cooperation in mixed-motive interdependence. *Journal of Experimental Social Psychology*, *31*, 575–593.

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice. *Journal of Social Issues*, *31*, 137–149.

Eek, D., & Gärling, T. (2000). *Effects of joint outcome, equality, and efficiency on assessments of social value orientations*. Unpublished manuscript, Goteborg University, Sweden.

Elliott, C. S., Hayward, D. M., & Canon, S. (1998). Institutional framing: Some experimental evidence. *Journal of Economic Behavior and Organization*, *35*, 455–464.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*, 137–140

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*, 889–906.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguereotype to laserphoto. *Journal of Personality and Social Psychology*, *63*, 877–889.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions*. New York: Norton.

Gallucci, M., & Perugini, M. (2003). Information seeking and reciprocity: A transformational analysis. *European Journal of Social Psychology*, *33*, 473–495.

Gärling, T. (1999). Value priorities, social value orientations, and cooperation in social dilemmas. *British Journal of Social Psychology*, *38*, 397–408.

Gneezy, U., & Rustichini, A. (2000). A fine is a prize. *Journal of Legal Studies*, *29*, 1–17.

Grzelak, J. L. (1982). Preferences and cognitive processes in interdependence situations: A theoretical analysis of cooperation. In

V. Derlage & J. Grzelak (Eds.), *Cooperation and helping behavior* (pp. 95–122). New York: Academic Press.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum games. *Journal of Economic Behavior and Organization*, *3*, 367–388.

Hertel, G., & Fiedler, K. (1994). Affective and cognitive influences in a social dilemma game. *European Journal of Social Psychology*, *24*, 131–145.

Hertel, G., & Fiedler, K. (1998). Fair and dependent versus egoistic and free effects of semantic and evaluative priming on the "Ring Measure of Social Values." *European Journal of Social Psychology*, *28*, 49–70.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). New York: Academic Press.

Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, *43*, 35–47.

Hobbes, T. (1996). *Leviathan*. Cambridge, UK: Cambridge University Press. (Original work published 1651)

Holmes, J. G. (2002). Interpersonal expectations as the building blocks of social cognition: An interdependence theory perspective. *Personal Relationships*, *9*, 1–26.

Homans, G. C. (1961). *Social behavior: Its elementary forms*. New York: Harcourt, Brace & World.

Iedema, J., & Poppe, M. (1994). Effects of social value orientation on expecting and learning others' orientations. *European Journal of Social Psychology*, *24*, 565–579.

Insko, C. A., & Schopler, J. (1998). Differential distrust of groups and individuals. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior: Toward a closer union* (pp. 75–107). Hillsdale, NJ: Erlbaum.

Jones, E. E., & Thibaut, J. W. (1958). Interaction goals as bases of inference in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 151–178). Stanford, CA: Stanford University Press.

Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior*, *13*, 100–111.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–292.

Karremans, J. C., Van Lange, P. A. M., Ouwerkerk, J. W., & Kluwer, E. S. (2003). When forgiving enhances psychological well-being: The role of interpersonal commitment. *Journal of Personality and Social Psychology*, *84*, 1011–1026.

Kelley, H. H. (2000). The proper study of social psychology. *Social Psychology Quarterly*, *63*, 3–15.

Kelley, H. H., Holmes, J. W., Kerr, N. L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. M. (2003). *An atlas of interpersonal situations*. New York: Cambridge University Press.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*, 66–91.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.

Ketelaar, T. (2004). Ancestral emotions, current decisions: Using evolutionary game theory to explore the role of emotions in decision making. In C. Crawford & C. Salmon (Eds.), *Evolutionary psychology, public policy, and personal decisions* (pp. 145–168). Mahwah, NJ: Erlbaum.

Knight, G. P., & Dubro, A. F. (1984). Cooperative, competitive, and individualistic social values: An individualized regression and clustering approach. *Journal of Personality and Social Psychology*, *46*, 98–105.

Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, *46*, 183–207.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*, 1044–1057.

Kramer, R. M., McClintock, C. G., & Messick, D. M. (1986). Social values and cooperative response to a simulated resource conservation crisis. *Journal of Personality*, *54*, 101–117.

Kruglanski, A. (2006). Theories as bridges. In P. A. M. Van Lange (Ed.), *Bridging social psychology: The benefits of transdisciplinary approaches* (pp. 21–32). Mahwah, NJ: Erlbaum.

Kuhlman, D. M., & Marshello, A. (1975). Individual differences in game motivation as moderators of preprogrammed strategic effects in prisoner's dilemma. *Journal of Personality and Social Psychology*, *32*, 922–931.

Lewin, K. (1935). *A dynamic theory of personality: Selected readings* (D. K. Adams & K. E. Zener, Trans.). New York: McGraw-Hill.

Liebrand, W. B., Jansen, R. W., Rijken, V. M., & Suhre, C. J. (1986). Might over morality: Social values and the perception of other players in experimental games. *Journal of Experimental Social Psychology*, *22*, 203–215.

Liebrand, W. B., & Van Run, G. J. (1985). The effects of social motives on behavior in social dilemmas in two cultures. *Journal of Experimental Social Psychology*, *21*, 86–102.

Lind, E. A. (2001). Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. In J. Greenberg & R. Cropanzano (Eds.), *Advances in organizational justice* (pp. 56–88). Stanford, CA: Stanford University Press.

Lomborg, B. (1996). Nucleus and shield: The evolution of social structure in the iterated prisoner's dilemma. *American Sociological Review*, *61*, 278–307.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. London: Wiley.

MacCrimmon, K. R., & Messick, D. M. (1976). A framework for social motives. *Behavioral Science*, *21*, 86–100.

Magnusson, D., & Endler, N. S. (Eds.). (1977). *Personality at the crossroads: Current issues in interactional psychology*. Hillsdale, NJ: Erlbaum.

Maki, J. E., & McClintock, C. G. (1983). The accuracy of social value prediction: Actor and observer influences. *Journal of Personality and Social Psychology*, *45*, 829–838.

Mansbridge, J. J. (1990). *Beyond self-interest*. Chicago: University of Chicago Press.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.

McClintock, C. G. (1972). Social motivation—A set of propositions. *Behavioral Science*, *17*, 438–454.

McClintock, C. G., & Allison, S. T. (1989). Social value orientation and helping behavior. *Journal of Applied Social Psychology*, *19*, 353–362.

McClintock, C. G., & Moskowitz, J. M. (1976). Children's preference for individualistic, cooperative, and competitive outcomes. *Journal of Personality and Social Psychology*, *34*, 543–555.

McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, *73*, 321–336.

Mealey, L. (1995). The socio-biology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, *18*, 523–599.

Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology*, *21*, 480–500.

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, *4*, 1–25.

Messick, D. M., & Thorngate, W. B. (1967). Relative gain maximization in experimental games. *Journal of Experimental Social Psychology*, *3*, 85–101.

Mikula, G., Petri, B., & Tanzer, N. K. (1990). What people regard as unjust: Types and structures of everyday experiences of injustice. *European Journal of Social Psychology*, *20*, 133–149.

Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology*, *74*, 53–62.

Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology*. Hillsdale, NJ: Erlbaum.

Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, *55*, 1–22.

Mulder, L. B., Van Dijk, E., De Cremer, D., & Wilke, H. A. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, *42*, 147–162.

Murray, S. L., & Holmes, J. G. (1993). Seeing virtues as faults. Negativity and the transformation of interpersonal narratives in close relationships. *Journal of Personality and Social Psychology*, *65*, 707–722.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.

Olweus, D. (1979). Stability of aggression patterns in males: A review. *Psychological Bulletin*, *86*, 852–875.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Parks, C. D. (1994). The predictive ability of social values in resource dilemmas and public goods games. *Personality and Social Psychology Bulletin*, *20*, 431–438.

Parks, C. D., & Rumble, A. C. (2001). Elements of reciprocity and social value orientation. *Personality and Social Psychology Bulletin*, *27*, 1301–1309.

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, *56*, 365–392.

Pillutla, M. M., & Murnighan, J. K. (1995). Being fair or appearing fair: Strategic behavior in ultimatum game bargaining. *Academy of Management Journal*, *38*, 1408–1426.

Poppe, M., & Valkenberg, H. (2003). Effects of gain versus loss and certain versus probable outcomes on social value orientations. *European Journal of Social Psychology*, *33*, 331–337.

Price, M. E., Tooby, J., & Cosmides, L. (2001). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior*, *23*, 203–231.

Ratner, R., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology*, *81*, 5–16.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, *35*, 395–405.

Rosenberg, S., & Sedlak, A. (1972). Structural representations of implicit personality theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 235–297). San Diego, CA: Academic Press.

Rusbult, C. E., & Van Lange, P. A. M. (1996). Interdependence processes. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 564–596). New York: Guilford Press.

Rusbult, C. E., & Van Lange, P. A. M. (2003). Interdependence, interaction, and relationships. *Annual Review of Psychology*, *54*, 351–375.

Rusbult, C. E., Verette, J., Whitney, G. A., Slovik, L. F., & Lipkus, I. (1991). Accommodation processes in close relationships: Theory and preliminary empirical evidence. *Journal of Personality and Social Psychology*, *60*, 53–78.

Samuelson, C. D. (1993). A multiattribute evaluation approach to structural change in resource dilemmas. *Organizational Behavior and Human Decision Processes*, *55*, 298–324.

Sattler, D. N., & Kerr, N. (1991). Might versus morality explored: Motivational and cognitive bases for social motives. *Journal of Personality and Social Psychology*, *60*, 756–765.

Shaver, P. R., & Brennan, K. A. (1992). Attachment styles and the "Big Five" personality traits: Their connection with each other and with romantic relationship outcomes. *Personality and Social Psychology Bulletin*, *18*, 536–545.

Sheldon, K. M. (1999). Learning the lessons of tit-for-tat: Even competitors can get the message. *Journal of Personality and Social Psychology*, *77*, 1245–1253.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1988). *The Robbers Cave experiment: Intergroup conflict and coopera-tion*. Middletown, CT: Wesleyan University Press. (Original work published 1961)

Singer, T., & Frith, C. (2006). The emergence of the "social" in cognitive neuroscience: The study of interacting brains. In P. A. M. Van Lange (Ed.), *Bridging social psychology: The benefits of transdisciplinary approaches* (pp. 97–102). Mahwah, NJ: Erlbaum.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131–142.

Smeesters, D., Warlop, L., Van Avermaet, E., Corneille, O., & Yzerbyt, V. (2003). Do not prime hawks with doves: The interplay of construct activation and consistency of social value orientation on cooperative behavior. *Journal of Personality and Social Psychology*, *84*, 972–987.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. (Reprinted in an edition by E. Cannan, ed., 1904. London: Methuen. Vol. 1)

Smith, A. (2000). *The theory of moral sentiments*. New York: Prometheus Books. (Original work published 1757)

Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 883–947). New York: Random House.

Stouten, J., De Cremer, D., & Van Dijk, E. (2005). All is well that ends well, at least for proselfs: Emotional reactions to equality violation as a function of social value orientation. *European Journal of Social Psychology*, *35*, 767–783.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, *44*, 684–707.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Erlbaum.

Tullock, G. (1976). *The vote motive*. London: Institute for Economic Affairs.

Tyler, T. R., & Degoey, P. (1995). Collective restraint in social dilemmas: Procedural justice and social identification effects on support for authorities. *Journal of Personality and Social Psychology*, *69*, 482–497.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–191). New York: Academic Press.

Tyler, T. R., Rasinski, K., & Spodick, N. (1985). The influence of voice on satisfaction with leaders: Exploring the meaning of process control. *Journal of Personality and Social Psychology*, *48*, 72–81.

Utz, S., Ouwerkerk, J., & Van Lange, P. A. M. (2004). Is cooperation smart or dumb? Effects of priming competence on cooperation. *European Journal of Social Psychology*, *34*, 317–332.

Van Beest, I., Wilke, H. A.., & Van Dijk, E. (2003). The excluded player in coalition formation. *Personality and Social Psychology Bulletin*, *29*, 237–247.

Van den Bos, K., Wilke, H. A., & Lind, E. A. (1998). When do we need procedural fairness? The role of trust in authority. *Journal of Personality and Social Psychology*, *75*, 1449–1458.

Van Dijk, E., De Cremer, D., & Handgraaf, M. (2004). Social value orientations and the strategic use of fairness in ultimatum bargaining. *Journal of Experimental Social Psychology*, *40*, 697–707.

Van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: Sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*, 1–25.

Van Dijk, E., & Wilke, H. (2000). Decision-induced focusing in social dilemmas: Give-some, keep-some, take-some and leave-some dilemmas. *Journal of Personality and Social Psychology*, *78*, 92–104.

Van Goozen, S. H., Frijda, N. H., & Van de Poll, N. E. (1995). Anger and aggression during role playing: Gender differences between hormonally treated male and female transexuals and controls. *Aggressive Behavior*, *21*, 257–273.

Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value ori-

entation. *Journal of Personality and Social Psychology*, 77, 337–349.

Van Lange, P. A. M. (Ed.). (2006). *Bridging social psychology: The benefits of transdisciplinary approaches*. Mahwah, NJ: Erlbaum.

Van Lange, P. A. M., Agnew, C. R., Harinck, F., & Steemers, G. (1997). From game theory to real life: How social value orientation affects willingness to sacrifice in ongoing close relationships? *Journal of Personality and Social Psychology*, 73, 1330–1344.

Van Lange, P. A. M., & Kuhlman, D. M. (1994). Social value orientations and impressions of a partner's honesty and intelligence: A test of the might versus morality effect. *Journal of Personality and Social Psychology*, 67, 126–141.

Van Lange, P. A. M., Otten, W., De Bruin, E. N, & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology*, 73, 733–746.

Van Lange, P. A. M., Ouwerkerk, J. W., & Tazelaar, M. J. A. (2002). How to overcome the detrimental effects of noise in social interaction: The benefits of generosity. *Journal of Personality and Social Psychology*, 82, 768–780.

Van Lange, P. A. M., Rusbult, C. E., Drigotas, S. M., Arriaga, X. B., Witcher, B. S., & Cox, C. L. (1997). Willingness to sacrifice in close relationships. *Journal of Personality and Social Psychology*, 72, 1373–1395.

Van Lange, P. A. M., & Semin-Goossens, A. (1998). The boundaries of reciprocal cooperation. *European Journal of Social Psychology*, 28, 847–854.

Van Lange, P. A. M., Van Vugt, M., Bekkers, R., Schuyt, T., & Schippers, M. (2005). *Ecological support for social value orientation: From games to prosocial behavior in the real world.* Unpublished manuscript, Free University, Amsterdam.

Van Vugt, M., & Van Lange, P. A. M. (2006). Psychological adaptations for prosocial behaviour: The altruism puzzle. In M. Schaller, D. Kenrick, & J. Simpson (Eds.), *Evolution and social psychology* (pp. 237–261). New York: Psychology Press.

Van Vugt, M., Van Lange, P. A. M., & Meertens, R. M. (1996). Commuting by car or public transportation? A social dilemma analysis of travel mode judgements. *European Journal of Social Psychology*, 26, 373–395.

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

PART IV

INTERPERSONAL SYSTEM

CHAPTER 24

Attitude Change

GREGORY R. MAIO
GEOFFREY HADDOCK

Attitudes are important because they involve large parts of our mental life. When we really like something or someone, our relevant feelings, beliefs, and behaviors are all swayed in a positive direction, and it is difficult for people to be "objective" and ignore these tendencies. This potent effect of attitudes is the reason why people often think of attitudes as the principal cause and the principal consequence of important social problems. For instance, if people wish to explain bitter fighting between two ethnic groups (e.g., Shia and Sunni Muslims or Serbs and Croats), the first explanation usually considered is "ethnic hatred," which is simply a negative attitude toward the other ethnic group. Consequently, if people seek a means to reduce the conflict, they will suggest ways to try and ameliorate the bitter intergroup attitudes.

The success of such interventions to change attitudes may vitally depend on developing a thorough understanding of the persuasion process. Attempts at persuading people to like another group may work well on some occasions but many terribly backfire on other occasions. To facilitate successful interventions, persuasion researchers have long searched for the factors that influence attitude change and for models that can explain how these factors are influential. This search has led to a large corpus of data spanning a variety of domains of application, including social groups (e.g., Blacks), policies (e.g., tuition fees), social values (e.g., equality), health-relevant behaviors (e.g., smoking), religious behavior (e.g., attending church or synagogue), environmental behaviors (e.g., recycling), food items (e.g., jams), household utensils (e.g., razors), simple tasks (e.g., choices be-

tween objects), and even humorous stimuli (e.g., cartoons). This chapter attempts to summarize some basic principles of persuasion that operate across these domains, drawing heavily on well-researched models of persuasion.

To begin, the chapter describes the content, structure, and function of attitudes, which helps to lay the foundation for understanding principles of attitude change. We then begin to foreshadow specific principles by reviewing models of attitude change and noting their points of agreement. Next, the chapter considers four basic principles in detail and briefly highlights other potential principles. Finally, we note relevant issues for future research.

ATTITUDE CONTENT, STRUCTURE, AND FUNCTION

The first step to understanding attitude change involves comprehending the content, structure, and function of attitudes. These terms have had varied uses in the past. For example, the phrase "attitude structure" has been used to reflect both the capacity of attitudes to reflect different beliefs and emotions and the capacity to reflect different combinations of positivity and negativity (Eagly & Chaiken, 1993). Similarly, the term "attitude function" is used to denote specifically the needs served by expressing or possessing attitudes, which are often related to the function served by the object of the attitude, but not always (Kruglanski & Stroebe, 2005; Shavitt, 1990). In this chapter, the phrase "attitude content" refers to the ways

in which attitudes may express different psychological constructs, such as beliefs and emotions. The phrase "attitude structure" refers to the ways in which attitudes summarize positivity and negativity toward the attitude object. The phrase "attitude function" refers to the ways in which the attitudes themselves may serve different psychological motivations, such as needs to be correct or to be liked by others. By understanding the content, structure, and function of attitudes, it becomes clearer what attitude change attempts must achieve in order to claim success at eliciting attitude change. So, what contents, structure, and functions must be addressed?

Attitude Content

Research on the content of attitudes has been dominated by two perspectives: the three-component model and the expectancy-value model. According to the three-component model, attitudes express people's feelings, beliefs, and past behaviors regarding the attitude object (Zanna & Rempel, 1988). For example, both authors of this chapter disliked the first two Oscar-winning films in the "Lord of the Rings" series because they thought the films were boring and abhorrent (affective component) and believed that the films were not particularly innovative and overly violent for children (cognitive component). Nonetheless, one of the authors sat through the third film in this series (behavioral component), despite his negative affective and cognitive reactions. The three-component model predicts that his behavior is also an important part of his attitude toward these films because, to some extent, it is difficult to believe that a person completely dislikes something that he or she has endured three times. Through the process of self-perception (Bem, 1972; Olson, 1990, 1992) and cognitive dissonance (Festinger, 1957), people tend to decide that they like something when they can recall doing it often.

This example points out an interesting feature of attitudes. Specifically, it is not necessary that all the attitude components express the same sentiment toward the object of an attitude. People's feelings, beliefs, and behaviors may imply conflicting evaluations of the attitude object, and this situation is often labeled "intercomponent ambivalence" (MacDonald & Zanna, 1998; Maio, Esses, & Bell, 2000). There may also be intracomponent ambivalence such that people possess both positive and negative feelings, beliefs, and behaviors. The amount of evaluative conflict within and between components can vary as a function of the attitude object (Bell, Esses, & Maio, 1996), the attitude holder (Thompson & Zanna, 1995), and the situation (Bell & Esses, 2002). Moreover, the net attitude itself may match or mismatch the evaluations in the components on which it is based (Chaiken, Pomerantz, & Giner-Sorolla, 1995).

Nevertheless, there is also a press toward "synergy" among the attitude components (Eagly & Chaiken, 1993; Zanna & Rempel, 1988). That is, despite the frequent existence of ambivalence, people also tend to have positive attitudes toward an object when their feelings, beliefs, and behaviors all express favorability toward an object. In contrast, people tend to have negative attitudes toward an object when their feelings, beliefs, and behaviors express unfavorability toward the object. This consistency may occur because, in general, internal psychological conflict is aversive to us (Festinger, 1957). Consequently, we seek information that confirms our preexisting beliefs, feelings, and behaviors (Frey, 1986; Frey & Stahlberg, 1987), while also processing new information in a manner that might help to reduce ambivalence (Jonas, Diehl, & Bromer, 1997; Maio, Bell, & Esses, 1996; Maio, Greenland, Bernard, & Esses, 2001).

If a result of these information-processing biases is consistency between the components, it could be argued that we should adopt a model of attitude content that focuses simply on one of the components. Some researchers also argue that a model focusing on one component alone has the advantage of simplicity, because there is a danger in making attitudes too much of an all-encompassing entity. These factors buttress the argument for models of attitudes that focus on the belief component. According to models with this focus, attitudes are based on beliefs that are evaluative in nature, and the most interesting issue is how people combine the beliefs to form an attitude (McGuire, 1960b; Wyer, 1970). This issue is central to the expectancy-value models of attitudes, such as the theory of reasoned action (Ajzen & Fishbein, 1980). This perspective proposes that an attitude is the *sum* of all the evaluative beliefs regarding the attitude object. For instance, if people believe that a new snack product is healthy and good tasting, they should hold a positive attitude toward it. Nevertheless, beliefs are rarely held with perfect certainty. For example, a person may be only 70% certain that the snack is healthy but 100% certain that it is good tasting. According to the expectancy-value model, the effect of each belief is a function of the certainty with which it is held. This notion is expressed using a well-known formula: $A = \Sigma b_i e_i$, where A is the total attitude toward the attitude object, b_i is the subjective belief that the object possesses attribute i (e.g., the probability that the snack is healthy) and e_i is the evaluation of attribute i (e.g., the positive value attached to health promotion).

Research has examined the utility of the three-component and expectancy-value models in different ways but has obtained support for both approaches. With regard to the expectancy-value approach, evidence indicates that people's reports of their own attitudes are at least moderately correlated with the summed products of the attitude-relevant expectancies and values (Bagozzi, 1981; Budd, 1986; van der Pligt & de Vries, 1998), although there have been statistical and methodological criticisms of these findings (Sparks, Hedderley, & Shepherd, 1991). In addition, it is clear that persuasive messages influence evaluative beliefs and that these beliefs influence attitudes when people are motivated and able to process persuasive messages in a systematic manner (Chaiken, Liberman, & Eagly, 1989; Petty & Cacioppo, 1986). With regard to the three-component model, evidence indicates that people's beliefs, feelings, and behaviors toward an attitude object are correlated but distinct (Breckler, 1984; Crites, Fabrigar, & Petty, 1994; Haddock & Zanna, 1998). There is also evidence that attitude-

relevant feelings and beliefs are clustered separately in memory (Trafimow & Sheeran, 1998) and that the affective and cognitive components each account for unique variance in attitudes (Breckler & Wiggins, 1989; Brinol, Petty, & Tormala, 2004; Crites et al., 1994; Trafimow & Sheeran, 1998).

The findings in support of the three-component model could be rebutted by the argument that affective reactions and past behaviors are simply different types of beliefs about the attitude objects. For example, people may form a positive attitude toward a snack that tastes good (an affective response) because, as the expectancy-value model proposes, people *believe* that the bar makes them happy and value their own happiness. Nevertheless, this "affective belief" is made salient only by considering the three-component model. Thus, the three-component model at least spurs the expectancy-value approach to consider different types of beliefs, which must be altered during the persuasion process.

Furthermore, as we shall see, there is evidence that appeals that are based on affective or cognitive information may be differentially effective at changing attitudes that are based primarily on affective information or cognitive information. Specifically, appeals can elicit more attitude change when they contain content that matches the content of the message recipient's attitude than when they contain content that mismatches the content of the attitude. For example, affective interventions influence affectively based attitudes more than cognitively based attitudes (Edwards & von Hippel, 1995; Fabrigar & Petty, 1999; Huskinson & Haddock, 2004). These effects form the basis for the third principle of persuasion that we describe below, while also providing further support for two components of the three-component model.

Attitude Structure

An important issue is how positive and negative evaluations are organized within and between components. The intuitive and most frequently assumed possibility is that the existence of positive feelings, beliefs, and behaviors inhibits the occurrence of negative feelings, beliefs, and behaviors. In other words, the positive and negative elements are stored in memory at opposite ends of a single dimension, and people tend to experience either end of the dimension or a location in between.

This one-dimensional view is opposed by a two-dimensional view, which suggests that positive and negative elements are stored along two separate dimensions (Cacioppo, Gardner, & Berntson, 1997; Thompson, Zanna, & Griffin, 1995). One dimension reflects whether people have few or many positive elements, and the other dimension reflects whether people have few or many negative elements. This view proposes that people can possess any combination of positivity or negativity in their attitudes. Consistent with the one-dimensional view, attitudes may subsume little positivity and high negativity, little negativity and high positivity, or no positivity or negativity (i.e., a neutral position). Inconsistent with the one-dimensional view, attitudes might occasionally subsume high amounts of positivity

and negativity, leading to attitudinal ambivalence. The bidimensional perspective explicitly allows for this ambivalence to occur, whereas the one-dimensional perspective does not. As a result, the bidimensional perspective challenges the meaning of the neutral point in one-dimensional scales for assessing attitudes (Kaplan, 1972). In these scales, a neutral response (e.g., neither like nor dislike) is halfway between the most extreme positive response option (e.g., "extremely likable") and the most extreme negative response option (e.g., "extremely dislikable"). People could choose this option because it is a compromise between many positive and negative elements of their attitude or because they have no positive or negative elements whatsoever.

This failure to distinguish between these two reasons for the neutral selection is consequential, because measures that directly assess ambivalence predict a variety of outcomes. The best known outcome is response polarization (Bell & Esses, 2002; Katz & Hass, 1988; MacDonald & Zanna, 1998): People who are highly ambivalent toward an object are more strongly influenced by features of their environment that make salient its positive or negative attributes, causing them to behave more favorably toward it when the positive elements are salient than when the negative elements are salient; nonambivalent people are less strongly influenced by the acute salience of the positive or negative attributes. More relevant to this chapter, ambivalence predicts attitude change, because ambivalent attitudes tend to elicit more message scrutiny (Jonas et al., 1997; Maio et al., 1996). Hence, ambivalent people exhibit more favorable attitudes in response to strongly argued messages than in response to weakly argued messages, whereas nonambivalent participants show little sensitivity to argument strength. Thus, ambivalence may enhance the motivation and ability to carefully consider the merits of persuasive information, causing more relevant information to be relied on more than less relevant information.

Ambivalence is therefore one of many variables that is relevant to another basic principle of persuasion that we describe next. This principle describes a tendency for the motivation to be correct to elicit greater use of relevant information in the persuasion context. Indeed, there is evidence that people who tend to experience ambivalence are more likely to fear invalidity in their judgments (Thompson & Zanna, 1995). Moreover, it is possible that ambivalence is relevant to the matching principle mentioned previously. That is, ambivalent individuals may be more persuaded by messages that recognize their ambivalence by presenting both sides of the arguments than by one-sided messages, whereas nonambivalent individuals may be more persuaded by one-sided messages than by two-sided messages. At present, however, we know of no evidence directly examining this hypothesis.

Attitude Function

Theorists have proposed that attitudes fulfill a variety of psychological needs or functions (Katz, 1960; Smith, Bruner, & White, 1956). Smith and colleagues' (1956) object-appraisal function is often considered the most ba-

sic motive served by attitudes: All attitudes serve to simplify reasoning and behavior by providing guides for how to interact with (e.g., approach or avoid) an attitude object. This function is served by any attitude that is easy to retrieve from memory. Consider a person whose attitude toward eating chocolate is strong and easy to retrieve. This person should have an easier time deciding whether to purchase a bar of chocolate than a person who has only a weak attitude toward chocolate (Fazio, 2000). Indeed, people spontaneously devote greater attention to objects of their more accessible attitudes (Fazio, Blascovich, & Driscoll, 1992), and they exhibit lower physiological arousal when making decisions about the objects of these attitudes (Blascovich et al., 1993). This spontaneous devotion of attention and low physiological arousal may help to make it easier to make decisions regarding the attitude object.

A second basic function of attitudes may be that they help people to experience emotions (Maio & Esses, 2001). Emotions themselves serve basic psychological needs (Damasio, 1994; Zajonc, 1980), and, as a result, people *seek* emotional experiences, even when the experiences are unpleasant (e.g., scary or sad films). One way of experiencing these emotions is through the possession and expression of strong positive or negative attitudes. For example, people might fulfill their need for affect by developing a strong hatred or affection for members of another ethnic group. Supporting this hypothesis, people who score high on an individual difference measure of the need for affect indicate more extreme attitudes across a variety of topics than people who are low in the need (Maio & Esses, 2001). Also, as we describe later, there is evidence linking this motive to individual differences in attitude structure and attitude change.

Both the object appraisal and the need for affect motivations are purported to influence the tendency to form attitudes in the first place, regardless of whether the attitudes are positive or negative. Other functions are used to predict when attitudes will be positive or negative. For instance, the utilitarian function of attitudes causes people to dislike things that are harmful to them and like things that are beneficial (Herek, 1986; Katz, 1960; Shavitt, 1989); the social-adjustive function causes people to like attitude objects (e.g., fashionable clothes) that are popular among people whom they admire and dislike attitude objects that are unpopular (DeBono, 2000; Smith et al., 1956); the value-expressive function causes people to favor objects that promote their personal values (e.g., the Socialist party), while disliking objects that threaten their values (Katz, 1960; Maio & Olson, 1995a, 1995b, 2000a, 2000b); the ego-defensive function causes people to like objects that promote their self-esteem and reject objects that threaten the self (Katz, 1960).

These functions do not constitute the full list of potential needs served by attitudes or the needs served by attitude objects themselves (Kruglanski & Stroebe, 2005). Additional attitude functions could be derived from findings within and outside social psychological research. Within this research domain, it has been found that people attempt to form attitudes that are consistent with their attitude-relevant beliefs and behaviors, because inconsistencies elicit an aversive arousal or "dissonance" (Festinger, 1957; Harmon-Jones & Mills, 1999). Perhaps, then, the maximization of consistency is a basic function of attitudes (Cialdini, Trost, & Newsom, 1995). A similar case can be made for the idea that attitudes can be used to protect oneself from feelings of mortality (Greenberg, Solomon, & Pyszczynski, 1997). Outside of social psychological research, many basic human motivations have been postulated, including some that are similar to the aforementioned attitude functions (Murray, 1938). For example, Murray suggested that there is a motivation to attain recognition, which is similar to the social-adjustment function. He also proposed other motivations that are not apparent in theories of attitude function, such as the motivation to dominate others. His account is particularly important because the motivations postulated by Murray map on to the so-called five basic dimensions of personality: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness (McCrae & John, 1992; Paunonen & Jackson, 1996; Piedmont, McCrae, & Costa, 1992). One could therefore argue that Murray's motivations should be used to further develop the list of potential and distinct attitude functions.

As with attitude content and structure, there are strong ramifications of attitude function for understanding the process of persuasion. As we describe later, a classic theory of persuasion proposed that messages are successful only insofar as they yield reinforcements for the individual (Hovland, Janis, & Kelley, 1953). Theories of attitude function complemented this view by suggesting that messages would be most effective when they addressed needs that are most important to the message recipients' attitudes—messages that address irrelevant needs would not really be reinforcing to the individual (Katz, 1960; Smith et al., 1956). Thus, similar to attitude content and structure, attitude function may serve as a basis for predicting the effects of messages that vary in their content, again supporting the matching principle.

Integrating the Role of Attitude Content, Structure, and Function

The similar potential roles of attitude content, structure, and function as determinants of the effects of message content illustrate how it would be a mistake to regard them as isolated entities. In a discussion of attitude theory, we have argued that attitude content, structure, and function are like three witches who make a better brew together than alone (Maio & Haddock, 2004). In areas of scientific exploration outside psychology (e.g., chemistry and physics), it is a virtual fact of life that the nature, structure, and function of physical entities are inextricably intertwined and mutually influential. In the attitudes literature, attitude content, structure, and function are intertwined. For example, the affective, cognitive, and behavioral components of attitudes may each express varying levels of positivity or negativity. Each feeling, belief, or behavior could subsume one-dimensional evaluations (i.e., negative to positive), but the resultant attitude

that integrates these may be one- or two-dimensional. Alternatively, it is conceivable that each feeling, belief, and behavior can subsume some degree of positivity *and* negativity (i.e., possess two-dimensionality), as is proposed in some research on emotion (Watson, Wiese, Vaidya, & Tellegen, 1999) and motivated behavior (Elliot & Harackiewicz, 1996). The net attitude might retain this two-dimensionality or not. Indeed, it has been proposed that attitudes may be two-dimensional at deep nonconscious levels of cognitive responding to objects, while being one-dimensional at higher levels that demand integration of the positive and negative elements (Cacioppo et al., 1997).

In addition, although it is possible to partly disentangle the effects of content and function (Maio & Olson, 2000a; Marsh & Julka, 2000; Murray, Haddock, & Zanna, 1996), it is apparent that they are often related. For example, if attitudes toward a brand of automobile are based on a need to conserve fuel, then these attitudes should be based on beliefs about the extent to which the car obtains good fuel economy. Similarly, if attitudes toward a style of clothing fulfill a psychological need to enhance social relations, then these attitudes should be based on beliefs about the extent to which the style is preferred among one's friends. These examples illustrate how attitudes that serve different functions should often differ in the content of the beliefs that support them. An additional consideration is that attitude content, structure, and function may interact (Carstensen, Isaacowitz, & Charles, 1999; Kempf, 1999; Maio & Haddock, 2004). For example, affective reactions may be stronger predictors of attitude when the attitude object has a hedonic purpose than when the object has a utilitarian purpose (Kempf, 1999). Full empirical description of these interactions remains an important topic for future research.

MODELS OF ATTITUDE CHANGE

Now that we have discussed what attitudes are and what they do, the foundation is in place to consider the manner in which attitudes change. Most of the research on this issue has examined attitude change in a somewhat indirect way, by examining the effects of diverse variables on the attitudes that people form about novel issues (e.g., comprehensive exams) or objects (telephone answering machines). This focus on novel attitude objects leaves a lot of research on "attitude change" open to the criticism that it is actually examining "attitude formation" or, at best, "attitude change" for weakly held attitudes. Notwithstanding this concern, the extant research is well summarized by six models that retain an emphasis on understanding attitude change: The Yale model of persuasion (Hovland et al., 1953), the information-processing paradigm (McGuire, 1968), the social judgment model (Sherif, 1980; Sherif & Sherif, 1967), the elaboration likelihood model (Petty & Cacioppo, 1986), the heuristic–systematic model (Chaiken et al., 1989), the unimodel (Kruglanski, Fishbach, Erb, Pierro, & Mannetti, 2004), and the cognition-in-persuasion model (Albarracin, 2002). This section briefly describes each model and

foreshadows the principles of attitude formation and change revealed within them.

Yale Model of Persuasion

After conducting seminal research on the effects of war propaganda in World War II (Hovland, Lumsdaine, & Sheffield, 1949) , Carl Hovland and his colleagues established the Yale Communication Research Program. This program initiated a systematic and empirical dissection of the factors that influence the success of persuasive messages. The program's researchers focused on three factors identified by Lasswell's (1948) famous statement that, to understand the effect of a communication, one must know "*Who* says *what* in which channel to *whom* with *what effect*" (emphasis added; p. 37); in other words, one must know the source of the message, the audience of the message, and the message itself. Their classic monograph, "Communication and Persuasion" (Hovland et al., 1953), was organized around various source, message, and audience variables. Their analysis included effects of source credibility, message threat, and audience personality on the attitudes targeted in the messages and relevant behaviors, while examining different dependent variables across studies (addressing the "what effect" issue in Lasswell's statement).

In addition, the group proposed that successful persuasion requires a sequence of events. Processing begins with attention to the message, leading to comprehension of the message and subsequent acceptance or rejection of it (Hovland et al., 1953). In theory, each persuasion variable could affect each of these processing stages. For example, people who are in a negative mood might be less likely to pay attention to a message, while being too distracted to comprehend it fully and too grumpy to agree with it. Thus, a complete understanding of the effects of a persuasive variable entails analysis of its effects on each processing stage.

Despite the importance of all these stages, the process of message acceptance has garnered the most empirical research from the Yale researchers and investigators of subsequent models. The Yale researchers proposed that message acceptance would be more likely when the audience silently rehearses the message arguments together with their initial attitude and the recommended attitudinal response. In this manner, the message content becomes associated in memory with the new and old attitude. After this double association is in place, the critical issue of interest is whether people choose to accept the new attitude over the old attitude.

To adopt the new attitude, an individual must anticipate receiving some reinforcement or incentive for attitude change. The incentive could be the presentation of a positive reinforcement (i.e., reward) or a negative reinforcement (i.e., the removal of something unpleasant). Either way, the incentive would address an important motive or function that is served by the individual's attitude. For example, a message that induces fear (e.g., about the effects of drunk driving) might motivate people to agree with the message in order to avoid suffering the cited negative consequences (a utilitarian function).

In addition, a persuasive message may heighten people's motivation to agree with others in order to smooth their social relations with them (social adjustment function). Also, agreement with an expert source may be a good means to satisfy the need to form correct and accurate knowledge (knowledge function).

The group conducted many studies that provided some support for their predictions. For example, they found that the inclusion of fear-eliciting stimuli in a message increases acceptance of the message, as long as the stimuli are not so fear inducing that they divert people's attention from the recommendations in the message (Janis & Feshbach, 1953). In addition, participants who scored high on a personality measure of social inadequacy showed more evidence of agreement with persuasive messages (Janis, 1954). Also, Hovland and Weiss (1951) found that participants evaluated articles from high-credibility sources (e.g., a reputable journal) more favorably than articles from low-credibility sources (e.g., a politically biased columnist) across several topics (e.g., atomic submarines and steel shortage). In these studies, the evidence was consistent with the notion that message acceptance is a function of the reinforcements that are elicited by message sources, message content, and the message recipients. It is also plausible that these reinforcements helped to satisfy motivations that were important functions of participants' prior attitudes rather than to address unimportant functions. Insofar as this assumption is correct, the Yale evidence can be construed as partial support for the matching principle mentioned earlier.

Subsequent investigations have greatly expanded the examination of source, message, and audience variables. One prominent area of expansion has focused on the link between the message source and important social groups, such as whether the source is a member of a minority or majority opinion groups. Despite initial classic observations of strong persuasion from members of majority groups (Asch, 1956), there is also evidence that minorities can have a significant impact on attitudes (Moscovici, 1980), and it is now clear that either source can elicit stronger persuasion and that majorities and minorities can trigger the same or different methods of message processing (Baker & Petty, 1994; Kruglanski & Mackie, 1991; Martin, 1998). This evidence fits data indicating that the effects of many persuasion variables (e.g., communicator credibility and communicator attractiveness) and message acceptance can be null or even reversed in different circumstances (Petty & Cacioppo, 1981). Such patterns call into question the Yale group's assumptions about the reinforcement-based learning mechanisms that elicit persuasion and, as described later, led to the creation of models that scrutinized the role of cognitions during message processing.

The Information-Processing Paradigm

One key element of the Yale group's model was the idea that there is a sequence of events from message attention to message comprehension and, finally, message acceptance. Yet, the implications of this sequence were not fully laid out and tested until McGuire's (1968) information-processing model. He suggested six core message processing states, which occur in the following order: presentation of the message, attention to the message, comprehension of the message, yielding to the argument, retention of the changed attitude in memory, and behavior relevant to the attitude. He later proposed that this model could include additional steps, such as two steps between retention of an attitude and behavior: the retrieval of an attitude from memory and the integration of this attitude with other relevant information (e.g., situational demands).

The important point of this sequential analysis is not how many specific stages exist; it is the idea that there are many steps at which the whole process can come to a halt. A behavior change can occur only through the completion of all the prior stages. This task looks challenging when expressed in terms of probabilities. If one assumes that there is a .70 chance of each stage occurring, the odds of passing through all the stages to the decision whether to perform a relevant behavior are $.70^5 = .17$. If you add to this process just two of the other stages indicated by McGuire (e.g., retrieval and integration), then attitude change becomes an even more daunting exercise, $.70^7 = .08$.

The second important feature of McGuire's (1968) model is that variables may have different effects on different stages during message processing. (Later, we will see how more recent models have further incorporated this notion that a single variable can act in different ways during the persuasion process.) This differential influence is most easily viewed by considering McGuire's simplified presentation of his model, which focuses on two stages. The first broad stage is reception, which occurs when people have attended to the message and understood it. The second stage, yielding, involves agreeing with the message content. According to McGuire, the impact of a message should be greatest when the separate probabilities of message reception and yielding are high:

$$p(I) = p(R) \times p(Y)$$

where $p(I)$ is the probability that the message exerts the desired influence, $p(R)$ is the probability of message reception, and $p(Y)$ is the probability of yielding to the message after it reception. An important caveat to this equation is that variables may have different influences on the processing stages, causing one probability to be high while the other is low. For example, intelligent audiences should be more likely to complete message reception than less intelligent audiences, but intelligent audiences might also be less likely to yield to the message after it is understood. Thus, intelligence might increase $p(R)$, but lower $p(Y)$. These opposite effects would lead to a curvilinear effect of audience intelligence. Intelligence would increase the impact of a message as intelligence changes from low to moderate levels, but intelligence would decrease the impact of a message as intelligence changes from moderate to high levels. McGuire also predicted that the precise shape of the curve would vary across contexts. For example, complex

messages increase the importance of reception, because many people will find it difficult to understand the message. For such messages, the positive impact of intelligence on reception may be more important than its negative impact on yielding.

These insights provided a valuable extension of the Yale model's simpler predictions about the effects of persuasion variables on different processing stages. McGuire's (1968) model drew more attention to this idea and provided a method for predicting the outcome of the distinct effects on the various processing stages. Nonetheless, applications of this model have received mixed support. Most tests have examined his specific predictions regarding the effects of audience intelligence and self-esteem on receptiveness to persuasive messages, and a recent meta-analysis of these studies (Rhodes & Wood, 1992) found evidence consistent with the hypothesis that high self-esteem increases reception (by enhancing comprehension) but decreases yielding (by enhancing skepticism). (Insufficient evidence existed to yield strong conclusions about intelligence.) There is also evidence that people can skip processing stages or use them in different order. For example, message recipients sometimes appear to pay little attention to message content and simply yield or not based on some salient aspect of the communicator, such as his or her expertise (Chaiken, 1980; Petty, Cacioppo, & Goldman, 1981). On other occasions, people are aware of and initially reject the conclusion of a message and then give the message greater scrutiny to defend their position (Edwards & Smith, 1996).

Of particular relevance is evidence of an automatic link between the comprehension phase and yielding. Some studies indicate that the mere comprehension of a false statement can be sufficient to elicit instantaneous belief in its validity, and this belief is later corrected only after further deliberate thinking (Gilbert, Tafarodi, & Malone, 1993). The idea is that the initial statements serve as anchors for our judgments, and we adjust belief in them up or down depending on the results of further processing (Tversky & Kahneman, 1974). Abundant social cognition research has made clear that people frequently take shortcuts to conclusions rather than exhaustively considering evidence beforehand (Fiske & Taylor, 1991). In addition, as we describe later, there is evidence that attitudes can be shaped outside conscious awareness, which raises questions about the necessity of many of the processing phases (e.g., attention and comprehension) and the levels of consciousness at which they occur. Thus, other models of persuasion have focused on describing shortcuts that appear to bypass some of the processing stages and on discovering when people use shortcuts instead of the more elaborate, sequential routes.

Social Judgment Model

The social judgment model was among the first to focus on a shortcut to persuasion (Sherif, 1980; Sherif & Sherif, 1967; Sherif, Sherif, & Nebergall, 1965). This model proposed that a quick and automatic perceptual mechanism of assimilation versus contrast affects the manner in which people react to message arguments. In this mechanism, message arguments are automatically compared with recipients' prior attitude, and their prior attitude can encompass a range of positions. This range of positions is tapped by three "latitudes." The latitude of acceptance encompasses the range of positions that a person accepts; the latitude of rejection encompasses the range of positions that a person rejects; the latitude of noncomittment encompasses the range of positions that a person neither accepts nor rejects. These latitudes are important because the social judgment model posits that message positions within the latitude of acceptance are perceived differently from message positions within the latitude of rejection. A message position that falls within the latitude of acceptance seems closer to the message recipient's attitudes than it actually is (assimilation), whereas a message position that falls within the latitude of rejection seems more distant from the recipients' attitude than it actually is (contrast). As a result, the model proposes that message recipients' attitudes move toward the message position when the position is within or close to the audience's latitude of acceptance. In contrast, little attitude change, or even reverse attitude change, can occur when the message position is in the latitude of rejection. In both instances, it is conceivable that the process of attitude change itself is very deliberate and thoughtful, but its direction is determined by the automatic perceptual shift in judgment of the message.

The model also proposes that wide latitudes of rejection and narrow latitudes of acceptance are likely to arise when people perceive the issue to be highly relevant to their self-concept or ego, which includes the cherished aspects of the self (Sherif et al., 1965). This self- or ego involvement (Sherif et al., 1965) has been labeled "value-relevant involvement" by other researchers (Johnson & Eagly, 1989; Maio & Olson, 1995a) because the social judgment model focuses on cherished aspects of the self that closely resemble social values, which are abstract ideals that serve as important guiding principles (Maio & Olson, 1998; Rokeach, 1973; Schwartz, 1992). Many empirical examinations of the model have tested whether higher levels of such ego or value involvement are associated with more message rejection (Hovland, Harvey, & Sherif, 1957; Johnson, Lin, Symons, Campbell, & Ekstein, 1995; Maio & Olson, 1995a; Ostrom & Brock, 1968), and reviews of this research have concluded that message recipients with high involvement exhibit greater resistance to message-based attitude change than those with low involvement (Johnson & Eagly, 1989; Sherif et al., 1965).

It is nonetheless unclear whether these effects occur because of assimilation and contrast mechanisms. Direct tests of the assimilation and contrast mechanisms have failed to support the model (Eagly & Telaak, 1972). It is possible that the effects of involvement occur because ego/value involvement causes message recipients to be more confident and certain about their attitude position (Eagly & Telaak, 1972; Hovland et al., 1957). Alternatively, ego/value involvement may reflect the presence of stronger ties between the recipients' attitudes and

other attitudes or beliefs (Krosnick, Boninger, Chuang, Berent, & Carnot, 1993).

Although these alternate mechanisms may explain the effects of ego/value involvement, it would be a mistake to consider these alternate explanations as evidence that assimilation and contrast play no role in the persuasion context. Assimilation and contrast processes in evaluative processes have been repeatedly demonstrated, particularly in examinations of the processes by which people form judgments of others (Abele & Petzold, 1998; Judd, Kenny, & Krosnick, 1983; Mussweiler, 2003; Schwarz & Bless, 1992; Stapel & Winkielman, 1998). What remains to be seen is how and when these processes occur in a persuasion context. For a message position to be assimilated or contrasted to attitudes, people have to be aware of the position, and this awareness often occurs only after people have at least partly considered the source, content, and context of the message. As a result, people might be influenced by these features (e.g., source attractiveness and logical coherence) prior to having completely inferred the message position. Thus, it is possible that the operation of assimilation and contrast is dependent on other factors in the persuasion context.

In fact, every subsequent model of persuasion has focused on one potentially important feature of the persuasion context: people's cognitive responses to the message content. This trend began with the insight that people might react differently to the same message content and that people's idiosyncratic cognitive responses could predict agreement with the message more strongly than qualities of the message itself (Greenwald, 1968). Many subsequent studies tested this prediction by including measures that asked people to indicate their thoughts about the presented message, after having read it, and the results frequently revealed that people who listed more favorable cognitive responses expressed more agreement with the message (Brock, 1967; Osterhouse & Brock, 1970; Petty, Wells, & Brock, 1976). Other studies have found that a direct manipulation of message-relevant thoughts predicts postmessage attitudes, supporting the notion that these thoughts play a causal role in the persuasion process (Killeya & Johnson, 1998). Thus, although the social judgment model introduced important ideas about the potential operation of quick perceptual processes of assimilation and contrast, later evidence indicated that cognitive responses to persuasive messages may be at least as important.

Elaboration Likelihood Model

The elaboration likelihood (ELM) model built on the previous research by proposing that there are times when attitude change occurs through shortcut, quick routes and times when attitude change occurs through more elaborate, cognitive processing (Petty & Cacioppo, 1986; Petty & Wegener, 1999). The former, *peripheral* route includes a number of potential antecedents of persuasion (Petty & Cacioppo, 1986; Petty & Wegener, 1999). For example, people might agree with a message because the communicator appears trustworthy (Petty et al., 1981). They might also be influenced by simple affec-

tive cues that are present in the message, such as physically attractive models (Chaiken, 1979). In addition, the length of the message and the number of its arguments can determine message acceptance, such that people agree with the longer messages that contain more arguments (Petty & Cacioppo, 1984). The common feature of all the peripheral processes is that people do not need to process the message carefully in order to reach their postmessage attitude. They can base their attitude on an easy-to-recognize aspect of the source, message, or context, without further reflection.

In the more elaborate, *central* route, persuasion occurs through its impact on cognitive responses to the message content (Petty & Cacioppo, 1986; Petty & Wegener, 1999). Strong arguments in the message should elicit favorable cognitive responses, whereas weak arguments in the message should elicit unfavorable cognitive responses. These cognitive responses, in turn, should shape subsequent attitudes, such that the attitudes are more favorable when the cognitive responses are more positive. Moreover, the bases of the postmessage attitude in elaborate cognitive responses should make them stronger than if the attitudes would have been formed through the peripheral route, because associations with cognitive support helps attitudes withstand persuasive attack and predict behavior.

According to this model, people are likely to use the peripheral route when their motivation and ability to process are low, while using the central route when their motivation and ability are high. Many factors can influence motivation and ability. For example, motivation can be high when the message appears highly relevant to personal goals (Petty & Cacioppo, 1984) or the message recipient has a dispositional inclination to engage in and enjoy effortful thinking (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Ability can be high when people possess sufficient knowledge to understand the message (Wood, Kallgren, & Preisler, 1985) and are not suffering from distractions (Petty et al., 1976). Using such factors to manipulate motivation and ability, research has found that manipulations of argument strength have more impact on attitudes when the motivation and ability to process the message are high than when they are low, whereas manipulations of peripheral cues (e.g., source attractiveness) have more impact on attitudes when motivation and ability are low (Chen & Chaiken, 1999; Eagly & Chaiken, 1993; Perloff, 2003; Petty & Wegener, 1998a). These results foreshadow the second principle that we describe later, which refers to people's ability to more greatly weigh highly relevant persuasive information (e.g., cogent arguments) over less relevant persuasive information (e.g., source attractiveness) when motivation and ability to form correct judgments is high. In addition, as the model predicts, postmessage attitudes are more stable and predictive of behavior when they are formed through the central route than when they are formed through the peripheral route (Petty, Haugtvedt, & Smith, 1995).

Although the central route appears to involve more detailed message processing and create stronger attitudes, an interesting issue is whether these attitudes are equally

responsive to weaknesses and strengths in the message. The ELM proposes that people are generally responsive to both the strengths and weaknesses of a message, but people can also weight the strengths or weaknesses more strongly when any of several "biasing" factors are present. For instance, individuals may possess prior knowledge that causes them to endorse either the strengths or weaknesses more strongly (Wood et al., 1985). Thus, although message processing is generally directed at objectively attaining a valid attitude, the ELM acknowledges that processing can be biased by other factors in the persuasion context.

Without doubt the strongest aspect of the ELM is its use of motivation and ability as determinants of the way in which people process persuasive messages. This approach is strong partly because it overlaps with abundant evidence of differences between shallow, quick processing and deeper, elaborate processing in the literature on basic cognitive processes (Craik & Lockhart, 1972; Moscowitz, Skurnik, & Galinsky, 1999). Moreover, like McGuire's information-processing paradigm, the ELM permits multiple roles for variables in the persuasion context. For example, source expertise usually acts as a peripheral cue but can also act as an argument for the validity of a message (Heesacker, Petty, & Cacioppo, 1983). However, it is worth considering that the model itself does not specify which role each persuasion variable should assume, which is an interesting and important issue for attitude change practitioners.

Heuristic–Systematic Model

The heuristic–systematic model (Chaiken et al., 1989; Chen & Chaiken, 1999) shares the ELM's emphasis on motivation and ability as determinants of the depth of message processing. Like the ELM, the heuristic–systematic model proposes that people expend more effort to assess the quality of message arguments when the motivation and ability to process the message are high than when they are low, and this type of processing is labeled "systematic." In addition, the heuristic–systemic model predicts that people will be more likely to utilize simple persuasive cues, or *heuristics*, when the motivation and ability to process a message are low (Chaiken, 1987). Thus, both models share an emphasis on two distinctly different routes to persuasion and are consistent with the notion that people will be more greatly affected by less relevant persuasive information when their motivation and ability to be correct is low rather than high.

Key differences begin in their lower processing routes. The heuristic route does not encompass the range of influences described in the ELM, such as affective mechanisms of association. Instead, the heuristic route includes only simple "if–then" decision guides as determinants of persuasion. For instance, "if a communicator is an expert, then he or she is likely to be correct" is a decision heuristic (or shortcut) that message recipients could use to accept or reject a message without further processing. It is possible that at least some of the processes covered in the ELM could be conceived of as heuristics. For example, some affective mechanisms of influence could be explained by a straightforward heuristic: "if I feel good while listening to a message, then it must be correct." The limits of the heuristic mechanisms in the lower processing route have yet to be fully delineated.

The models also differ in their assumptions about the compatibility between the lower and higher processing routes. The heuristic–systematic model proposes that the systematic route can co-occur with heuristic processing when the systematic route is insufficient to reach a confidently held attitude. In contrast, stronger activity through the ELM's central route is supposed to suppress activity in the peripheral route, although both routes can be used simultaneously at moderate levels of motivation and ability and cues within the peripheral route (e.g., mood) can often act as arguments within the central route (Petty & Wegener, 1998a, 1999). Consistent with the heuristic–systematic model, there is evidence that enhancements of the motivation to process a message also lead to increased use of source credibility information when the message arguments conflict and are incapable of supporting a strong conclusion (Maheswaran & Chaiken, 1994), suggesting that use of cues is not excluded by higher levels of message processing. (Recall, however, that source credibility information could be regarded as a persuasive argument within the central processing route of the ELM, although source credibility is usually treated as a peripheral cue.)

Finally, the models differ in their assumptions about people's goals during message processing. The ELM and the heuristic–systematic model both propose that people focus on attaining a valid and correct attitude, but the heuristic–systematic model proposes that people are also motivated by the desire to project a desirable impression to others and by the desire to protect their values and self-concept. The model proposes that the latter two motives should elicit somewhat more biased processing than the accuracy motive, and this hypothesis has been supported by a meta-analysis of past studies and by several experiments (Johnson & Eagly, 1989; Johnson et al., 1995; Maio & Olson, 1995a). In addition, this multiple motive approach better fits abundant theory and research suggesting that social behavior and judgment in general tend to serve multiple motives (Chaiken, Duckworth, & Darke, 1999; Katz, 1960; Maio & Olson, 2000a; Smith et al., 1956). Awareness of these multiple motives is important because the more specific motive to be correct may elicit the most differential use of relevant and irrelevant persuasive information.

Unimodel

Despite the differences between the ELM and the heuristic–systematic model, they share an emphasis on the notion that message processing is affected by motivation and ability to consider the information that is relevant to the judgment at hand. The unimodel of persuasion (Thompson, Kruglanski, & Spiegel, 2000) shares many of the dual-process models' assumptions about the roles of motivation and ability. As in the dual-process models, the unimodel proposes that increases in motiva-

tion and ability to reach a correct judgment enhance the scrutiny of presented information. With a few possible exceptions (Kruglanski & Thompson, 1999), the factors that affect motivation and ability are similar to those examined in the dual-process models of persuasion (e.g., personal relevance and distraction).

The model's unique attributes stem from its use of lay epistemic theory (Kruglanski, 1989) to add three important stipulations about the processing of information in persuasion contexts. The first stipulation is that *any* information that is relevant to the persuasive conclusion can be considered persuasive evidence (Kruglanski & Thompson, 1999). This relevant evidence can go beyond the message content to include information about the source and context. The unimodel proposes that such information might receive high or low processing and does not distinguish between extramessage information (e.g., source expertise) and message arguments. It indicates that prior evidence of differences in the effects of extramessage cues and message arguments have been confounded by differences in levels of processing difficulty. Specifically, the typical research presents the extramessage information briefly and early in the messages, whereas the message arguments were longer, more complex, and later in the message. As a result, the extramessage cues did not require higher levels of processing motivation and cognitive capacity to affect attitudes, whereas the message arguments did require higher levels of motivation and ability. When extramessage information is allowed to be higher in processing complexity (e.g., by presenting source expertise through a lengthy vita at the end of a message), it can act in the same way as message arguments. In contrast, research on the ELM and the heuristic–systematic model has largely confined source and context information to their lower processing routes, although each model notes exceptions in principle to this confinement (Chaiken et al., 1999; Petty, Wheeler, & Bizer, 1999b).

These exceptions often involve instances wherein the extramessage information is considered to be relevant to the conclusion by the message recipient. For example, the attractiveness of a spokesmodel's hair is relevant to her claims about the shampoo she promotes. This *subjective* judgment of relevance is the unimodel's second unique stipulation and its cornerstone (Kruglanski et al., 2004; Kruglanski & Thompson, 1999). The unimodel proposes that persuasion involves a single process of hypothesis testing and inference, during which people attempt to weight the relevant information more strongly than the less relevant information. Any information that has subjective relevance to the conclusion is considered more deeply when the motivation and ability to process this information is high, which applies also to broader classes of situations wherein people form subjective judgments (e.g., impressions of others and self-assessment). This emphasis on the differential use of relevant and irrelevant information is again supportive of the second principle of persuasion, which we describe next.

The third stipulation is that this deeper consideration of relevant information also depends on the structuring of the message. According to the model, earlier information affects the processing and construal of later information. By virtue of occurring earlier, information that is presented first should be easier to process than information that is late in a message. As a result, manipulations of motivation (e.g., personal relevance and accountability) and ability (e.g., distraction) should have an impact on the use of relevant information more strongly when it is presented later in a message sequence. Several experiments have supported this prediction (Pierro, Mannetti, Kruglanski, & Sleeth-Keppler, 2004).

Cognition-in-Persuasion Model

Similar to the unimodel, the cognition-in-persuasion model (Albarracin, 2002) builds on the assumption that motivation and ability affect message processing, while reintroducing the notion of processing sequence. In this model, message recipients' reaction to a message begins with their interpretation of the message and any other information (e.g., source characteristics and own affect) that is available at the time. Recipients retrieve prior knowledge from memory to help evaluate the message, and then they identify and select information to serve as the basis for their final attitude and subsequent behavior. Any of these processing stages can be disrupted by reductions in the motivation and ability to process the information available. Moreover, unlike earlier stage models (McGuire, 1968), message processing can occasionally bypass early stages, perhaps creating more stable attitude change (Albarracin, 2002).

The separate analysis of information identification, retrieval, and selection has interesting implications for the use of relevant information (e.g., message arguments) and irrelevant information (e.g., mood) in the model (Albarracin & Kumkale, 2003). According to the model, decreases in motivation and ability should gradually decrease the use of message arguments, because people attempt to use them but simply fail to process them fully. In contrast, the impact of irrelevant information depends on it being identified (or noticed) and then selected (or utilized). As a result, there should be a curvilinear impact of irrelevant information on attitude change: Irrelevant information should be used when people are sufficiently motivated and able to identify it but insufficiently motivated and able to go a step further and recognize its irrelevance. There should be no impact of the irrelevant information when motivation and ability are lower (thereby permitting its identification) or higher (thereby permitting recognition of irrelevance) than this moderate level. Of interest, the potential null role of less relevant information at low levels of motivation and ability acts as an important caveat to the notion that this information is used at lower levels of motivation and ability than at higher levels of motivation and ability (our second principle).

In addition, the model emphasizes a potential for effects of conflict within a message, similar to the heuristic–systematic model (Chaiken et al., 1989; Maheswaran & Chaiken, 1994). The cognition-in-persuasion model asserts that the detection of conflict triggers attempts at further analysis (during message

processing) and integration (postmessage processing). This proposal fits some evidence about the effects of conflicting information within a persuasion context (Baker & Petty, 1994; Jonas et al., 1997; Maheswaran & Chaiken, 1994) and is also relevant to the earlier discussion of attitude structure, because conflict within premessage attitudes can increase message scrutiny (Maio et al., 1996).

FOUR PRINCIPLES IN PERSUASION

Having considered the psychological nature of attitudes and models of how they change, we now turn to the challenge of articulating basic principles that are shared across the research and theory that we have identified. In this section, we present more detail about each of the four basic principles that we have highlighted throughout our review. We summarize arguments and evidence regarding these principles, and then, consistent with admonitions that every hypothesis is true in only a restricted set of conditions (McGuire, 1983), we outline potential caveats to each principle.

Extraneous Information Can Influence Persuasion

Although attitudes research has rightly focused on the fine-grained determination of the mechanisms through which variables influence attitudes, this mechanistic approach has often led to a focus on the trees at the expense of the wood. With at least one exception (Cialdini, 1993, 2001), persuasion researchers have underemphasized one of the most interesting and consistent findings in persuasion research: People are often persuaded by factors that they would not see as being logically relevant. The significance of this finding can be appreciated merely by attempting to explain persuasion research to a lay audience. Whenever we describe classic experiments to practitioners and our students, they are continually surprised by the range of *extraneous* variables that influence attitudes and often appear to disbelieve that these variables would actually influence *them*.

Nonetheless, one of the earliest empirical demonstrations of this principle tends to be the easiest for people to accept. Specifically, people can be more amenable to an argument when they are in a good mood than when they are in a bad mood (Razran, 1938, 1940). This effect of mood can be reversed (Wegener, Petty, & Klein, 1994), but this variance does not alter the fact that mood does indeed affect persuadability. Yet, if you were to ask someone whether it is a good thing that her judgment of someone's message is affected by her current mood, she would likely respond with concern. If, for example, a person agreed to buy a product partly because he or she was in a good mood, the person would be giving insufficient weight to factors that are actually far more relevant and important, such as the cost and utility of the product. Thus, even though this result probably confirms human intuition about persuasion, it is disturbing and robust enough to ask whether this finding is part of a larger phenomenon.

The extant literature reveals that mood is, indeed, only one of many extraneous variables that affect persuasion. Another important variable is prior commitment to a relevant attitude or behavior. For example, people may become favorable toward buying a product after being told that the product will feature certain desirable attributes (e.g., a high discount) and then remain favorable even after subsequently learning that the desirable attributes will not occur (Burger & Petty, 1981; Cialdini, 1993). People become somewhat obligated to the source of the message and bound to their initial attitude, even though it is not the attitude they would have formed with the correct information.

Whereas both mood and prior commitment are attributes of the message recipient, other extraneous variables are attributes of the source of a message or the message itself. For example, messages paired with likable (Chaiken, 1980), attractive (Chaiken, 1979), powerful (French & Raven, 1959), famous (Petty, Cacioppo, & Schumann, 1983), and ingroup sources (Whittler & Spira, 2002) tend to elicit more favorable attitudes, although these characteristics can also occasionally cause more negative attitudes (Petty, Fleming, & White, 1999; Roskos-Ewoldsen, Bichsel, & Hoffman, 2002). In addition, somewhat extraneous message variables include rate of speech (Smith & Shaffer, 1995; Woodall & Burgoon, 1984), humor (Conway & Dube, 2002), number of arguments (Petty & Cacioppo, 1984), citations of consensus (Darke et al., 1998), and framing of the attributes (Blanton, Stuart, & VandenEijnden, 2001).

Of course, these audience, source, and message variables can vary in the extent to which they are (1) irrelevant and extraneous to the veracity of the message and (2) cause attitude change. With regard to the degree of relevance to the message's validity, there are occasions when it is reasonable for someone to be more persuaded by an attractive source than by an unattractive source. For example, an advertisement for a brand of shampoo should be very unpersuasive from someone who has dull, lifeless hair, because, if the shampoo were really effective, the person advocating its use should have tried it and found it to be useful (Kahle & Homer, 1985). Similarly, rate of speech might be relevant if it is a useful cue to the spokesperson's credibility and expertise (Smith & Shaffer, 1995) and the framing of messages might be relevant to acute or chronic goals (Rothman & Salovey, 1997; Tykocinski, Higgins, & Chaiken, 1994). There may also be individual differences and cultural differences in what is regarded as relevant or diagnostic information (Cialdini, Wosinska, Barrett, Butner, & Gornik-Durose, 1999; Livingston, 2001).

When the information is believed to be irrelevant, however, its effects may diminish when it is followed by highly relevant information that people are motivated and able to process (Albarracin, 2002; Pierro et al., 2004). One potential reason for this effect is that people tend to accept arguments automatically, even when their validity is weak or obviously false (Gilbert et al., 1993). People who are motivated and able might then attempt to correct for the potential impact of the extraneous information on their attitudes and beliefs, but these at-

tempts at correction should depend on having awareness of the existence of a potential extraneous effect and a theory about its impact (Wegener & Petty, 1997, 2001). To this point, however, research has revealed comparatively little about when, where, and how we manifest awareness of extraneous influences and attempt to correct them. This neglect has occurred partly because we tend to assume that people possess little awareness of the influence of extraneous variables. Yet, people do have lay theories of persuasion that occasionally recognize the extraneous variables, such as mood (Friestad & Wright, 1999; Wilson & Brekke, 1994; Wilson, Houston, & Meyers, 1998). Now that the day-to-day impact of extraneous variables is clear, the next important question for research is what people can do to resist their influence and how lay theories play their part (Sagarin, Cialdini, Rice, & Serna, 2002).

Motivation and Ability to Form a Correct Attitude Increase the Impact of Relevant Information

One of the mechanisms for resisting the impact of extraneous variables involves increasing people's motivation and ability to consider deeply a persuasive message and its implications for their attitudes, as is suggested by most of the models of persuasion that we review. This deeper consideration of a message may cause the relevant information to override the impact of the irrelevant information, particularly when the relevant information is difficult to process (Pierro et al., 2004). In addition, this override effect should arise without any awareness or theories about the impact of the extraneous information, because the effect simply relies on a much greater weighting being accorded to the relevant information. More important, our foregoing review describes how numerous manipulations of motivation and ability increase the impact of relevant information on attitudes and notes that all the contemporary models agree about this potential effect of motivation and ability. Therefore, the tendency for motivation and ability to increase the use of relevant information is a viable candidate for our second principle of persuasion.

To make this principle most defensible, we explicitly focus on the motivation and ability to form a *correct* attitude. This focus on the pursuit of the correct attitude is important because manipulations of motivation can confound this goal with other goals, such as the goal to protect values or self-esteem. For example, it is often assumed that manipulations of the personal relevance of a topic increase the motivation to form a correct attitude, but they may also occasionally prime motives to protect existing values or attitudes (Johnson & Eagly, 1989; Maio & Olson, 1995a). An example is the frequently used topic of comprehensive examinations. Students who believe that the exams will take place during their degree should be more concerned about the issue than students who believe that the exams will not occur during their studies, so the timing of exam implementation is used to manipulate the personal relevance of messages favoring these exams in numerous studies. Nonetheless, a person who is used to thinking in terms of the costs and benefits of such

policies may be more interested in an accurate assessment of the impact of the exams than a person who is used to linking policies with personally cherished values (e.g., freedom and equality). As a result, the high "personal relevance" of taking the exams might motivate the former student to accurately assess messages on the topic, while motivating the latter student to reject the message with little thought because of a perceived threat to values (Maio & Olson, 1995a).

Given the widespread use of personal relevance to manipulate motivation in the persuasion context, it is important to note its ability to activate quite diverse motives. Any need that is strong and salient for the individual could be activated by personal relevance, and research on personality and social processes has identified a plethora of such needs (Murray, 1938). Personal relevance may have a particularly strong effect on some basic self-motives, including needs to protect self-esteem, assess the self accurately, verify self-conceptions, and improve the self (Sedikides & Strube, 1997). These self-motives help to illustrate the potentially diverse effects of personal relevance. In theory, needs for self-assessment and self-improvement should promote more objective processing, while self-verification might promote less effortful scans for confirming information. In contrast, the need to protect self-esteem might facilitate closer scrutiny of a personally relevant message but in a biased fashion. This can happen when, for example, people receive messages about personally relevant healthy behaviors (Liberman & Chaiken, 1992; Sherman, Nelson, & Steele, 2000). Thus, when using personal relevance to manipulate the motivation to form a correct attitude, researchers should ensure that the motives concerned with correctness and improvement are dominant. This goal can be achieved by verifying that effects of the manipulations are mediated by their effects on motives to be accurate, and there are extant measures that might be used as prototypes for the development of such tests (Kruglanski et al., 2000; Kruglanski & Webster, 1996; Neuberg, Judice, & West, 1994). In addition, researchers should be mindful that each motive may be pursued in different ways. For instance, people may attempt to fulfill a need for achievement by seeking to prevent failure or to approach success (Elliot & Harackiewicz, 1996; Higgins, 1997). Matching the method of pursuit of the individual's preferred orientation tends to elicit better success (Higgins & Spiegel, 2004).

A final caveat is that the motivation to be correct can, by itself, cause the use of irrelevant information on occasion. The heuristic–systematic model predicts that people use less relevant persuasive information (e.g., source credibility) when people are highly motivated to form a correct attitude and the only available relevant information is ambiguous or contradictory (Chaiken et al., 1989; Maheswaran & Chaiken, 1994). This effect may occur because the less irrelevant information becomes somewhat relevant in the presence of contradictory information. Indeed, past research showing this effect has focused on a source characteristic—credibility or trustworthiness—that has potential relevance (Maheswaran & Chaiken, 1994). Another possibility is that people in this predica-

ment simply need to base their views on some piece of information—whether it be a random toss of a coin or something with dubious validity. Also, people in this situation may expand their theories of the attitude object to include the information as a plausible guide. The result can be an overinterpretation of irrelevant evidence, at least when people are faced with difficulties identifying and correcting for this evidence (Pelham & Neter, 1995; Tetlock & Boettger, 1989; Tetlock, Lerner, & Boettger, 1996). Thus, increased motivation to be veridical can occasionally increase the use of irrelevant information, in addition to increasing the use of relevant information, and this effect on the use of irrelevant information may be particularly evident when the relevant information is conflicting and difficult to disentangle from the irrelevant information.

Persuasion Is Enabled by Congruence between Message Content and Recipients' Accessible Knowledge and Goals

A long-standing topic in research on persuasion concerns whether the effectiveness of a persuasive appeal is enhanced by matching its arguments to properties of the recipient's attitude. Over 50 years ago, it was suggested that an appeal is more likely to elicit persuasion when it highlights an incentive for attitude change (Hovland et al., 1953). Such an incentive may come from a recipient perceiving a match between the content of an appeal and the content and function of their attitude. If this hypothesis is correct, it may provide a useful third principle of persuasion.

Extant research has focused on matches to salient knowledge (elicited by preexisting attitude content) and salient goals (elicited by preexisting attitude function) separately. Using the multicomponent model of attitude, a number of investigators have considered how single attitudes based on affective or cognitive information might be differentially susceptible to affective and cognitive persuasive appeals. On the whole, this line of research has provided relatively strong evidence for structural matching. One design involved the creation of novel attitudes toward a fictitious beverage named "Power-Plus" (Edwards, 1990). These attitudes were deemed to be affective or cognitive dependent on the order in which participants were exposed to positive affective and cognitive information about Power-Plus. After forming an initial attitude about the product, participants entered the persuasion stage of the experiment, in which they were provided with negative affective and cognitive information. As in the attitude formation stage, the basis of the persuasive appeal was manipulated by varying the order of presentation of the affective or cognitive information, the assumption again being that the type of information presented first would represent the basis of the appeal. Edwards (1990) found strong evidence of a structural matching effect: The affect-based appeal was significantly more effective in changing attitudes that were affect based, while the cognition-based appeal was somewhat (but not significantly) more effective in changing cognition-based attitudes. With few exceptions (Millar &

Millar, 1990), similar patterns have since been obtained using a variety of experimental materials, procedures, and attitude objects (Edwards & von Hippel, 1995; Fabrigar & Petty, 1999).

In addition, recent research has considered whether individual differences in attitude structure elicit matching effects. One line of work has demonstrated that individuals differ in the degree to which their attitudes are guided by affective and cognitive information (Huskinson & Haddock, 2004). That is, some individuals rely primarily on affective information when forming attitudes, whereas others rely primarily on cognitive information. More relevant to structural matching, an affective appeal is more effective in eliciting attitude change among individuals whose attitudes were affect based, whereas a cognitive appeal is slightly more effective in eliciting attitude change among individuals whose attitudes were cognition based (Huskinson & Haddock, 2004). In a similar vein, research has found that individual differences in the need for affect (Maio & Esses, 2001) and need for cognition (Cacioppo et al., 1996) moderate the effectiveness of affect- and cognition-based appeals. Consistent with a structural matching perspective, an affective message was more persuasive among individuals who were high in the need for affect and low in the need for cognition (Maio, Esses, Arnold, & Olson, 2004).

There is also abundant evidence for the importance of matching persuasive content to the goals of the message perceiver. Using an individual difference perspective, research has tested whether individuals high versus low in self-monitoring (Snyder, 1974, 1986) are differentially influenced by appeals that highlight either social-adjustive or value-expressive concerns. A number of experiments demonstrated that high self-monitors (whose attitudes are posited to primarily fulfill a social-adjustive function) are more persuaded by appeals targeting social-adjustive concerns, whereas low self-monitors (whose attitudes are posited to primarily fulfill a value-expressive function) are more persuaded by appeals targeting value-expressive concerns (DeBono, 2000; Prentice, 1987).

Additional evidence supporting the functional matching effect comes from research that has employed paradigms in which an attitude's function is manipulated at either the level of the object or via an experimental induction technique. In one experimental induction, participants were led to believe that their attitudes typically fulfilled either a social-adjustive or value-expressive function (Murray et al., 1996). Subsequently, these participants were more persuaded by an appeal that matched their perceived attitude function. At the level of the attitude object, research carried out by Shavitt and colleagues (Shavitt, 1990; Shavitt & Nelson, 2000) demonstrated that attitude objects that primarily fulfill an instrumental function (e.g., coffee) were more likely to change in response to an instrumental-based appeal (e.g., emphasizing the product's flavor and freshness) than to a social-identity-based appeal (e.g., emphasizing the message recipient's personality and discernment), whereas the attitude objects that primarily fulfill a social-identity function (e.g., greeting cards) were less likely to change

in response to an instrumental-based appeal than to a social-identity-based appeal.

While the existing research provides convergent evidence that persuasion is enhanced by matching an appeal to an attitude's functional or structural basis, it is clear that matching does not always enhance persuasion (Petty, Wheeler, & Bizer, 2000). One reason stems from the ELM, which suggests that the functional matching effect might be attributable to the enhanced scrutiny of matched appeals (Petty & Wegener, 1998b). Interestingly, an explanation based on enhanced scrutiny implies that a matched appeal could produce more or less persuasion, dependent on the strength of the arguments contained within the message. To the extent that the arguments are strong, enhanced scrutiny should produce greater attitude change. However, if the arguments in an appeal are weak, enhanced scrutiny should produce less attitude change. To test this proposed mechanism, Petty and Wegener created strong and weak appeals that matched (or not) the functional basis of an individual's attitude. The results revealed that postmessage attitudes were more strongly affected by argument quality when the message matched the functional basis of the recipient's attitude.

The notion that enhanced scrutiny might account for matching effects is consistent with the previous research, because that research has tended to use arguments that the researchers expected to be cogent. In these instances, scrutiny of the messages causes positive message-relevant thoughts, which should increase agreement with the message (Lavine & Snyder, 1996). Exceptions to this pattern have occurred primarily in instances in which the appeals may have been counterargued with relative ease (Millar & Millar, 1990).

This pattern indicates that matching operates as an *enabler* of persuasion but not as an elicitor of persuasion. Through functional matching, people may be more likely to consider the message content, thereby detecting informational matches and mismatches. Functional matching may elicit this effect by affecting both the motivation and the ability to process message information. Motivation may be augmented by the perception of matching knowledge and goals, which may increase the sense of personal relevance (Petty & Wegener, 1998b; Petty et al., 2000). Ability may be increased by greater comprehension of the message content, which is facilitated because the message is framed in a way that is compatible with the individual's accessible knowledge and goals. As a result of this greater comprehension of message content, people may be better able to detect matches between the message content and their own prior knowledge and beliefs (e.g., when arguments seem "strong") and mismatches between the message content and their prior knowledge and beliefs (e.g., when arguments seem "weak") (Lavine & Snyder, 1996). In other words, matching is akin to pitching a message in the recipient's own language, rather than a foreign language—the common language makes the message more interesting, easy to understand, and compelling.

If matching works in this way, it may also be the case that matching is effective *within* categories of attitude content, structure, and function. DeSteno, Petty, Rucker, Wegener, and Braverman (2004) have obtained interesting evidence consistent with this view. In their research, participants who were induced to feel sad were subsequently more persuaded by arguments that induce sadness than participants who were previously made to feel anger. In contrast, participants who were induced to feel anger were subsequently more persuaded by arguments that induce anger than participants who had been made to experience sadness. These effects are driven by increases in the perceived validity of the arguments when the affect-matching occurs. Nonetheless, this specific within-component matching does not always arise (Fabrigar & Petty, 1999), and the effect uncovered by DeSteno and colleagues is partly dependent on a high motivation to process the messages. Additional research is needed to ascertain whether within- and between-component matching effects are sufficiently similar to conclude that they arise from the same common mechanism, such as a common "language" or frame-of-reference for processing.

Persuasion Can Occur without Awareness

Although the vast majority of persuasion research has considered the degree to which conscious thought processes influence how individuals respond to persuasive appeals, a number of experiments have tested whether attitudes might change in the absence of conscious awareness (Dijksterhuis, Aarts, & Smith, 2005). For instance, experiments using classical conditioning paradigms have revealed small but significant effects of subliminal affective primes on diverse stimuli such as Chinese ideographs (Murphy & Zajonc, 1993), neutral words (De Houwer, Hendrickx, & Baeyens, 1997), and unfamiliar people (Krosnick, Betz, Jussim, & Lynn, 1992). In addition, such effects can be found using both implicit and explicit postconditioning measures of attitude, despite participants' unawareness of the covariation between the conditioned stimulus and the unconditioned stimulus during the conditioning procedure (Olson & Fazio, 2001, 2002). Taken together, this research suggests that attitudes toward novel stimuli can be formed and changed in the absence of participants being aware of the link between an attitude object and evaluative information.

Betsch and colleagues (Betsch, Plessner, & Schallies, 2004; Betsch, Plessner, Schwieren, & Gütig, 2001) have suggested that such associative conditioning effects may occur through a different mechanism than the traditional mechanism of attitude formation and change. Their dual-process *value-account model* postulates that attitudes can be produced by implicit processing (which occurs unintentionally and without awareness) and explicit processing (which occurs intentionally and with awareness). With respect to the implicit formation of attitudes, evaluative information about an attitude is implicitly *added* into a hypothetical memory structure they refer to as a value account. With respect to explicit attitude formation, summary evaluations are deliberately normalized, or *averaged* with, the amount of sampled evaluative information.

To date, a number of studies by Betsch and colleagues (2004) have supported the model. For instance, one experiment used a dual-task procedure in which participants processed television advertisements while at the same time reading aloud share price information scrolled at the bottom of the screen (Betsch et al., 2001). The results indicated that spontaneous attitude judgments toward individual shares reflected the objective sum of their return values, with a perfect rank-order correlation between intuitive attitude judgments and the actual total yield of shares. These differences were found despite participants being unable to recall evaluative information about individual shares. Another experiment found direct evidence supporting the use of a summation mechanism for implicit attitude formation and an averaging mechanism for explicit attitude formation (Betsch et al., 2001).

Recent work has also demonstrated that subliminal priming can, under certain conditions, enhance effects of subsequent persuasive messages. Some research has indicated that subliminal primes can be used to activate goal-relevant cognitions (Fitzsimons & Bargh, 2003), which, when combined with a motivation to pursue the goal, enhances the effectiveness of a relevant persuasive appeal (Dijksterhuis et al., 2005; Strahan, Spencer, & Zanna, 2002). According to this perspective, both the activation of goal-relevant cognitions and the motivation to pursue the goal are necessary to make an appeal especially persuasive. In one study testing this hypothesis, participants who were either thirsty or satiated were subliminally primed with either thirst-related or neutral words. When subsequently offered the opportunity to sample different beverages, thirsty participants who had been primed with thirst-related words drank significantly more than all other participants (Strahan et al., 2002). A subsequent study demonstrated that thirsty participants who had been primed with thirst-related words were significantly more persuaded by an advertisement for a thirst-quenching beverage than by an advertisement for an electrolyte-restoring beverage. Furthermore, participants who received the thirst-related primes were also more likely to select coupons for the thirst-quenching beverage (Strahan et al., 2002). In essence, this research provides further support for the function-matching principle that we described earlier, but using a subliminal activator of attitude function.

An important distinction in this context is between awareness of the variables that instigate attitude formation and change and awareness of the attitude itself. Participants in the studies described earlier exhibited no knowledge of the subliminal primes or conditioning that affected their attitudes but revealed clear awareness of their attitude in explicit measures of the attitudes. These studies thus provide elegant instantiations of a more general limitation of people's introspective abilities: People can often report attitudes without accurately ascertaining why they feel the way that they do (Nisbett & Wilson, 1977). People may even hold incorrect theories about the factors impinging upon their attitude and overcorrect or undercorrect for these factors as a result (Wegener & Petty, 1997; Wilson & Brekke, 1994). Thus, an important question is whether people will on occasion exhibit more veridical knowledge of the bases of their attitudes and also be more able to resist the influence of factors beyond their awareness.

There is evidence that this veridical knowledge and resistance may occur for strongly held attitudes that are based on a large amount of knowledge. Such attitudes tend to be more predictive of behavior and resistant to change (Kallgren & Wood, 1986), and they are less vulnerable to interventions that elicit introspection about the bases of attitudes (Wilson, Dunn, Kraft, & Lisle, 1989; Wilson, Kraft, & Dunn, 1989). Introspection about attitudes typically causes people to access an assortment of inaccurate theories of the bases of the attitudes, which causes them to change (Wilson, Dunn, et al., 1989). This effect disappears when the attitudes are based on extensive knowledge (Maio & Olson, 1998; Wilson, Kraft, et al., 1989), suggesting that some veridical access of the true bases of the attitude may be occurring. Thus, an interesting issue is whether, for such attitudes, subliminal interventions fail to elicit attitude change, because people are fully able to access a set of representations to support their attitude. We anticipate that this moderating effect of attitude strength may be an important caveat to the nonconscious operation of persuasion.

Summary of Principles and Future Directions

All the models of persuasion that we have encountered appear to assume that, regardless of differences in how the effects of extraneous information on attitudes are modeled, such information does exert an important impact. Put simply, all models agree that there are times when a person is influenced by information that has, at best, equivocal relevance to the object of judgment (principle 1). Fortunately, increasing people's motivation and ability to form a correct attitude often attenuates the impact of this information, except perhaps when the relevant information is difficult to interpret and the irrelevant information is difficult to identify (principle 2). Nonetheless, people appear more motivated and able to make such distinctions when there is congruence between their attitude-activated knowledge and goals and the contents of a message (principle 3), and the fact that persuasion can occur without conscious awareness of the persuasive factors (principle 4) provides an interesting potential limitation to this process of relevance detection and correction. Together, these four principles help to summarize significant patterns across the literature.

The need to be concise prevents us from elaborating on other potential principles, but a couple of possibilities merit some mention. First, there is consistent evidence that attitudes are more resistant to change when they are highly accessibility from memory or based in knowledge that is extensive and evaluatively consistent (Bassili, 1996; Chaiken et al., 1995; Wood, Rhodes, & Biek, 1995). These characteristics can be viewed as indicators of attitude strength, which is another variable property of attitudes (Krosnick & Petty, 1995). This convergence tempts the conclusion that "strong attitudes are more resistant to change," but this statement is too circular for inclusion

as a principle of attitude change. In addition, all the traditional indicators of attitude strength are dependent on attitude content, structure, and function, as we outlined earlier. For example, attitudes with more supporting behavioral experiences and structurally consistent content are stronger (Chaiken et al., 1995; Wood et al., 1995), and attitudes that strongly serve a utilitarian function are easier to retrieve from memory (Fazio, 2000). Consistent with other researchers (Krosnick et al., 1993), we expect that there is more to be gained by examining the unique roles and interactions among these facets than by treating them solely as common indicators of a single underlying principle.

We are particularly intrigued by the possibility of delineating a sequential operation in the persuasion process. Ever since the seminal models of persuasion outlined a sequence of processing stages, experiments have attempted to understand attitude change as a sequence of information-presentation and -processing stages, including classic demonstrations and many recent experiments and models (Albarracin, 2002; Pierro et al., 2004). The extant evidence leaves us with little doubt that the sequence of information presentation is important. We would go so far as to suggest a tentative fifth principle: Information presented early in a persuasion sequence affects the construal of later information. This principle emerged with classic research on impression formation (Asch, 1946; Asch & Zukier, 1984; Hamilton & Zanna, 1974) and may be consistent with basic principles in the nature of verbal persuasive dialogue (Rips, Brem, & Bailenson, 1999). Nonetheless, we have not yet seen this potential principle integrated in a consistent manner in research on persuasion, despite some promising beginnings (Kruglanski et al., 2004). A vital issue for such integrations is the role of people's initial attitudes as biasing "information" early in the processing sequence—a possibility that is emphasized in the social judgment, elaboration likelihood, and cognition-in-persuasion models—because it appears that initial attitudes can bias information processing, at least when the attitudes are accompanied by a strong sense of conviction (Edwards & Smith, 1996; Siero & Doosje, 1993) or because of easy retrieval from memory (Houston & Fazio, 1989; Schuette & Fazio, 1995). Thus, this potential principle merits further analysis and evaluation.

In any attempt to extrapolate basic principles from the literature on persuasion, it is inevitable that attention is also drawn to basic issues that have not yet been fully addressed. One important issue is the need for an integrated perspective on how persuasive interventions affect nontargeted attitudes. Fishbein and Ajzen (1981; Ajzen & Fishbein, 1980) described how persuasive arguments might occasionally influence beliefs that were not directly part of the message content. For example, people who learn that a detergent is inexpensive might also come to believe that it is also ineffective, because of a belief that inexpensive items tend to be of less use. Thus, ironically, an argument that seems positive on the surface may have negative implications for the attitude, and it is important to consider the indirect effects of a message through such associations (McGuire, 1960a). The overwhelming majority of past research on persuasion has focused heavily on numerous, diverse *antecedents* (e.g., source characteristics) of changes in the valence of the target attitude. Consideration of the impact of these variables on nontargeted attitudes is an important step toward broadening our examination of the potential consequences of persuasive processes.

An additional approach to broadening the consequences of persuasive processes involves looking at changes in other properties of the targeted and nontargeted attitudes. The ELM partly helps to address this issue because it predicts that deliberate consideration of a persuasive message causes the targeted attitude to become stronger (Petty & Cacioppo, 1986). In other words, the attitudes formed after systematic processing should be more predictive of behavior, stable, and resistant to change. These effects could occur because the post-message attitude is held with more certainty, seen as more important, more accessible from memory, less ambivalent, and so on. Understanding the effects on these attributes is important because it may help to predict the likelihood that effects of a message persist over time. With some notable exceptions (Gruder et al., 1978; Pratkanis, Greenwald, Leippe, & Baumgardner, 1988; Zanna, Fazio, & Ross, 1994), most recent studies have examined attitude change immediately after message presentation, and it is important to develop principles for describing when and why change might persist over a longer period.

It would also be useful to examine the effects of message processing on the functions of attitudes. For example, do messages that mismatch the attitude's pre-existing function diminish the extent to which the attitude subsequently serves that function? Also, in the specific circumstances wherein a mismatching message *is* relatively effective (e.g., when people are motivated to process arguments that are weak) (Petty et al., 2000), does this lead to an attitude that serves multiple functions? If multiple functions do arise, the attitude might potentially become stronger and more predictive of behavior because of its new relevance to motives that may arise across different situations. Such possibilities make clear the importance of moving beyond paradigms that treat attitude valence as the sole (or principal) dependent variable following exposure to a persuasive message.

It is also important to consider a more diverse set of attitude topics in persuasion research. Most highly cited persuasion research focuses on educational policies (e.g., tuition fees and comprehensive exams) or commercial products (telephone answering machines, mobile phones) and is ominously silent about the predictors of attitude change on noncommercial topics. Yet, noncommercial messages are highly visible from many sources: political interest groups continually promote particular political candidates (especially at election time), healthy behaviors (e.g., safe driving and use of contraception), environmentally friendly behaviors (e.g., recycling and nonlittering), religious beliefs (e.g., Church of Latter Day Saints), legislative agendas (e.g., National Rifle Association promotions), and volunteer behaviors (e.g., blood

donation). Marketers have an entire subfield of research examining such campaigns, which they label "social marketing," and one of the guiding clichés in this area is that selling social ideas and behaviors is more difficult than selling commodities (Rothschild, 1979; Wiebe, 1951), perhaps because people are resistant to being persuaded about such self-relevant issues (Cohen, Aronson, & Steele, 2000; Correll, Spencer, & Zanna, 2004). As a result, social marketing may often be ineffective. For example, several studies of antiracism advertising in the United Kingdom have found that the messages may even backfire among those who are ambivalent toward ethnic minorities (Maio, Haddock, Watt, & Hewstone, in press). Similarly, other research has found that common approaches to advocating environmentally friendly behavior in advertising actually encourage less of the behavior, by making the environmentally unfriendly behaviors appear to be common and something that everyone does (Cialdini, 2003). The challenge is for persuasion researchers to collect more data on applications to social issues and to understand factors that predict resistance to persuasion more generally.

CONCLUSION

We have reviewed the literature on persuasion, described six models that attempt to integrate this literature, and extracted four principles that appear evident thus far: (1) attitude change can be elicited by extraneous features of a message or persuasion context; (2) motivation and ability to form a correct attitude increase the impact of relevant information; (3) persuasion is enabled by congruence between the message and accessible knowledge and goals; and (4) persuasion can occur without awareness. These four principles are not the only principles that may be extracted from the research that we have considered, and others may emerge after research addresses the basic and applied issues that we identified (e.g., antecedents of argument strength and effects on nontargeted attitudes).

We are always mindful of the argument that the opposite of every principle might also be true (McGuire, 1983). Nevertheless, the four principles that we have described provide good testimony to the remarkable achievements of persuasion research in the past 50 years. These principles have endured many rigorous tests through this period, and, despite McGuire's sage warning, they may be evident in research for years to come. We hope that our articulation helps to make the principles more salient and leads to research that fully delineates their scope.

REFERENCES

Abele, A. E., & Petzold, P. (1998). Pragatic use of categorical information in impression formation. *Journal of Personality and Social Psychology*, *75*, 347–358.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Albarracin, D. (2002). Cognition in persuasion: An analysis of information processing in response to persuasive communications. *Advances in Experimental Social Psychology*, *34*, 61–130.

Albarracin, D., & Kumkale, G. T. (2003). Affect as information in persuasion: A model of affect identification and discounting. *Journal of Personality and Social Psychology*, *84*, 453–469.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290.

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs*, *70*(9).

Asch, S. E., & Zukier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology*, *46*, 1230–1240.

Bagozzi, R. P. (1981). An examination of the validity of two models of attitude. *Multivariate Behavioral Research*, *16*, 323–359.

Baker, S. M., & Petty, R. E. (1994). Majority and minority influence: Source-position imbalance as a determinant of message scrutiny. *Journal of Personality and Social Psychology*, *67*, 5–19.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*, 637–653.

Bell, D. W., & Esses, V. M. (2002). Ambivalence and response amplification: A motivational perspective. *Personality and Social Psychology Bulletin*, *28*, 1143–1152.

Bell, D. W., Esses, V. M., & Maio, G. R. (1996). The utility of open-ended measures to assess intergroup ambivalence. *Canadian Journal of Behavioural Science*, *28*, 12–18.

Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, *6*, 1–62.

Betsch, T., Plessner, H., & Schallies, E. (2004). The value-account model of attitude formation. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 251–274). New York: Psychology Press.

Betsch, T., Plessner, H., Schwieren, C., & Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin*, *27*, 242–253.

Blanton, H., Stuart, A. E., & VandenEijnden, R. J. (2001). An introduction to deviance-regulation theory: The effect of behavioral norms on message framing. *Personality and Social Psychology Bulletin*, *27*, 848–858.

Blascovich, J., Ernst, J. M., Tomaka, J., Kelsey, R. M., Salomon, K. L., & Fazio, R. H. (1993). Autonomic reactivity as a moderator of autonomic reactivity during decision making. *Journal of Personality and Social Psychology*, *64*, 165–176.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, *47*, 1191–1205.

Breckler, S. J., & Wiggins, E. C. (1989). Affect versus evaluation in the structure of attitudes. *Journal of Experimental Social Psychology*, *25*, 253–271.

Brinol, P., Petty, R. E., & Tormala, Z. L. (2004). Self-validation of cognitive responses to advertisements. *Journal of Consumer Research*, *30*, 559–573.

Brock, T. C. (1967). Communication discrepancy and intent to persuade as determinants of counterargument production. *Journal of Experimental Social Psychology*, *3*, 269–309.

Budd, R. J. (1986). Predicting cigarette use: The need to incorporate measures of salience in the theory of reasoned action. *Journal of Applied Social Psychology*, *16*, 663–685.

Burger, J. M., & Petty, R. E. (1981). The low-ball compliance technique: Task or person commitment? *Journal of Personality and Social Psychology*, *49*, 492–500.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*, 3–25.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*, 165–181.

Chaiken, S. (1979). Communicator physical attractiveness and persuasion. *Journal of Personality and Social Psychology*, *37*, 1387–1397.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752–766.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Mahwah, NJ: Erlbaum.

Chaiken, S., Duckworth, K. L., & Darke, P. (1999). When parsimony fails. . . . *Psychological Inquiry*, *10*, 118–123.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chaiken, S., Pomerantz, E. M., & Giner-Sorolla, R. (1995). Structural consistency and attitude strength. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 387–412). Mahwah, NJ: Erlbaum.

Chen, S., & Chaiken, S. (1999). The heuristic–systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). New York: Guilford Press.

Cialdini, R. B. (1993). *Influence: Science and practice* (3rd ed.). New York: HarperCollins College.

Cialdini, R. B. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, *12*, 105–109.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, *69*, 318–328.

Cialdini, R. B., Wosinska, W., Barrett, D. W., Butner, J., & Gornik-Durose, M. (1999). Compliance with a request in two cultures: The differential impact of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin*, *25*, 1242–1253.

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin*, *26*, 1151–1164.

Conway, M., & Dube, L. (2002). Humor in persuasion on threatening topics: Effectiveness is a function of audience sex role orientation. *Personality and Social Psychology Bulletin*, *28*, 863–873.

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology*, *40*, 350–356.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, *11*, 671–684.

Crites, S. L., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: Conceptual and methodological issues. *Personality and Social Psychology Bulletin*, *20*, 619–634.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Avon.

Darke, P. R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H. P., & Hazlewood, J. D. (1998). Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin*, *24*, 1205–1215.

DeBono, K. G. (2000). Attitude functions and consumer psychology: Understanding perceptions of product quality. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 195–222). Mahwah, NJ: Erlbaum.

De Houwer, J., Hendrickx, H., & Baeyens, F. (1997). Evaluative learning with "subliminally" presented stimuli. *Consciousness and Cognition*, *6*, 87–107.

DeSteno, D., Petty, R. E., Rucker, D. D., Wegener, D. T., & Braverman, J. (2004). Discrete Emotions and Persuasion: The Role of Emotion-Induced Expectancies. *Journal of Personality and Social Psychology*, *86*, 43–56.

Dijksterhuis, A., Aarts, H., & Smith, P. K. (2005). The power of the subliminal: On subliminal persuasion and other potential applications. In R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 77–106). Oxford, UK: Oxford University Press.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Eagly, A. H., & Telaak, K. (1972). Width of the latitude of acceptance as a determinant of attitude change. *Journal of Personality and Social Psychology*, *23*, 388–397.

Edwards, K. (1990). The interplay of affect and cognition in attitude formation and change. *Journal of Personality and Social Psychology*, *59*, 202–216.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, *71*, 5–24.

Edwards, K., & von Hippel, W. (1995). Hearts and minds: The priority of affective and cognitive factors in person perception. *Personality and Social Psychology Bulletin*, *21*, 996–1011.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*, 461–475.

Fabrigar, L. R., & Petty, R. E. (1999). The role of affective and cognitive bases of attitudes in susceptibility to affectively and cognitively based persuasion. *Personality and Social Psychology Bulletin*, *25*, 363–381.

Fazio, R. H. (2000). Accessible attitudes as tools for object appraisal: Their costs and benefits. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 1–36). Mahwah, NJ: Erlbaum.

Fazio, R. H., Blascovich, J., & Driscoll, D. (1992). On the functional value of attitudes: The influence of accessible attitudes upon the ease and quality of decision making. *Personality and Social Psychology Bulletin*, *18*, 388–401.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fishbein, M., & Ajzen, I. (1981). Acceptance, yielding, and impact: Cognitive processes in persuasion. In R. E. Petty, T. M. Ostrom, & T. C. Brock (Eds.), *Cognitive responses in persuasion* (pp. 339–359). Mahwah, NJ: Erlbaum.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fitzsimons, G. M., & Bargh, J. A. (2003). Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology*, *84*, 148–164.

Frey, D. (1986). Recent research on selective exposure to information. *Advances in Experimental Social Psychology*, *19*, 41–80.

Frey, D., & Stahlberg, D. (1987). Selection of information after receiving more or less reliable self-threatening information. *Personality and Social Psychology Bulletin*, *12*, 434–441.

Friestad, M., & Wright, P. (1999). Everyday persuasion knowledge. *Psychology and Marketing*, *16*, 185–194.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, *65*, 221–233.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology*, *29*, 61–139.

Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. A. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147–170). New York: Academic Press.

Gruder, C. L., Cook, T. D., Hennigan, K. M., Flay, B. R., Alessis, C., & Halamaj, J. (1978). Empirical tests of the absolute sleeper effect predicted from the discounting cue hypothesis. *Journal of Personality and Social Psychology*, *36*, 1061–1074.

Haddock, G., & Zanna, M. P. (1998). On the use of open-ended measures to assess attitudinal components. *British Journal of Social Psychology*, *37*, 129–149.

Hamilton, D. L., & Zanna, M. P. (1974). Context effects in impression formation: Changes in connotative meaning. *Journal of Personality and Social Psychology*, *29*, 649–654.

Harmon-Jones, E., & Mills, J. (1999). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 3–21). Washington, DC: American Psychological Association Press.

Heesacker, M. H., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality*, *51*, 653–666.

Herek, G. M. (1986). The instrumentality of attitudes: Toward a neofunctional theory. *Journal of Social Issues*, *42*, 99–114.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T., & Spiegel, S. (2004). Promotion and prevention strategies for self-regulation: A motivated cognition perspective. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 171–187). New York: Guilford Press.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition*, *7*, 51–66.

Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *Journal of Abnormal Social Psychology*, *55*, 244–252.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. New Haven, CT: Yale University Press.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on mass communication*. Princeton, NJ: Princeton University Press.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, *15*, 487–518.

Huskinson, T. L., & Haddock, G. (2004). Individual differences in attitude structure: Variance in the chronic reliance on affective and cognitive information. *Journal of Experimental Social Psychology*, *40*, 82–90.

Janis, I. L. (1954). Personality correlates of susceptibility to persuasion. *Journal of Personality*, *22*, 504–518.

Janis, I. L., & Feshbach, S. (1953). Effects of fear-arousing communications. *Journal of Abnormal Social Psychology*, *48*, 78–92.

Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, *106*, 290–314.

Johnson, B. T., Lin, H., Symons, C. S., Campbell, L. A., & Ekstein, G. (1995). Initial beliefs and attitudinal latitudes as factors in persuasion. *Personality and Social Psychology Bulletin*, *21*, 502–511.

Jonas, K., Diehl, M., & Bromer, P. (1997). Effects of attitudinal ambivalence on information processing and attitude-intention consistency. *Journal of Experimental Social Psychology*, *33*, 190–210.

Judd, C. M., Kenny, D. A., & Krosnick, J. A. (1983). Judging the positions of political candidates: Models of assimilation and contrast. *Journal of Personality and Social Psychology*, *44*, 952–963.

Kahle, L. R., & Homer, P. M. (1985). Physical attractiveness of the celebrity endorser: A social adaptation perspective. *Journal of Consumer Research*, *11*, 954–961.

Kallgren, C. A., & Wood, W. (1986). Access to attitude-relevant information in memory as a determinant of attitude-behavior consistency. *Journal of Experimental Social Psychology*, *22*, 328–338.

Kaplan, K. J. (1972). On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin*, *77*, 361–372.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*, 163–204.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893–905.

Kempf, D. S. (1999). Attitude formation from product trial: Distinct roles of cognition and affect for hedonic and functional products. *Psychology and Marketing*, *16*, 35–50.

Killeya, L. A., & Johnson, B. T. (1998). Experimental induction of biased systematic processing: The directed thought technique. *Personality and Social Psychology Bulletin*, *24*, 17–33.

Krosnick, J. A., Betz, A. L., Jussim, L. J., & Lynn, A. R. (1992). Subliminal conditioning of attitudes. *Personality and Social Psychology Bulletin*, *18*, 152–162.

Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K., & Carnot, C. G. (1993). Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology*, *65*, 1132–1151.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). Mahwah, NJ: Erlbaum.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.

Kruglanski, A. W., Fishbach, A., Erb, H.-P., Pierro, A., & Mannetti, L. (2004). The parametric unimodel as a theory of persuasion. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 399–422). New York: Psychology Press.

Kruglanski, A. W., & Mackie, D. M. (1991). Majority and minority influence: A judgmental process analysis. *European Review of Social Psychology*, *1*, 229–261.

Kruglanski, A. W., & Stroebe, W. (2005). The influence of beliefs and goals on attitudes: Issues of structure, function, and dynamics. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes* (pp. 323–368). Mahwah, NJ: Erlbaum.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Kruglanski, A. W., Thompson, E. P., Higgins, E. T., Atash, M. N., Pierro, A., Shah, J. Y., et al. (2000). To "do the right thing" or to "just do it": Locomotion and assessment as distinct self-regulatory imperatives. *Journal of Personality and Social Psychology*, *79*, 793–815.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263–283.

Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), *The communication of ideas: Religion and civilization series* (pp. 37–51). New York: Harper & Row.

Lavine, H., & Snyder, M. (1996). Cognitive processing and the functional matching effect in persuasion: The mediating role of subjective perceptions of message quality. *Journal of Experimental Social Psychology*, *32*, 580–604.

Liberman, A., & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin*, *18*, 669–679.

Livingston, R. W. (2001). What you see is what you get: Systematic variability in perceptual-based social judgment. *Personality and Social Psychology Bulletin*, *27*, 1086–1096.

MacDonald, T. K., & Zanna, M. P. (1998). Cross-dimension ambivalence toward social groups: Can ambivalence affect intentions to hire feminists? *Personality and Social Psychology Bulletin*, *24*, 427–441.

Maheswaran, D., & Chaiken, S. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, *66*, 460–473.

Maio, G. R., Bell, D. W., & Esses, V. M. (1996). Ambivalence and persuasion: The processing of messages about immigrant groups. *Journal of Experimental Social Psychology*, *32*, 513–536.

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual

differences in the motivation to approach or avoid emotions. *Journal of Personality*, *69*, 583–616.

Maio, G. R., Esses, V. M., Arnold, K., & Olson, J. M. (2004). The function-structure model of attitudes: Incorporating the need for affect. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitude* (pp. 7–34). New York: Psychology Press.

Maio, G. R., Esses, V. M., & Bell, D. W. (2000). Examining conflict between components of attitudes: Ambivalence and inconsistency are distinct constructs. *Canadian Journal of Behavioural Science*, *32*(1), 58–70.

Maio, G. R., Greenland, K., Bernard, M. M., & Esses, V. M. (2001). Effects of intergroup ambivalence on information processing: The role of physiological arousal. *Group Processes and Intergroup Relations*, *4*, 355–372.

Maio, G. R., & Haddock, G. (2004). Theories of attitude: Creating a witches' brew. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes* (pp. 425–453). New York: Psychology Press.

Maio, G. R., Haddock, G. G., Watt, S. E., & Hewstone, M. (in press). Implicit measures and applied contexts: An illustrative examination of anti-racism advertising. In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new wave of implicit measures*.

Maio, G. R., & Olson, J. M. (1995a). Involvement and persuasion: Evidence for different types of involvement. *Canadian Journal of Behavioral Science*, *27*, 64–78.

Maio, G. R., & Olson, J. M. (1995b). Relations between values, attitudes, and behavioral intentions: The moderating role of attitude function. *Journal of Experimental Social Psychology*, *31*, 266–285.

Maio, G. R., & Olson, J. M. (1998). Values as truisms: Evidence and implications. *Journal of Personality and Social Psychology*, *74*, 294–311.

Maio, G. R., & Olson, J. M. (2000a). Emergent themes and potential approaches to attitude function: The function–structure model of attitudes. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 417–442). Mahwah, NJ: Erlbaum.

Maio, G. R., & Olson, J. M. (2000b). What is a "value-expressive" attitude? In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 249–269). Mahwah, NJ: Erlbaum.

Marsh, K. L., & Julka, D. L. (2000). A motivational approach to experimental tests of attitude functions theory. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitude* (pp. 271–294). Mahwah, NJ: Erlbaum.

Martin, R. (1998). Majority and minority influence using the afterimage paradigm: A series of attempted replications. *Journal of Experimental Social Psychology*, *34*, 1–26.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*, 175–215.

McGuire, W. J. (1960a). Direct and indirect persuasive effects of dissonance-producing messages. *Journal of Abnormal and Social Psychology*, *60*, 354–358.

McGuire, W. J. (1960b). A syllogistic analysis of cognitive relationships. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change: An analysis of consistency among attitude components*. New Haven, CT: Yale University Press.

McGuire, W. J. (1968). Personality and attitude change: An information-processing theory. In A. G. Greenwald, T. C. Brock, & T. A. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 171–196). San Diego, CA: Academic Press.

McGuire, W. J. (1983). A contextualist theory of knowledge: Its implications for innovation and reform in psychological research. *Advances in Experimental Social Psychology*, *16*, 1–47.

Millar, M. G., & Millar, K. U. (1990). Attitude change as a function of attitude type and argument type. *Journal of Personality and Social Psychology*, *59*, 217–228.

Moscovici, S. (1980). Toward a theory of conversion behavior. *Advances in Experimental Social Psychology*, *13*, 209–239.

Moscowitz, G. B., Skurnik, I., & Galinsky, A. D. (1999). The history of dual-process notions, and the future of preconscious control. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 12–36). New York: Guilford Press.

Murphy, S., & Zajonc, R. B. (1993). Affect. cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*, 723–739.

Murray, H. (1938). *Explorations in personality*. New York: Oxford University Press.

Murray, S. L., Haddock, G., & Zanna, M. P. (1996). On creating value-expressive attitudes: An experimental approach. In C. Seligman, J. M. Olson, & M. P. Zanna (Eds.), *The psychology of values: The Ontario symposium* (Vol. 8, pp. 107–133). Mahwah, NJ: Erlbaum.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489.

Neuberg, S. L., Judice, T. N., & West, S. G. (1994). What the need for closure scale measures and what it does not: Toward differentiating among related epistemic motives. *Journal of Personality and Social Psychology*, *72*, 1396–1412.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal report on mental processes. *Psychological Review*, *84*, 231–259.

Olson, J. M. (1990). Self-inferences processes in emotion. In J. M. Olson & M. P. Zanna (Eds.), *Self-inference processes: The Ontario Symposium* (Vol. 6, pp. 17–41). Mahwah, NJ: Erlbaum.

Olson, J. M. (1992). Self-perception of humor: Evidence for discounting and augmentation effects. *Journal of Personality and Social Psychology*, *62*, 369–377.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, *12*, 413–417.

Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition*, *20*, 89–104.

Osterhouse, R. A., & Brock, T. C. (1970). Distraction increases yielding to propaganda by inhibiting counterarguing. *Journal of Personality and Social Psychology*, *15*, 344–358.

Ostrom, T. M., & Brock, T. C. (1968). A cognitive model of attitudinal involvement. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 373–383). Chicago: Rand-McNally.

Paunonen, S. V., & Jackson, D. N. (1996). The Jackson Personality Inventory and the five-factor model of personality. *Journal of Research in Personality*, *30*, 42–59.

Pelham, B. W., & Neter, E. (1995). The effect of motivation on judgment depends on the difficulty of the judgment. *Journal of Personality and Social Psychology*, *68*, 581–594.

Perloff, R. M. (2003). *The dynamics of persuasion: Communication and attitudes in the 21st century* (2nd ed.). Mahwah, NJ: Erlbaum.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Brown.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, *46*, 69–81.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*, 847–855.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*, 135–146.

Petty, R. E., Fleming, M. A., & White, P. H. (1999). Stigmatized sources and persuasion: Prejudice as a determinant of argument scrutiny. *Journal of Personality and Social Psychology*, *76*, 19–34.

Petty, R. E., Haugtvedt, C. P., & Smith, S. M. (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are

persistent, resistant, and predictive of behavior. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 93–130). Mahwah, NJ: Erlbaum.

Petty, R. E., & Wegener, D. T. (1998a). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 323–390). Boston: McGraw-Hill.

Petty, R. E., & Wegener, D. T. (1998b). Matching versus mismatching attitude functions: Implications for scrutiny of persuasive messages. *Personality and Social Psychology Bulletin*, *24*, 227–240.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 41–72). New York: Guilford Press.

Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, *34*, 874–884.

Petty, R. E., Wheeler, S. C., & Bizer, G. Y. (1999b). Is there one persuasion process or more? Lumping versus splitting in attitude change theories. *Psychological Inquiry*, *10*, 156–163.

Petty, R. E., Wheeler, S. C., & Bizer, G. Y. (2000). Attitude functions and persuasion: An elaboration likelihood approach to matched versus mismatched messages. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 133–162). Mahwah, NJ: Erlbaum.

Piedmont, R. L., McCrae, R. R., & Costa, P. T. J. (1992). An assessment of the Edwards Preference Schedule from the perspective of the five-factor model. *Journal of Personality Assessment*, *58*, 67–78.

Pierro, A., Mannetti, L., Kruglanski, A. W., & Sleeth-Keppler, D. (2004). Relevance override: On the reduced impact of "cues" under high-motivation conditions of persuasion studies. *Journal of Personality and Social Psychology*, *86*, 251–264.

Pratkanis, A. R., Greenwald, A. G., Leippe, M. R., & Baumgardner, M. H. (1988). In search of reliable persuasion effects: III. The sleeper effect is dead. Long live the sleeper effect. *Journal of Personality and Social Psychology*, *54*, 203–218.

Prentice, D. A. (1987). Psychological correspondence of possessions, attitudes, and values. *Journal of Personality and Social Psychology*, *53*, 993–1003.

Razran, G. H. S. (1938). Conditioning away social bias by the luncheon technique. *Psychological Bulletin*, *35*, 693.

Razran, G. H. S. (1940). Conditioned response changes in rating and appraising sociopolitical slogans. *Psychological Bulletin*, *37*, 481.

Rhodes, N., & Wood, W. (1992). Self-esteem and intelligence affect influenceability: The mediating role of message reception. *Psychological Bulletin*, *111*, 156–171.

Rips, L. J., Brem, S. K., & Bailenson, J. B. (1999). Reasoning dialogues. *Current Directions in Psychological Science*, *8*, 172–177.

Rokeach, M. (1973). *The nature of human values*. New York: Free Press.

Roskos-Ewoldsen, D. R., Bichsel, J., & Hoffman, K. (2002). The influence of accessibility of source likability on persuasion. *Journal of Experimental Social Psychology*, *38*, 137–143.

Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, *121*, 3–19.

Rothschild, M. L. (1979). Marketing communications in nonbusiness situations or why it's so hard to sell brotherhood like soap. *Journal of Marketing*, *43*, 11–20.

Sagarin, B. J., Cialdini, R. B., Rice, W. E., & Serna, S. B. (2002). Dispelling the illusion of invulnerability: The motivations and mechanisms of resistance to persuasion. *Journal of Personality and Social Psychology*, *83*, 526–541.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin*, *21*, 704–710.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Mahwah, NJ: Erlbaum.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1–65.

Sedikides, C., & Strube, M. J. (1997). Self evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. *Advances in Experimental Social Psychology*, *29*, 209–269.

Shavitt, S. (1989). Operationalizing functional theories of attitude. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 311–337). Hillsdale, NJ: Erlbaum.

Shavitt, S. (1990). The role of attitude objects in attitude functions. *Journal of Experimental Social Psychology*, *26*, 124–148.

Shavitt, S., & Nelson, M. R. (2000). The social identity function in persuasion: Communicated meanings of product preferences. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 37–57). Mahwah, NJ: Erlbaum.

Sherif, C. W. (1980). Social values, attitudes, and the involvement of the self. In H. E. J. Howe & M. M. Page (Eds.), *Nebraska symposium on motivation, 1979* (Vol. 27, pp. 1–64). Lincoln: University of Nebraska Press.

Sherif, C. W., & Sherif, M. (Eds.). (1967). *Attitude, ego-involvement, and change*. New York: Wiley.

Sherif, C. W., Sherif, M., & Nebergall, R. E. (1965). *Attitude and attitude change*. Philadelphia: Saunders.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, *26*, 1046–1058.

Siero, F. W., & Doosje, B. J. (1993). Attitude change following persuasive communication: Integrating social judgment theory and the elaboration likelihood model. *European Journal of Social Psychology*, *23*, 541–554.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. New York: Wiley.

Smith, S. M., & Shaffer, D. R. (1995). Speed of speech and persuasion: Evidence for multiple effects. *Personality and Social Psychology Bulletin*, *21*, 1051–1060.

Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537.

Snyder, M. (1986). *Public appearances/private realities: The psychology of self-monitoring*. New York: Freeman.

Sparks, P., Hedderley, D., & Shepherd, R. (1991). Expectancy-value models of attitudes: A note on the relationship between theory and methodology. *European Journal of Social Psychology*, *21*, 261–271.

Stapel, D. A., & Winkielman, P. (1998). Assimilation and contrast as a function of context-target similarity, distinctness, and dimensional relevance. *Personality and Social Psychology Bulletin*, *24*, 634–646.

Strahan, E. J., Spencer, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: Striking while the iron is hot. *Journal of Experimental Social Psychology*, *38*, 556–568.

Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology*, *57*, 388–398.

Tetlock, P. E., Lerner, J. S., & Boettger, R. (1996). The dilution effect: Judgmental bias, conversational convention, or a bit of both? *European Journal of Social Psychology*, *26*, 915–934.

Thompson, E. P., Kruglanski, A. W., & Spiegel, S. (2000). Attitudes as knowledge structures and persuasion as a specific case of subjective knowledge acquisition. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 59–95). Mahwah, NJ: Erlbaum.

Thompson, M. M., & Zanna, M. P. (1995). The conflicted individual: Personality-based and domain-specific antecedents of ambivalent social attitudes. *Journal of Personality*, *63*, 259–288.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not

be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 361–386). Mahwah, NJ: Erlbaum.

Trafimow, D., & Sheeran, P. (1998). Some tests of the distinction between cognitive and affective beliefs. *Journal of Experimental Social Psychology*, *34*, 378–397.

Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Tykocinski, O., Higgins, E. T., & Chaiken, S. (1994). Message framing, self-discrepancies, and yielding to persuasive messages: The motivational significance of psychological situations. *Personality and Social Psychology Bulletin*, *20*, 107–115.

van der Pligt, J., & de Vries, N. (1998). Belief importance in expectancy-value models of attitudes. *Journal of Applied Social Psychology*, *28*, 1339–1354.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). Two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*, 820–838.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. *Advances in Experimental Social Psychology*, *29*, 141–208.

Wegener, D. T., & Petty, R. E. (2001). Understanding effects of mood through the elaboration likelihood and flexible correction models. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 177–210). Mahwah, NJ: Erlbaum.

Wegener, D. T., Petty, R. E., & Klein, D. J. (1994). Effects of mood on high elaboration attitude change: The mediating role of likelihood judgments. *European Journal of Social Psychology*, *24*, 25–43.

Whittler, T. E., & Spira, J. S. (2002). Model's race: A peripheral cue in advertising messages? *Journal of Consumer Psychology*, *12*, 291–301.

Wiebe, G. D. (1951). Merchandising commodities and citizenship on television. *Public Opinion Quarterly*, *15*, 679–691.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117–142.

Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. *Advances in Experimental Social Psychology*, *22*, 287–343.

Wilson, T. D., Houston, C. E., & Meyers, J. M. (1998). Choose your poison: Effects of lay beliefs about mental processes on attitude change. *Social Cognition*, *16*, 114–132.

Wilson, T. D., Kraft, D., & Dunn, D. S. (1989). The disruptive effects of explaining attitudes: The moderating effect of knowledge about the attitude object. *Journal of Experimental Social Psychology*, *25*, 379–400.

Wood, W., Kallgren, C. A., & Preisler, R. M. (1985). Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes. *Journal of Experimental Social Psychology*, *21*, 73–85.

Wood, W., Rhodes, N., & Biek, M. (1995). Working knowledge and attitude strength: An information-processing analysis. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 283–313). Mahwah, NJ: Erlbaum.

Woodall, W. G., & Burgoon, J. K. (1984). Talking fast and changing attitudes: A critique and clarification. *Journal of Nonverbal Behavior*, *8*, 126–142.

Wyer, R. S., Jr. (1970). Quantitative prediction of belief and opinion change: A further test of a subjective probability model. *Journal of Personality and Social Psychology*, *16*, 559–570.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151–175.

Zanna, M. P., Fazio, R. H., & Ross, M. (1994). The persistence of persuasion. In E. Langer & R. C. Schank (Eds.), *Beliefs, reasoning, and decision making: Psycho-logic in honor of Bob Abelson* (pp. 347–362). Mahwah, NJ: Erlbaum.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge, UK: Cambridge University Press.

CHAPTER 25

Foundations of Interpersonal Trust

JEFFRY A. SIMPSON

Trust: "confidence that [one] will find what is desired [from another] rather than what is feared."
—DEUTSCH (1973, p. 148)

According to Morton Deutsch, who many consider the founder of modern theory and research on trust, trust involves the delicate juxtaposition of peoples' loftiest hopes and aspirations in relation to their deepest worries and darkest fears. For this reason, situations in which trust is relevant often generate strong approach–avoidance gradients, particularly when individuals feel vulnerable and must count on the benevolence of their partners to receive important outcomes. Although not a complete definition of the construct, Deutsch's crisp observation captures the quintessential features of interpersonal trust, which is the topic of this chapter.

Trust is one of the most important components—and perhaps the most essential ingredient—for the development and maintenance of happy, well-functioning relationships (Fehr, 1988; Regan, Kocan, & Whitlock, 1998). Several lifespan theories, ranging from Bowlby's (1969, 1973) attachment theory to Erikson's (1963) theory of psychosocial development, contend that early exposure to relationships defined by strong trust lays the foundation on which most future relationships are constructed. Without some basic level of trust, individuals are reluctant to initiate, invest in, or sustain most voluntary relationships (e.g., with friends, recreation partners, and romantic partners). Indeed, trust appears to be crucial for the emergence of healthy and secure relationships (Holmes & Rempel, 1989; Larzelere & Huston, 1980), and the betrayal of trust is one of the most commonly mentioned reasons for the demise of relationships (Miller & Rempel, 2004). Outside the realm of relation-

ships, trust acts as a social lubricant that promotes cooperation between group members, sustains social order, and permits beneficial long-term exchanges that otherwise might never occur (Cook & Cooper, 2003; Ostrom & Walker, 2003).

Given the central importance of trust in interpersonal affairs, one might suspect that it has received widespread theoretical and empirical attention. Though there have been significant pockets of theory (e.g., Holmes & Rempel, 1989; Kelley & Thibaut, 1978) and research (e.g., Mikulincer, 1998; Rempel, Holmes, & Zanna, 1985) on trust, surprisingly little is known about how trust develops, how it is maintained, how it shapes and interacts with major interpersonal processes (e.g., the development of intimacy and closeness) and outcomes (e.g., relationship satisfaction and stability), and how it unravels when betrayed.

Why has trust received such limited attention? There are a variety of viable reasons. To begin with, trust is a complex, multidimensional construct, rendering it amenable to diverse interpretations in different social situations (Kramer & Carnevale, 2001). Second, trust might be construed differently and take on varying importance at different stages of relationship development (Fletcher, Simpson, & Thomas, 2000; Larzelere & Huston, 1980). Views about what constitutes sufficient trust in a partner/relationship during the initial stages of relationship development (such as a partner's general reliability and predictability) may be quite different from those used to gauge trust in long-term relationships (such as a partner's

dependability and one's confidence that he or she will remain loyal and supportive over time). Third, trust develops and changes in situations that are notoriously difficult to observe and study, such as in "strain test" situations (Holmes, 1981; Kelley, 1983). Creating or simulating these situations in the lab is challenging, especially when one wants to study partners in established relationships. Fourth, relationship satisfaction has often been treated as a proxy for many conceptually related yet distinct relationship constructs, including trust (Huston, 2000). The failure of researchers to clearly define, measure, and conceptually tease apart relationship constructs that are correlated with yet theoretically distinct from interpersonal trust—constructs such as satisfaction, love, commitment, passion, and intimacy—has hampered our understanding of trust.

This chapter addresses a series of fundamental questions central to understanding interpersonal aspects of trust. What is interpersonal trust? When, how, and why does it develop? What are the dispositional, relational, structural, and situational factors that instigate, facilitate, inhibit, or destroy it? Why do some people repeatedly experience greater trust in their close relationships than others? Why are certain relationships characterized by greater or more mutual levels of trust? What happens when trust is seriously questioned or betrayed? Provisional answers to these and other questions are offered.

The first section of the chapter reviews basic definitions, conceptualizations, and operationalizations of interpersonal trust. After reviewing some of the linguistic origins of trust, both individualistic (dispositional) and interpersonal (dyadic) definitions and conceptualizations of trust are presented. The second section highlights some of the major theoretical foundations and bases of trust at different levels of conceptual analysis. At the ultimate level of analysis, traditional genetic evolutionary models relevant to trust as well as multilevel selection/cultural coevolutionary models are showcased. At the ontogenetic level, some prominent lifespan models of social and personality development that are most pertinent to interpersonal trust are highlighted. At the proximate level, a few of the most significant social and psychological processes bearing on trust are outlined. Following this, major models specifying the normative (i.e., typical or modal) and individual-difference processes believed to govern the development, maintenance, and deterioration of trust in close relationships are discussed. The third section provides a selective yet representative overview of research on trust, with most attention focusing on interpersonal (rather than intergroup) trust. This overview begins with the seminal contributions of Deutsch and the early Prisoner's Dilemma Game (PDG) studies conducted prior to the mid-1960s, progresses to the dispositional movement that was popular from the late 1960s through the mid 1970s, and concludes with more recent dyadic formulations of trust. In the final section, six core principles of trust are identified. Following this, important constructs from different interpersonal models are merged to form an integrative process model, which suggests how trust might develop and be maintained in relationships.

DEFINITIONS AND CONCEPTUALIZATIONS OF TRUST

As alluded to previously, the study of interpersonal trust has been constrained by several interrelated factors. Two of the greatest hindrances have been the complex, multifaceted nature of the construct and the tendency of investigators to define and operationalize trust differently in different studies. Another part of the problem, however, stems from how trust is defined and expressed in different languages. In French, for example, the word for trust indicates that one merely has "confidence" in someone or something. Several other languages either do not have specific words for trust or have created new words only recently. For instance, there is no noun in Norwegian that corresponds to the English definition of trust, and the Japanese invented a new word to capture the construct merely a century ago (Hardin, 2003). Even in English, which contains more words than any other language and has a fairly nuanced definition of trust, the origins of the word come from the root "tryst," which refers to "an agreement . . . to meet" or "an appointed meeting or meeting place." Thus, even languages that have a clear word or phrase connoting trust might have derived the term from slightly different roots.

Most of the confusion surrounding trust, however, emanates from disparities in, or the lack of precision with which, trust has been defined and operationalized in different studies by different investigators. Some of the earliest research-based definitions of trust adopted a dispositional view of human nature (e.g., Rotter, 1971; Sato, 1988; Wrightsman, 1991). According to this perspective, trust entails generalized beliefs and attitudes about the degree to which other people are likely to be reliable, cooperative, or helpful, independent of the specific context or situation in which an interaction with them might take place. This global, context-free conceptualization of trust, which is assessed by measures such as the Machiavellianism Scale (Christie & Geis, 1970) and Wrightsman's (1974) Philosophies of Human Nature Scales, conceptualizes trust as a stable dispositional orientation toward the world and the people in it. As a rule, individuals who score higher on dispositional trust have warmer, more communal, and more benevolent perceptions of others, whereas less dispositionally trusting individuals harbor colder, more individualistic, and more cynical views of others (Wrightsman, 1991).

In the early 1980s, conceptualizations and measures of trust started to become more partner and relationship specific (Kelley & Thibaut, 1978; Lewis & Weigert, 1985; Rempel et al., 1985). According to this more dyadic (interpersonal) perspective, trust is a psychological state or orientation of an actor (the truster) toward a specific person (the trustee) with whom the actor is in some way interdependent (i.e., the truster needs the trustee's cooperation to attain valued outcomes or resources). What makes trust a particularly complex construct is that it has three components (e.g., "*I* trust *you* to do *X*"; Hardin, 2003). Accordingly, trust is a function of properties and characteristics of the self (I), the specific partner with whom one is interacting (you), and the unique features,

requirements, or constraints of the current situation (to do X). When any one of these components changes, an individual's thoughts, perceptions, feelings, and actions with regard to trusting a particular other may also change.

To complicate matters, individuals also vary in how broadly they define what "X" is or could be. For some people involved in certain relationships, "X" might constitute almost anything (e.g., to do the shopping, to be helpful and considerate, or to offer emotional support in times of need). Individuals who define "X" broadly should think, feel, and behave in a more stable manner across different trust-relevant situations, at least with reference to the same partner. For other individuals, however, "X" may be limited to a small set of circumscribed behaviors or activities (e.g., to do the shopping, *or* to be helpful and considerate, *or* to provide emotional support, but not all three). Individuals who subscribe to a more activity-specific view should adjust their trust-relevant thoughts, perceptions, feelings, and actions as situations and events change. Elements of Hardin's (2003) tripartite definition of trust can be found in many definitions and conceptualizations of the construct that have been proposed by various interpersonal and intergroup scholars.

Several scholars have also advanced more content-specific definitions of trust. Intergroup researchers, for example, have defined trust as a specific set of socially learned *expectations* that people hold about various social systems, ranging from other people to social organizations to the larger moral social order (Barber, 1983). Trust has also been defined as a constellation of *beliefs* regarding the extent to which others are or will be concerned about one's personal welfare and best interests (Pruitt & Rubin, 1986), and as a set of *attributions* that individuals routinely make when inferring motives beneath the actions of their partners (Tyler, 2001). According to this attributional perspective, trust is evident when individuals repeatedly presume that specific others are concerned about their welfare, will take their views and personal interests into account when making decisions, and will work toward fair and equitable outcomes.

Trust has also been defined in relation to specific situational contexts. Bacharach and Gambetta (2001), for instance, conceptualize trust as involving situations in which an individual expects his or her partner to pursue a particular course action (i.e., to do X) and two conditions hold: (1) if the partner does *not* perform the expected action, the individual would have benefited more by doing something else; and (2) the individual encourages the partner to complete the expected action. This viewpoint suggests that trust should be highest when both partners' "raw" payoffs (i.e., the outcomes that are best for each of them personally) coincide with their "all-in" payoffs (i.e., the outcomes that are best, on average, for themselves, their partner, *and* the relationship). Trust should also be higher when both individuals believe that their *partner* will base his or her decisions and actions on what is best for the relationship (i.e., on joint-interest payoffs) in a given situation, even if the partner's "raw" and "all-in" payoffs do not completely coincide. Over time,

self-expansion processes (Aron, Aron, & Norman, 2001) might help relationship partners transform their raw (i.e., selfish and self-centered) payoffs so they overlap more closely with their all-in (i.e., more partner- and relationship-centered) payoffs. Doing so could, in turn, generate better or more stable prorelationship outcomes (Kelley & Thibaut, 1978). Kelley and colleagues (2003) have recently proposed that trust can also be conceptualized as a particular type of interpersonal situation. In particular, trust situations involve the unique configuration of high interdependence between partners, a mixture of coordination *and* exchange as the basis for interdependence, and moderately corresponding interests. This new model is discussed in greater detail later.

Blending features of various definitions, Kramer and Carnevale (2001) claim that trust involves a suite of specific beliefs, expectations, *and* attributions that the actions of partners with whom one is interdependent will be beneficial (or at least not detrimental) to one's self-interest, especially in situations in which one must count on partners to provide unique benefits or valuable outcomes. They argue that trust-relevant situations elicit two interlocking cognitive processes: (1) feelings of vulnerability, which arise from uncertainty about the partner's true motives, intentions, or actions with regard to the self, and (2) perceptions and expectations about how the partner will behave across time, particularly in "strain test" situations in which an individual is highly outcome dependent and specific actions or decisions that would promote his or her own best interests are at odds with those that would maximally benefit the partner (Holmes, 1981; Kelley, 1983). When the partner promotes the individual's best interests over his or her own interests, *both* parties may experience heightened trust.

Scholars have also offered dyadic definitions of trust. Working within interdependence theory (Thibaut & Kelley, 1959), Wieselquist, Rusbult, Foster, and Agnew (1999) suggest that trust is evident when individuals believe that their partners are highly committed, harbor benevolent intentions, and are willing to undergo prorelationship transformations of motivation that result in self-sacrificial or accommodative behaviors (e.g., "voice" reactions during relationship conflict). Adopting a slightly different theoretical angle, Holmes and his colleagues (Holmes & Rempel, 1989; Rempel et al., 1985) propose that trust has three components, namely, the degree to which (1) partners are perceived as reliable (predictable); (2) partners are perceived to be concerned about one's welfare and are willing to support one's best interests, especially in times of need (high dependability); and (3) individuals are confident about the continued strength and permanence of the partner and relationship (faith). In recent years, these investigators have shifted their conceptualization of dyadic trust away from predictability and more toward dependability and faith (see Rempel, Ross, & Holmes, 2001).

Interpersonal trust can also be conceptualized in terms of the core motives that may drive and sustain it. According to McClintock (1972, 1976), five core social motives can shape the level or quality of trust that exists in a particular relationship. Specifically, an individual's inclina-

tion to trust another person can be motivated to advance (1) his or her own gain maximization (reflecting an individualistic or egoistic orientation), (2) his or her relative gain maximization (reflecting a competitive orientation), (3) his or her joint gain maximization (reflecting a cooperative orientation), (4) others' gain maximization (reflecting an altruistic orientation), and/or (5) others' gain minimization (reflecting an aggressive, spiteful orientation). Greater dyadic trust should be witnessed when individuals and their partners *both* display—or believe that they display—either joint gain maximization or others' gain maximization social motives, depending on the specific social and situational circumstances.

Deutsch (1973) also has speculated that individuals trust others for assorted motivational reasons. Individuals may, for instance, trust others out of despair (when the possible consequences of *not* trusting someone could be worse than trusting them), social conformity (trusting to avoid ostracism or violations of social norms), innocence (trusting others due to lack of knowledge, information, or experience), or impulsiveness (failing to give proper weight to the future consequences of an act). In addition, individuals might trust others out of virtue (to affirm their core values), masochism (in the hope of eventually being betrayed), faith (hoping that dreaded consequences never unfold), confidence (when individuals believe they will get what they desire rather than what they fear), or a desire for risk taking (trusting as a way of taking greater risks).

Trust might also have further motivational foundations. For example, interpersonal trust that is based on important *personal* goals and motives (e.g., I trust you because I know that you love me and will always consider my best interests) should be different from trust that is based on either *moral* goals/motives (e.g., I trust you because you are morally committed to fulfilling your promises to me) or *structural* goals/motives (e.g., I trust you because you have other relationships or valued possessions that would be damaged if you did not fulfill your obligations to me). Trust that is anchored in personal goals and motives should remain stable as long as personal goals and motives remain constant, even if moral or structural factors change. Trust that is grounded in moral or structural factors, on the other hand, should remain stable unless major moral or structural variables shift across time. Insufficient attention has been devoted to determining whether and how different motivational sources influence the nature, quality, and strength of trust in close relationships.

In summary, the multidimensional nature of trust makes the construct challenging to define, operationalize, and measure. The different definitions and operationalizations of trust that have been proposed do not constitute rival alternatives; rather, they highlight different facets and components of trust. Hardin's (2003) tripartite definition of the construct—"*I* trust *you* to do *X*"—accentuates the complex, multidimensional, and very interpersonal nature of trust. This complexity is simply compounded by the fact that assorted social motives and personal needs can serve as the basis of trust with different partners and in different relationships.

THEORETICAL FOUNDATIONS AND BASES OF TRUST

Virtually all constructs are situated within, and require explanation at, three levels of conceptual analysis (Sherman, 1988; Tinbergen, 1963): ultimate causation, ontogeny, and proximate causation. Questions of ultimate causation focus on the possible phylogenetic, evolutionary, or cultural/coevolutionary origins of a specific trait or behavior. Why, for instance, do many people trust total strangers, at least until they demonstrate that they are untrustworthy? Why do people continually enter complicated, long-term exchange relationships with so many different individuals? What genetic/evolutionary and cultural/coevolutionary forces might have instigated and sustained these proclivities?

Questions of ontogeny address the experiential factors that shape how a given trait or behavior develops and changes across the lifespan. Ontogenetic questions fall into two categories (Tooby & Cosmides, 1990). Some address how or why certain early environmental experiences shunt individuals down different developmental pathways, culminating in the enactment of different behavioral strategies in adulthood. What early life events or experiences, for example, are associated with the tendency to develop more versus less trusting relationships in adulthood? Other ontogenetic questions focus on how and why specific developmental experiences produce different activation thresholds. What sorts of early social experiences, for instance, motivate certain people to value trust highly, to think about it often, or to respond more intensely when it is violated?

Questions of proximate causation deal with how and why specific stimuli or events in the current environment activate, maintain, or regulate a given trait or behavior. Which classes of situations or events, for example, activate working models (relationship schemas) relevant to trust? Which specific experiences or events make people more versus less likely to trust others at particular points during a relationship? Constructs cannot be fully understood unless cornerstone questions relevant to each level of analysis are asked and sufficiently answered.

Ultimate Causation Explanations

To comprehend the distal causal factors that might be responsible for the development of the capacity to trust others witnessed in humans, one must first look back to the most stable features of the environments in which humans most likely evolved. During more than 98% of human evolutionary history, our ancestors lived as hunters and gatherers (Hill, 2002; Kelly, 1995), most likely in small, cooperative tribes or bands (Richerson & Boyd, 2005). Many people in a tribe or band were biologically related (Foley, 1987), and complete strangers were probably encountered infrequently, most often during intertribal trading or war (Wright, 1994). Although some people migrated in and out of their original tribes/bands, most individuals probably lived in the same tribe/band their entire life. Children probably were raised with considerable help from extended family members and the

entire tribe/band, and older children—especially older siblings—most likely assumed important roles in socializing younger children (Eibl-Eibesfeldt, 1989). Both genders participated in securing food, with men doing most of the hunting and women doing most of the gathering (Konner, 1982; Richerson & Boyd, 2005). Thus, extensive cooperation with other tribe/band members—both kin as well as biologically unrelated individuals—was mandatory, particularly in light of the changing and precarious nature of the climate, competing tribes, and the food supply. These conditions happen to be ideal for reciprocal altruism to evolve (see Cosmides & Tooby, 1992). Brewer and Caporeal (1990) suggest that active participation in cooperative groups might have been the primary "survival strategy" of early humans. Willingness to enter and maintain mutually cooperative, long-term alliances with others, therefore, may have been essential for survival, successful reproduction, and adequate parenting.

Various gene-centered evolutionary models can explain how and why humans developed the capacity to trust others and become so disturbed when trust is betrayed. According to inclusive fitness theory (Hamilton, 1964), genes are replicated not only through one's own reproduction but also via the reproduction of biological relatives who carry the same genes. Selection, therefore, should have favored individuals who invested in and made sacrifices for their relatives, provided that the average cost of giving benefits (C) was less than the value of the benefits to relatives (B) multiplied by their degree of genetic relatedness (r) (i.e., Hamilton's rule, where C < rB). Although inclusive fitness theory can explain how and why altruism, cooperation, and the capacity for trust could have evolved with respect to biological relatives, it fails to explain the evolution of these traits with regard to interactions and social exchanges involving biologically unrelated individuals.

To solve this conundrum, Trivers (1971) developed reciprocal altruism theory. According to Trivers (1971), altruism is a "behavior that benefits another organism, not closely related, while being apparently detrimental to the organism performing the behavior, benefit and detrimental being defined in terms of contribution to inclusive fitness" (p. 35). According to this view, if two biologically unrelated individuals provide mutual benefits to each other that are greater than the costs each individual incurs by providing the benefits, both individuals should benefit through the economic principle of gains in trade. Consequently, genes that led our ancestors to recognize and selectively enter certain mutually beneficial transactions with nonkin (e.g., long-term tit-for-tat exchanges with highly resourceful and trustworthy partners) could have been preferentially selected.

Simulation research has confirmed that tit-for-tat strategies (whereby positive partner overtures are immediately rewarded and negative ones are immediately punished) tend to develop quickly and remain stable as long as interaction partners continue to make cooperative choices, at least in two-person experimental games (Axelrod, 1984). Moreover, all the conditions necessary for the evolution of reciprocal altruism in humans—important benefits can be conferred, individuals have re-

peated interactions with the same people, individuals can remember whom they have given benefits to and received benefits from, and exchange decisions are based on the outcomes of earlier interactions with certain people—were probably present during much of evolutionary history.

Among others, Kurzban (2003) suggests that the need for cooperative hunting could have been one of the major selection pressures that jump-started reciprocal altruism in humans. Moreover, delayed exchanges of goods and resources may have been more common in evolutionary environments than simultaneous exchanges, requiring that trust in others be carefully and judiciously placed. Interestingly, tit-for-tat strategies require a willingness to trust partners and to be cooperative on the first "move" (trial), after which decisions are governed by whether partners behave cooperatively or noncooperatively on subsequent trials. Extending these ideas, Clutton-Brock and Parker (1995) propose that spite (i.e., the inclination to punish or ostracize defecting or uncooperative individuals, even when such actions are costly to the self) could have evolved to "back up" trust if partners reneged on important promises. Tooby and Cosmides (1990, 1992) conjecture that certain specialized cognitive abilities in humans (e.g., cheater detection and superior memory for faces) may have evolved to help individuals identify and envision new ways in which valuable resources could be exchanged, further fueling the evolution of reciprocal altruism. Certain cognitive adaptations, therefore, might have accelerated the evolution of trust, including humans' specialized abilities to discern when trust is warranted and well-placed and when it might be violated (see Kurzban, 2003).

Other ultimate-level accounts of trust have been formulated in response to the fact that gene-centered evolutionary models do not fully explain the pervasiveness and depth of human altruism. These accounts, which are collectively known as multilevel selection or gene-cultural coevolutionary models, claim that humans are unique among species in their tendency to display "strong reciprocity" (Fehr & Fischbacher, 2003). Strong reciprocity is evident when individuals assume the costs of rewarding or punishing others in situations in which cooperation is required to secure vital resources or good outcomes, even if "enforcers" receive no personal benefits or incur major costs. Unlike reciprocal altruism, which presumes that individuals should reward or punish others only if tangible benefits are likely to be received (Axelrod & Hamilton, 1981), strong reciprocity indicates that individuals are willing to enforce important social rules or norms to ensure that cheaters and noncooperators do not destroy cooperation and goodwill within groups. Laboratory studies using the Ultimatum Game have confirmed that most individuals closely monitor and quickly punish people who behave unfairly (e.g., who cheat or fail to reciprocate cooperation) or who offer others unfair outcomes, even if providing sanctions harms their own rational self-interest (Fehr & Fischbacher, 2003). Research using intergenerational Ultimatum Games has also revealed that receiving advice from previous players increases altruistic punishments and rewards enacted by

current players, and that players who receive advice achieve greater cooperation from others over time.

What might explain this strong willingness to make personal sacrifices in the service of maintaining cooperative norms and behavior? The answer probably lies in how easily cooperation can disintegrate. Because a very small percentage of free riders or chronic cheaters can destroy cooperation in most groups, cooperative systems usually fail unless a clear majority of group members vigilantly monitor and sanction norm violaters, even if they have no personal stake or investment in a given interaction (Fehr & Schmidt, 1999). Individuals who regularly police and enforce important rules and norms, however, may also gain personal benefits through being seen as highly altruistic, which either could enhance their reputation within groups (Alexander, 1987; Nowak & Sigmund, 1998) or signal that they are sufficiently viable to withstand the costs of engaging in altruistic acts (Gintis, Smith, & Bowles, 2001; Zahavi, 1995)

Advocates of gene–cultural coevolutionary models have also questioned whether tit-for-tat strategies could have been responsible for the evolution of reciprocal altruism in humans. Although the results of repeated two-person interactions suggest that tit-for-tat strategies develop readily and remain fairly stable (Axelrod, 1984), these strategies are less stable in *n*-person PDGs unless nearly *all* group members cooperate on every trial (see Boyd & Richerson, 1988). Moreover, tit-for-tat strategies typically stipulate that individuals cannot "leave the game" (exit) and that third parties cannot intervene unless they can personally benefit from rewarding fair players or punishing unfair ones. These conditions rarely exist in most real-world settings.

In summary, gene–culture coevolutionary models posit that traditional genetic selection models (e.g., inclusive fitness theory and reciprocal altruism theory) cannot fully explain the evolution and existence of strong reciprocity, whereas theories of cultural group selection (Henrich & Boyd, 2001) and gene–culture coevolution (Gintis, 2003) can. These latter models propose that certain norms and institutions (e.g., food sharing, hunting, and serial monogamy) could have been maintained only if nearly all group members monitored and sanctioned important norm violations. This propensity may have spawned altruism in humans, which might have launched both the capacity for trust and the need to gauge the trustworthiness of others.

Ontogenetic Causation Explanations

Ontogenetic explanations address when, how, and why a particular trait or behavior develops and changes across the lifespan. The earliest ontogenetic theoretical accounts highlighted the critical role that experiencing trust with a primary caregiver assumes in social and personality development. In his psychosocial theory of development, for example, Erikson (1963) posits that all individuals encounter a series of basic conflicts at different stages of social development. The first major conflict, which takes place in infancy and early childhood, involves issues of basic trust versus mistrust. During this stage, highly dependent children first learn whether their basic needs will be met, ignored, or overindulged by their primary caregivers. Children who experience sufficient levels of satisfaction relative to frustration typically develop a sense of hope and confidence that their needs will be met in the future, which then influences how they perceive and behave when later psychosocial conflicts are encountered. Children who are ignored or overindulged tend to experience mistrust, which sets the stage for maladaptive responses to later life conflicts. Trust, therefore, serves as the foundation of personality and social development.

In developing attachment theory, Bowlby (1969, 1973, 1980) argued that children who receive warm, supportive, and situationally contingent care when distressed develop positive working models (relationship schemas) of themselves, significant others, and the world in general. One cardinal component of positive working models of others is having confidence in their availability and responsiveness, particularly during times of acute need. This core feature of attachment security is also a defining feature of trust. According to attachment theory, having trusting and supportive relationships early in life instills a sense of felt security (Sroufe & Waters, 1977), which provides an "inner resource" that allows individuals to take leaps of faith and place greater trust in significant others. Children who are subjected to rejection, unpredictable care, neglect, or abuse normally develop negative working models of others, which should retard or hinder the development of trust in later adult relationships (cf. Simpson, Ickes, & Grich, 1999).

Family systems theory (Bowen, 1976, 1978) has also been invoked to explain the development of interpersonal trust. According to this perspective, individuals who develop more differentiated self-concepts based on interactions with family members in the past should be able to establish and maintain emotional ties with future significant others more readily. Highly differentiated individuals tend to feel both connected to yet independent from significant others (Allison & Sabatelli, 1988; Shapiro, 1988). Because their self-concepts are more diversified, well balanced, and fully integrated, highly differentiated people can achieve closeness and intimacy without overidentifying with partners or becoming enmeshed with them (Bartle & Sabatelli, 1995). These characteristics should facilitate interaction patterns that result in greater trust, partly because highly differentiated people may feel more comfortable taking the risks necessary to forge deeper levels of trust.

Recent life-history models might also explain some of the developmental antecedents of dispositional trust. Adopting an evolutionary-based lifespan perspective, Belsky, Steinberg, and Draper (1991) contend that certain early experiences provide children with diagnostic information about the kinds of environments that they will most likely encounter in adulthood. As they develop, children implicitly use this information to adopt an appropriate reproductive strategy—one that is most likely to increase their inclusive fitness—in future environments (see also Hinde, 1986). According to this model, contextual factors in the family of origin (e.g., the amount of

stress, spousal harmony, and financial resources) systematically affect early childrearing experiences (e.g., the degree of sensitive, supportive, and responsive care). These experiences then influence children's psychological and behavioral development (e.g., their attachment patterns/styles and the nature of their working models), which in turn influence the rate of somatic development (how quickly individuals reach sexual maturity) and the mating orientation that they adopt in adulthood (short term vs. long term). Belsky and colleagues (1991) surmise that high levels of early stress and family dissention are associated with less sensitive parenting, which leads children to develop more negative working models of themselves and/or others and, therefore, more insecure attachment patterns. This information ostensibly signals that future pair-bonds may be short term and unstable, which prompts insecure individuals to mature more quickly and adopt a short-term, opportunistic orientation toward mating in which sexual intercourse with multiple partners occurs earlier in life and parental investment in children is lower. One central feature of short-term mating orientations is lower dyadic trust. The Belsky and colleagues model, therefore, suggests that the seeds of distrust may be planted early in life when individuals first learn about the world, primarily in their families of origin. Chisholm (1993, 1996) has proposed a similar lifespan model but argues that mortality rates in local environments are the focal cues that shunt individuals toward different adult mating orientations (see Simpson, 1999).

Proximate Causation Explanations

Various theories and models have outlined the proximal factors and processes that might facilitate or impede the development of trust in relationships.

Kelley and Colleagues (2003)

Informed by interdependence theory (Kelley & Thibaut, 1978), Kelley and colleagues (2003) have proposed that the development and deterioration of trust in relationships can be understood via the properties of the social situations in which relationships are embedded. They argue that "trust situations" reflect the unique combination of high interdependence between partners, coordination (the need to synchronize interactions), and exchange (the need to receive good outcomes from the partner) as the basis for interdependence, and corresponding (similar or joint) interests. Prototypical trust situations, therefore, entail both mutual partner control (e.g., quid pro quo situations) and corresponding mutual joint control (e.g., situations that require synchrony of choices).

Figure 25.1 depicts the unique 2 × 2 pattern of payoffs (benefits and costs) associated with choices that partners can make in prototypical trust situations. If both partners choose option 1 (they work on a difficult but important task together, reflected in the a1/b1 cell), each partner receives 20 units of benefit because the task gets accomplished while partners share the pleasure of each other's

FIGURE 25.1. The payoffs (outcomes) associated with Kelley et al.'s (2003) trust situation. Payoffs above the diagonal in each cell are for Partner A; those below the diagonal in each cell are for Partner B.

company. If both partners select option 2 (neither works on the task, represented by the a2/b2 cell), neither partner receives benefits because nothing gets accomplished. If partner A chooses option 2 (not to work on the task: a2) whereas partner B chooses option 1 (works solo on the task: b1), partner A benefits by 10 units because some progress is made on the task, but partner B experiences a net loss of 10 units because he or she is saddled with all the work. The reverse pattern exists when partner A chooses a1 (to work on the task) and partner B chooses b2 (to not work on the task).

Trust situations have several special properties. First, they involve collective rationality in that cooperative behavior by both partners (a1/b1) *always* yields superior outcomes compared to when partners fail to cooperate (a1/b2 or a2/b1). Second, the best outcome *always* occurs when both partners make the cooperative choice (a1/b1). Third, cooperative choices are risky because if one's partner decides to make a noncooperative choice, a cooperative choice generates the worst possible outcomes given that one has been exploited. Kelley and colleagues (2003) suggest that cooperation is undermined primarily by fear in trust situations and that, when individuals fail to make cooperative choices, they frequently do so for competitive or self-protective reasons. Partners who repeatedly enter or find themselves in trust situations and make joint cooperative choices should not only receive superior outcomes over time but also should develop greater trust. The "added value" that comes with joint cooperative decisions (i.e., the 10 additional units of benefit that accrue when both partners complete a valued activity *together*) is crucial to understanding what makes joint cooperative decisions so powerful and transforming. Though not much is known about what "added value" represents, most individuals find it intrinsically rewarding when, rather than taking advantage of their vulnerabilities, their partners repeatedly make choices that are beneficial to themselves or the relationship (Holmes, 1981).

Tit-for-tat tactics may generate trust more rapidly in newly formed relationships because they encourage rela-

tionship partners to gravitate toward a1/b1 decisions. When using tit-for-tat tactics, individuals remain cooperative until their partners make the first noncooperative overture, after which they respond noncooperatively until their partners revert back to cooperative choices. Because individuals never make the first noncooperative move, they are not perceived as aggressive or exploitative by their partners. Because noncooperative behavior is always responded to with noncooperation, partners also quickly learn the importance of cooperative norms. And because individuals immediately switch to cooperative choices when their partners do, they convey a willingness to forgive, which is likely to strengthen the inference that the individual yearns to establish and maintain cooperation. The strongest trust, however, may be forged in somewhat asymmetrical relationships in which one individual has power over the partner but consistently makes decisions that benefit the low-power partner at the expense of his or her own self-interest. In these situations, both the low-power partner and the high-power partner should perceive the benevolent and caring motives harbored by the high-power partner. These perceptions might contribute to the "added value" of a1/b1 decisions discerned by *both* partners.

Holmes and Rempel (1989)

Holmes and Rempel (1989) have proposed a dyadic model of trust that contains both normative (i.e., typical or modal) and individual difference components. According to their model, trust should be enhanced when relationship partners successfully resolve basic concerns regarding dependency issues, such as what happens when partners encounter noncorrespondent or conflicting payoff structures. The normative component of the model specifies the progression of events and psychological processes that should promote or diminish trust in relationships *in general*. The individual-difference component clarifies how these normative features may be qualified by the level of trust (high, medium, or low) that characterizes a current relationship.

From a normative standpoint, Holmes and Rempel (1989) suggest that the development of trust involves a process of uncertainty reduction as individuals gradually migrate from having confidence in their partner's general predictability to having confidence in their benevolent relationship values, motives, goals, and intentions. During the initial stages of relationship development, individuals typically project their hopes and ideals onto their partners. These projections tend to be vague and rarely challenged, usually because partners are not highly interdependent, their interactions are positive and limited in scope, they are well mannered, they want to view one another in the best possible light, and they have not yet experienced major conflicts. At this stage, trust is based chiefly on individuals' hopes that their partners will, over time, live up to these idealized images.

As individuals become increasingly dependent on their partners, concerns about being vulnerable to loss, exploitation, and rejection begin to surface. To assuage the anxiety associated with this predicament, individuals

try to quell relationship uncertainties by searching for evidence that their partners truly care for and are genuinely concerned about them, a process that Holmes and Rempel (1989) term "reciprocal reassurance." Reassurance should be strongest when partners make costly personal sacrifices for the individual, take bold risks such as providing costly "no strings attached" outcomes to the individual, or adopt more vulnerable positions than the individual has previously taken (Pruitt, 1965). Trust is not likely to grow, however, unless exchanges between relationship partners are reasonably reciprocal and balanced, and partners view one another as being neither overly anxious nor overly reluctant to venture interpersonal risks. Indeed, Holmes and Rempel contend that equal and gradually increasing mutual involvement by *both* partners should facilitate greater trust by further dampening relationship uncertainties. Equal involvement ensures that both partners are similarly dependent and, therefore, equally vulnerable. For this reason, equal involvement may function as a temporary substitute for trust until genuine trust can develop.

As relationships move into the accommodation stage, resolved relationship uncertainties allow individuals to take greater leaps of faith and to place increasing confidence in their partners. Gradually, most individuals begin to amass an "economy of surplus," which enables them to ignore, discount, or excuse minor partner transgressions that, if committed during earlier stages, could have shaken the emerging foundation of trust. During the accommodation stage, however, individuals start to view their partners more realistically, which can reignite relationship uncertainties as partners become increasingly interdependent and their self-interests occasionally diverge. To resolve this dilemma, individuals begin to view the relationship prospectively, whereby the costs of remaining in the relationship are weighed against the prospect of future rewards (Huesmann & Levinger, 1976; Kelley & Thibaut, 1978). Individuals also contemplate whether they are in the best possible relationship and can count on their partners to fulfill their fundamental needs. As a result, decisions to invest further in the relationship become more contingent upon subjective forecasts about what the future may hold. At this point, trust becomes less highly correlated with other markers of relationship quality such as love (Larzelere & Huston, 1980).

When relationship uncertainties resurface at the accommodation stage, individuals may enter or create "strain test" situations (Holmes, 1981; Kelley, 1983) to gauge their partner's true degree of commitment and devotion. If the partner (1) demonstrates care, concern, and responsiveness when the individual is highly vulnerable or dependent and (2) takes actions that run counter to his or her own best interests, trust should increase. Trust should also be enhanced if partners repeatedly acknowledge an individual's core needs and reaffirm his or her self-worth, if partners' actions are perceived as being unconditional, if partners are willing to relinquish control to individuals in important situations, if they have behaved fairly and benevolently in difficult (e.g., noncorrespondent) situations in the past, and if they engage in acts that lead individuals to make positive, intrinsi-

cally motivated attributions for the partner's behavior (Holmes & Rempel, 1989; see also Kruglanski, 1971).

Holmes and Rempel (1989) also elucidate how individual differences in trust may alter or qualify these normative stages and psychological processes. Individuals involved in relationships characterized by greater trust should have more benevolent, long-term views of their partner's motives and actions. They also should harbor more positive, well-integrated, and well-balanced working models that instill more benevolent expectations and, therefore, more constructive, problem-focused styles of resolving relationship difficulties.

Individuals involved in relationships containing moderate levels of trust should have more suspicious and guarded views of their partner's motives and behavior. Because they possess chronically accessible relationship uncertainties, "medium trust" individuals should closely monitor and frequently test their partner's care and responsiveness, even if such tests might confirm their darkest fears. Given their hypervigilant, emotion-focused style of coping with relationship problems, these individuals may unwittingly amplify or create the very outcomes they wish to avert by constantly questioning and testing their partner's trust and commitment, effectively driving the partner away.

Individuals involved in relationships defined by low trust have minimal confidence in their partner's ability or willingness to ever be caring or responsive. As a consequence, they hold highly cynical and negative views and expectations about relationships. This tendency may generate distress-maintaining attributions, whereby the implications of possible negative partner behaviors are ac-

centuated and positive partner actions are disregarded or dismissed. These cognitive tendencies should lead "low trust" individuals to display heightened anger or to withdraw when major relationship problems arise. Because a fair amount of research has addressed some of these speculations, we revisit the individual-difference component of Holmes and Rempel's model in the next section.

Wieselquist, Rusbult, Foster, and Agnew (1999)

Guided by the investment model (Rusbult, 1983), Wieselquist and colleagues (1999) claim that trust is most evident when partners display large prorelationship transformations of motivation, which are indexed by the discrepancy between a partner's initial self-interests (captured in the given matrix) versus his or her final, transformed outcomes and observable actions (represented in the effective matrix). Individuals who make larger transformations from purely self-interested outcomes (e.g., MaxOwn) to either partner (e.g., MaxOther) or relationship (e.g., MaxJoint) outcomes should and do have relationships characterized by greater trust (Yovetich & Rusbult, 1994). Given that commitment generates so many crucial relationship-maintenance acts (e.g., disparagement of alternative partners, willingness to sacrifice, accommodative behaviors, positive partner and relationship illusions, and a sense of cognitive interconnectedness), Wieselquist and colleagues believe that commitment should play a central role in both the generation of and changes in dyadic trust. According to their mutual cyclical growth model, shown in Figure 25.2, greater dependence by one partner in a relationship should in-
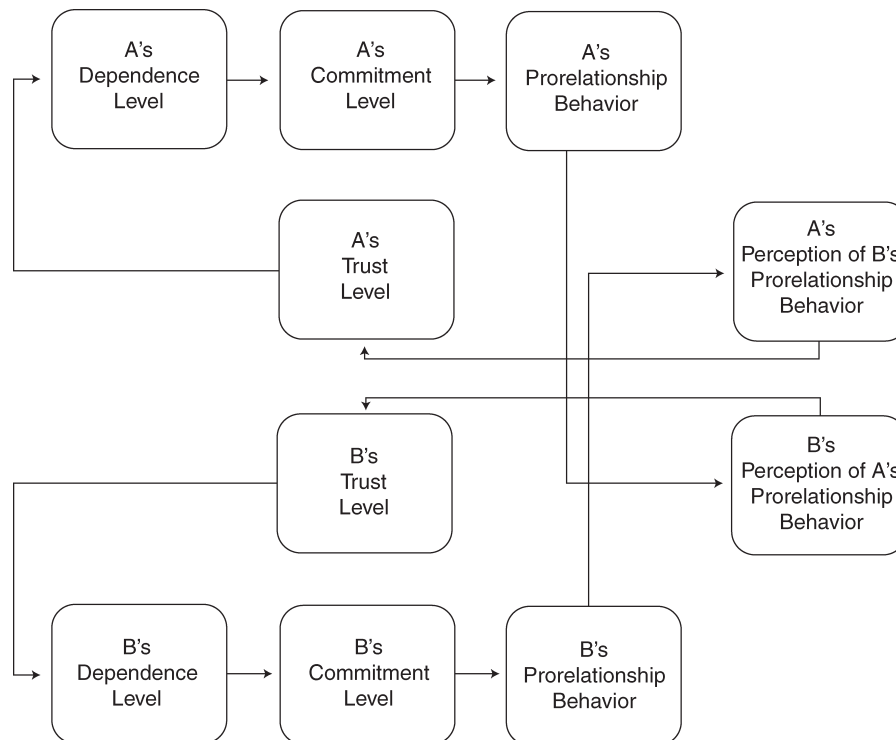
FIGURE 25.2. Wieselquist et al.'s (1999) mutual cyclical growth model.

still greater commitment, which should then launch prorelationship actions that yield MaxJoint outcomes. These behaviors should then be witnessed by the other partner who should experience greater trust. This, in turn, should promote greater dependence, greater commitment, and more prorelationship acts, which are witnessed by the first partner, who subsequently experiences greater trust. According to the model, the process then begins again, at least to some asymptotic level. As we shall see, although cross-sectional data support certain paths in the model, most of the antecedent variables do not predict temporal *changes* in the downstream variables.

Deutsch (1973)

Deutsch (1973) has proposed one of the most detailed proximate-level accounts of trust. He suggests that trust is evident when individuals select an "ambiguous path" in which the strength of their positive motivations to pursue a course of action is *less* than the strength of their negative motivations. In addition, individuals' subjective estimates of positive outcomes occurring must be greater than their subjective estimates of negative outcomes occurring, and they must have sufficiently high "security" to take the action. Subjective estimates ought to be based on past experiences in similar situations, past experiences with or beliefs about specific partners, more global working models of self and others, and one's confidence in being able to achieve positive outcomes or avoid negative ones.

A series of specific hypotheses can be derived from these central premises. Deutsch claims, for instance, that the longer the time frame in which events that would have negative motivational significance might take place, the more individuals should make trusting choices and decisions in their relationships. After an individual has made a trusting choice, he or she should then seek reasons that justify and support the wisdom of the choice, especially if the choice was originally difficult to make. Furthermore, an individual's perception of his or her *partner's* intention to perform a future action or behavior should depend on the partner's presumed strength of motivation to engage in the action or behavior, the partner's commitment to engaging in it, the source of the partner's motivation (e.g., internal vs. external or altruistic vs. exchange focused), and the individual's own attentional focus.

In terms of perceiving altruistic or trust-promoting intentions in others, Deutsch speculates that positive intentions should be more likely to be inferred when individuals believe that others (e.g., their partners) like them. Liking, in turn, should depend on the amount, frequency, and diversity of benefits that an individual has received from a partner in the past. It should also be greater if individuals are aware that their partners were *not* forced to provide benefits and if individuals believe that partners knew a priori that their actions would have positive consequences. Liking should also be heightened if the partner is perceived as having little to gain by providing benefits to the individual, if the partner's actions

provide greater gains to the individual than to the partner, and if the partner knows that his or her benevolent actions would be costly from the start.

In terms of perceptions of power, Deutsch claims that individuals who perceive that their partners have relatively greater power should be either more trusting *or* more suspicious (i.e., an amplification effect). More powerful individuals should be less inclined to infer altruistic motives and intentions in their partners and, hence, less likely to trust them. These perceptions should be easier to change, however, if individuals enter interactions or relationships with the belief or goal that mutual trust is possible. Deutsch argues that communicating a desire for cooperation (which can be conveyed by expectations, intentions, or reactions to norm agreements and norm violations) should increase dyadic trust and render the communicator more trustworthy.

In sum, more theoretical attention has been paid to the proximal processes that might facilitate or impede interpersonal trust than to ontogenetic processes. Models of proximate causation have viewed the antecedents of trust as anchored in how relationship partners routinely make choices in certain types of critical situations (e.g., prototypical trust situations; Kelley et al., 2003) and the typical manner in which trust usually unfolds as relationships develop, enter new stages and partners grow more interdependent (Holmes & Rempel, 1989). Proximate models have also highlighted the important role that commitment and the responses of relationship partners may assume in the development of trust in relationships (Wieselquist et al., 1999) and some of the more detailed psychological processes and mechanisms that might guide specific decisions to enter situations or make decisions that could promote or stall the development of trust (Deutsch, 1973).

A REPRESENTATIVE REVIEW OF RESEARCH ON INTERPERSONAL TRUST

Though there was a smattering of theory and research on trust prior to the late 1950s, Morton Deutsch was largely responsible for launching interest in the topic. In his early work, Deutsch (1960) marshaled evidence that the development of mutual trust between interaction partners was essential for cooperation to emerge and be maintained. He argued that mutual trust should develop when interaction partners adopt a positive and benevolent orientation toward one another's welfare. Using the PDG as the primary investigative paradigm, Deutsch confirmed that people who enter single-trial PDGs with a cooperative orientation tend to make more cooperative choices that yield greater gains for both interaction partners. On the other hand, those who have a competitive orientation frequently make noncooperative choices that culminate in losses for both partners. Deutsch also found that cooperative individuals do not necessarily communicate to enact cooperative choices, whereas competitive individuals often convey motives or behaviors that block or subvert trustworthy communication. Later studies

(e.g., Ellison & Firestone, 1974; MacDonald, Kessel, & Fuller, 1972; McAllister & Bergman, 1986) revealed that more intimate self-disclosure, which frequently includes expressions of vulnerability and willingness to trust others, generates greater liking for and trust in partners who also disclose personal information.

Unfortunately, there were problems with this early line of experimental work (see Cook & Cooper, 2003). First, Deutsch did not operationalize and measure trust directly in these initial studies. Cooperative behavior was treated as an indirect proxy of trust. Second, this initial wave of research confounded trust and cooperation, which are distinct constructs. Third, the early studies were grounded in a social exchange perspective, which left little room to explain cases in which trust develops in the absence of direct exchanges. Fourth, the majority of the early PDG studies examined single-trial interactions between strangers. By not permitting multiple interactions over time between people who shared a past or a future, it proved difficult to generalize the results of early PDG experiments to everyday interactions and relationships. Consequently, these early single-trial PDG studies produced outcomes that tend to be different than those found when iterative, higher-stakes games are used (e.g., Fehr & Fischbacher, 2003).

The rise of dispositional views of trust in the late 1960s and early 1970s was a reaction to these shortcomings. Rotter (1967), in fact, developed self-report scales of generalized (dispositional) trust directly in response to Deutsch's initial program of trust research. Rotter believed that single-trial PDG studies generated situation-specific results that were confined to a small set of competitive situations not representative of everyday interactions. Once dispositional measures of generalized trust were developed, trust research in the 1970s identified systematic differences between people who reported being more versus less trusting of people in general. This body of work revealed that individuals who score higher in dispositional trust are more trustworthy partners in different types of social interactions, find it more difficult to lie, are more well liked by others, contribute more to philanthropic projects, are better at judging whom to trust and whom *not* to trust, and are more willing to trust strangers (Rotter, 1971). Although greater dispositional distrust was initially tied to poorer psychological adjustment and greater antisocial behavior (Rotter, 1980), later research indicated that links between dispositional trust, maladjustment, and distress are modest (Comrey & Schiebel, 1983; Folkman, Lazarus, Gruen, & DeLongis, 1986; Wrightsman, 1974). When highly distrusting individuals have problems, they usually report greater interpersonal difficulties, especially with excessive competitiveness, envy, resentfulness, vindictiveness, and lack of feelings or concern for others (Gurtman, 1992). Highly trusting persons, in contrast, adopt more kind and benevolent orientations toward others. However, they are not more gullible or exploitable than less trusting persons (Rotter, 1980), perhaps because more trusting individuals respond to partner transgressions by immediately withdrawing cooperation when partners behave badly in a tit-for-tat, situation-contingent manner.

In the early 1980s, research shifted toward dyadic conceptions of trust. Kelley (1980), for example, speculated that certain interpersonal traits (dispositions) might be systematically correlated with dyadic measures of trust. He surmised that people who score lower on dominance and nurturance might be particularly inclined to distrust relationship partners (for the theoretical reasons, see Carson, 1979). Testing relations between various interpersonal traits and dyadic trust, Gaines and colleagues (1997) documented that affiliative–dominance (e.g., extraversion) is the best predictor of high-faith scores on Rempel and colleagues' (1985) dyadic trust scale. Aloofness–introversion, on the other hand, is the strongest predictor of low-faith scores.

Holmes (1991) also examined ties between dyadic measures of trust and global interpersonal orientations. He found that individuals who either yearn for extreme emotional intimacy *or* fear being too close tend to trust their current romantic partners less. Moreover, dyadic trust tends to be higher among people who value moderate amounts of autonomy in relationships and lower in those who place greater emphasis on being independent and self-sufficient. These effects might reflect the fact that the core motives underlying autonomy and affiliation are associated with different styles of coping with relationship uncertainty (Holmes & Rempel, 1989). They may also reflect the fact that these higher-level interpersonal orientations contain some of the cardinal features that define attachment security and insecurity (Bowlby, 1973; Mikulincer, 1998). Correlations between interpersonal traits (dispositions) and dyadic measures of trust, however, tend to be modest (Holmes & Rempel, 1989).

The principal drawback of the dispositional approach was that it did not address the inherently dyadic, partner-specific, and context-dependent nature of trust that characterizes many close relationships. For this reason, dyad-centered measures of trust (e.g., Johnson-George & Swap, 1982; Rempel et al., 1985) and dyad-centered programs of research (e.g., Rempel et al., 1985) rose in the mid-1980s, partially in response to theoretical advances spearheaded by Deutsch (1973), Kelley and Thibault (1978), and Holmes and Rempel (1989), and partially in response to landmark empirical studies that cast trust within a dyadic framework (e.g., Kelley & Stahelski, 1970; Larzelere & Huston, 1980). Several key findings have emerged from the dyad-centered approach, only some of which can be highlighted. Some of the most influential research has been guided by two of the most dominant theoretical perspectives in interpersonal relationships: interdependence theory (Kelley & Thibaut, 1978; Thibaut & Kelley, 1959) and attachment theory (Bowlby, 1969, 1973, 1980).

Interdependence Theory-Based Approaches

Working from an interdependence theory perspective, Holmes, Rempel, and their colleagues have engaged in one of the most systematic and sustained programs of empirical work on dyadic trust. While interested in both normative and individual difference components of trust, the bulk of their research has focused on how indi-

vidual differences in trust are related to the way in which individuals appraise, construe, feel, and behave in situations when partners must grapple with dependency issues and their self-interests are discrepant. Their research has highlighted several crucial differences in how persons involved in high-, medium-, or low-trust relationships process information and respond when dependency issues are salient.

Individuals involved in relationships defined by high levels of trust (whom Holmes & Rempel, 1989, term "high-trust" people) harbor more optimistic and benevolent expectations about their partner's motives, make more positive and global attributions about their partner's behaviors, and have more integrated and well-balanced working models that remain open to assimilating new information (Rempel et al., 2001). High-trust individuals also disregard or downplay their partner's negative relationship-relevant actions, which isolates and minimizes the potential negative impact of partner indiscretions. When resolving relationship conflicts, high-trust individuals display more positive affect and less negative affect (Holmes & Rempel, 1986), and their evaluations of their partners and relationships are less strongly tied to the emotions they experience during these intense discussions. High-trust individuals also evaluate their partners more positively, not only when they are asked to recall positive or neutral relationship experiences but particularly when they recount *negative* relationship experiences (Holmes & Rempel, 1989). That is, when high-trust individuals encounter relationship threats, they step back and contemplate their partner's probable benevolent goals and motives within a broader, long-term perspective (Holmes, 1981).

On the basis of these and other findings, Holmes and Rempel (1989) suggest that the working models of high-trust individuals (e.g., their trust-related hopes and fears vis-à-vis the current partner and relationship) should be more unified and better integrated than those of medium- or low-trust persons. This greater integration might result from several cognitive mechanisms, including limiting or dispelling the implications of negative relationship actions or events when they occur (Rempel, 1987), altering the meaning of negative actions/events so positive perceptions of the partner and relationship can be maintained (e.g., using "yes, but . . . " refutations; Murray & Holmes, 1993), or framing negative actions/events within the broader constellation of past positive relationship experiences (Murray & Holmes, 1993). Holmes and Rempel (1989) also surmise that greater trust should enhance felt security, which might liberate high-trust individuals from having to monitor their partners and relationships and could motivate them to err on the side of benevolence when interpreting ambiguous partner behaviors. In addition, high-trust individuals may have higher thresholds for perceiving relationship threats, they might experience or construe negative events as less aversive, or they may recover from adverse relationship experiences more quickly (see Holmes & Rempel, 1989).

Individuals involved in relationships that contain moderate levels of trust (termed "medium-trust" individuals)

have less coherent and more fragmented working models in which trust-relevant hopes and fears are intermingled in strange and sometimes contradictory ways (Holmes & Rempel, 1989). These individuals are often trapped in strong approach–avoidance situations in which positive partner behaviors are viewed as hopeful signs of possible relationship improvement but any negative behavior is taken as clear evidence that relationship dissolution could be imminent. For this reason, medium-trust individuals continually monitor and test for signs of their partner's care, concern, and responsiveness. Ironically, this hypervigilance may lead medium-trust individuals to perceive or unwittingly create the negative relationship outcomes they wish to avoid, given their likely overreliance on the diagnostic value of negative relationship information (see Reeder & Brewer, 1979). Moreover, when medium-trust individuals recall positive relationship events from the past, they judge their partner's behavior positively yet make cynical attributions about the motives behind their partner's actions (Holmes & Rempel, 1986; Rempel et al., 2001). In other words, medium-trust individuals superficially acknowledge their partner's positive actions but fail to infer that their partner's motives might be constructive and benevolent. As a result, positive partner actions may trigger latent worries about what could eventually "go wrong" in the minds of medium-trust people, a process that might impede or derail the development of what they crave— deeper intimacy and greater security.

Thus, contrary to high-trust individuals, who probably adopt a promotion focus toward relationships (Gable & Reis, 2001), medium-trust people adopt a risk-averse, prevention-focused strategy, one that is revealed by their chronic feelings of vulnerability, their incessant worries about mistakenly making positive attributions for their partner's actions, and their lingering concerns about eventually being hurt by their partners. These defensive, self-protective tendencies should reduce the likelihood that constructive partner behaviors will be accepted as evidence of the partner's genuine care and concern. To compound matters, medium-trust individuals also have more poorly integrated views of their partners and relationships (Holmes & Rempel, 1989). Memories of hurts and fears in the minds of medium-trust individuals, therefore, could be more sharply segregated from memories of positive relationship experiences. In addition, medium-trust individuals are likely to possess negative relationship models and memories that are more easily activated, and they may experience more volatile emotional swings in their relationships across time, perhaps owing to the greater compartmentalization or the haphazard connections that define their working models (see Mikulincer, 1998).

Individuals involved in relationships characterized by low trust (termed "low-trust" persons) have minimal confidence in their partner's care and responsiveness. These individuals are convinced that their partners will not take their needs or concerns into account when important decisions are made, and they are least likely to attribute their partner's positive actions to either benevolent or altruistic motives (Holmes & Rempel, 1989). Instead, low-

trust people exhibit distress-maintaining attributions (Holtzworth-Munroe & Jacobson, 1985), accentuating the ramifications of negative partner behaviors and ignoring or downplaying positive partner behaviors. When trying to resolve relationship problems, low-trust individuals consciously report having more specific (i.e., less global) neutral or slightly negative attributions regarding their partner's actions. However, more indirect measures suggest that they actually harbor strong negative attributions about their partner's real underlying motives (Rempel et al., 2001). These destructive attributional patterns may sustain the cynical and suspicious working models of low-trust persons. Because they believe that their partners cannot be trusted and will never be responsive to their basic needs, low-trust individuals rarely take the interpersonal risks necessary to build greater trust and felt security. In fact, they typically disengage or withdraw from interactions in which greater trust could be forged (such as in initially noncorrespondent situations with their partners; Rempel et al., 2001), possibly to protect themselves from the anguish of further rejection.

Testing the dependency regulation model (Murray, Holmes, & Griffin, 2000), Murray and colleagues (2005) have confirmed that, before they can fully trust their partners, low-self-esteem people may first need to either view themselves more positively *or* view their partners less positively to believe that their partner's love and benevolent actions are genuine, justified, and valid. Specifically, when low-self-esteem individuals are given feedback that they are more desirable as romantic partners than most people *or* that their current partners have treated them badly, they report greater felt security and more positive views of themselves *and* their romantic partners. Furthermore, the connection between the low-self-esteem/negative feedback experimental condition and improved self-perceptions is mediated by enhanced felt security. These psychological processes may play a critical role in nurturing trust.

Extending Rusbult's (1983) investment model, Wieselquist and colleagues (1999) have found that the degree of dyadic trust in relationships is likely to be contingent on aspects of the self, the partner, *and* the relationship (see Figure 25.2). In two longitudinal studies, they have shown that greater trust is concurrently associated with greater dependence on the partner, which in turn forecasts greater concurrent commitment. Heightened commitment concurrently predicts the enactment of more prorelationship behaviors (e.g., accommodation and personal sacrifice), which is perceived by the partner. These partner perceptions are then linked to a string of positive relations between the partner's level of trust, dependence, commitment, and prorelationship actions, which subsequently predict the *partner's* perceptions of these actions.

Wieselquist and colleagues (1999) propose that perceiving greater prorelationship accommodation and sacrifice by one's partner might be responsible for promoting trust. Indeed, when individuals perceive that their partners have experienced a major transformation away from purely self-interested outcomes (MaxOwn) toward partner-centered outcomes (MaxOther) or relationship-centered outcomes (MaxJoint), they do report greater trust (see also Yovetich & Rusbult, 1994). If, however, one partner has an anxious–ambivalent attachment style (i.e., chronically worries about whether partners love or will stay with him or her), lower levels of trust tend to be reported by *both* relationship partners. Wieselquist and colleagues also suggest that partners may occasionally use "diagnostic" or "strain test" situations to test the strength of their partner's prorelationship motives or demonstrate their *own* prorelationship motives. Trust is less likely to develop or grow if partners never experience these situations, if one or both partners fail these critical tests by not showing expected transformation tendencies, or if one or both partners do not perceive or fail to infer the other's genuinely positive, prorelationship motives. Wieselquist and colleagues also conjecture that trust may be more likely to develop and grow if mutual (similar) levels of vulnerability, mutual changes in commitment, and reciprocal positive behaviors are exchanged between partners. Sometimes, however, trust may grow stronger if partners do *not* reciprocate evenly, at least over short time periods (see Drigotas, Whitney, & Rusbult, 1995).

Miller and Rempel (2004) have extended this line of work by testing whether charitable attributions might be responsible for increasing dyadic trust. Noting that Wieselquist and colleagues (1999) found little evidence for which variables predict actual *changes* in trust over time in relationships, Miller and Rempel conjecture that partner-enhancing attributions regarding the motives behind a partner's relationship-relevant actions and behaviors might predict actual increases in dyadic trust. Studying married couples, they found that partner-enhancing attributions (e.g., the belief that one's spouse has good intentions when discussing a conflict) predicted significant gains in trust 2 years later. Increases in trust also predicted increases in the tendency to make partner-enhancing attributions, suggesting that both processes may operate in a mutually reinforcing fashion. Changes in trust, therefore, do not appear to be driven by attributions about a partner's specific interaction behaviors; rather, they arise from the more global attributions that individuals make regarding the *motives* that guide a partner's habitual relationship-relevant actions.

Attachment Theory–Based Approaches

Additional research on trust has been informed by attachment theory. The earliest work linking adult attachment styles to dyadic trust (e.g., Simpson, 1990) indicated that more securely attached people (who have positive models of themselves and relationship partners) report greater dyadic trust than either more avoidantly attached people (who have variable self-views and predominately negative views of partners) or anxiously attached people (who have negative self-views and hopeful yet guarded views of partners). More secure people also sustain higher levels of trust in their relationships across time than do their insecure counterparts (Keelan, Dion, & Dion, 1994). The working models of more securely attached individuals may partially explain their stronger in-

clination to trust others. Baldwin, Fehr, Keedian, Seidel, and Thompson (1993), for example, have found that more secure people expect hypothetical partners to behave more benevolently in situations that call for trust. They also respond more rapidly to words connoting positive relationship outcomes (e.g., care), whereas more insecure individuals react more quickly to words suggesting negative relationship outcomes (e.g., hurt). Highly secure individuals also possess more available and accessible positive relationships exemplars, whereas highly avoidant and highly anxious individuals have more available and accessible negative relationship exemplars (Baldwin, Keelan, Fehr, Enns, & Koh-Rangarajoo, 1996). When individuals feel more secure in a given relationship, they report heightened trust, above and beyond any effects associated with their chronic (dispositional) attachment style. Pistole (1994) suggests that the lower trust displayed by more insecurely attached people might also be attributable to how these individuals regulate important social interactions. Highly anxious individuals, who crave emotional closeness and intimacy, may smother their partners, unwittingly pushing them away. Highly avoidant individuals, who desire self-reliance and emotional independence, may erect barriers that create and maintain emotional distance, a tendency that could launch or exacerbate destructive "distance/pursuer" interaction cycles.

The most detailed research linking adult attachment styles and dyadic trust has been conducted by Mikulincer. Mikulincer (1998) claims that confidence in and positive expectations of a partner's availability and responsiveness are core elements not only of dyadic trust but of secure working models as well. He argues that in response to the quality of previous care and support they have received (or *perceive* they have received), individuals with different attachment styles should have different trust-related goals. More securely attached persons ought to organize their interpersonal behavior around creating and maintaining greater closeness and intimacy in their relationships, partly because they do not have to worry about attaining felt security. Achieving greater intimacy may, in fact, be essential for developing greater trust. According to Mikulincer, the tendency of securely attached individuals to be more trusting should facilitate their quest for greater intimacy because heightened trust should serve as an "internalized secure base" that permits more secure people to take the interpersonal risks and leaps of faith necessary to build deeper trust. Many of these risks are likely to include the disclosure of personal goals, hopes, feelings, and fears.

The central interaction goals of insecurely attached individuals, by comparison, should mirror the defenses that insecure people use to combat attachment-related distress (Bowlby, 1988). Given their unrelenting concerns about receiving deficient care, attention, and support (Bowlby, 1973), highly anxious people should adopt interaction goals that are geared toward attaining what they desire the most—greater felt security. In light of their prolonged history of rejection and subsequent desire to avoid emotional entanglements (Crittenden & Ainsworth, 1989), highly avoidant people should adopt

goals to attain what they ultimately desire—gaining and maintaining control and autonomy in relationships.

Mikulincer (1998) has provided compelling evidence for the existence and operation of these goals in people who have different attachment styles. More securely attached people do have more accessible memories of trust-validating experiences, they feel more positive affect when remembering these experiences, and they report more trust-validating experiences. They also place greater importance on trust-validating experiences and view them as being attributable to the benevolent traits and motives of their partners. More secure individuals also talk more openly with their partners about trust-violation events, they have less accessible negative trust-related memories, they react less negatively to negative memories, and they do *not* attribute trust violations to their partner's stable dispositions or enduring motives (i.e., they make situation-specific and transient attributions for negative partner behaviors). Although more secure individuals tend to be involved in more satisfying relationships with perhaps better adjusted romantic partners than many insecure individuals (Simpson, 1990), the cognitive and emotional tendencies documented by Mikulincer ought to accentuate positive views of partners and relationships and curtail negative ones, possibly fueling greater trust.

More insecure individuals, in contrast, report less trust and have more accessible negative trust-related memories (Mikulincer, 1998). In addition, they view trust as a means to achieve security-based goals (in the case of highly anxious people) or control/autonomy-based goals (in the case of highly avoidant people). More insecure individuals also exhibit more maladaptive coping strategies when distressed. Highly anxious persons, for instance, use more emotion-focused coping strategies, whereas highly avoidant persons use more avoidance/distancing coping strategies. Finally, highly anxious individuals respond to negative trust-related experiences with greater negative affect and place more weight on these negative experiences.

Other Theoretical and Empirical Approaches

There have been other notable studies of dyadic trust. Several, for example, have documented that trust tends to be high during the initial stages of relationship development (Fletcher et al., 2000), after which it gradually declines once the honeymoon phase is over (Larzelere & Huston, 1980). What is not known, however, is whether relationship partners define and weigh trust similarly at different stages of relationship development. Other research has indicated that individuals who engage in more self-disclosure and communicate greater trust often have partners who are more trusting (Butler, 1986; Haas, 1981; Larzelere & Huston, 1980). Trust may be especially facilitated by the combination of intimate self-disclosures and reciprocated trust from partners (Deutsch, 1960; Ellison & Firestone, 1974; McAllister & Bergman, 1986), though some research has questioned the importance of reciprocity (Bartle, 1996). Other lines of work have discovered small but stable gender differences in dyadic

trust. Women typically score slightly higher than men on trust measures (Butler, 1986; Canary & Cupach, 1988), perhaps because certain family-of-origin experiences might affect women and men differently (see Cooper & Grotevant, 1987). These gender differences might also stem from how men and women are socialized in Western cultures. Given that most men are taught to value independence and autonomy (Block, 1983; Gilligan, 1982), many men may eschew the greater dependence on others that higher levels of trust entail. Given that most women are socialized to be interdependent and relationship oriented, many women may feel more comfortable entering and maintaining more dependent relationships that involve greater trust.

Work conceptualizing trust from family systems theory (Bowen, 1976, 1978) suggests that women who are less empathetic, who are poorer at solving relationship problems, who are less likely to take responsibility for their actions, and whose families-of-origin did not express thoughts and emotions accurately tend to be less trusting (Reeves & Johnson, 1992). In addition, Bartle (1996) has found that greater emotional and behavioral reactivity to emotion-evoking events involving one's *parents* predicts less dyadic trust of one's current romantic partners. Men's trust, however, correlates most highly with their level of emotional/behavioral reactivity (such that less reactive men report the most trust), whereas women's trust is most strongly tied to their comfort with making intimate self-disclosures to significant others. Borrowing principles of family systems theory, some research insinuates that the *lack* of sufficient self-differentiation from significant others (e.g., parents, romantic partners, and close friends) may foreclose the kinds of social interactions needed to develop stronger trust (Ryder & Bartle, 1991). According to this view, if an individual's self-concept is more diversified, he or she should be more willing to risk placing deeper confidence and faith in others. If relationships fail, the self-concepts of these "well-diversified" individuals should be buoyed by the other sources that define and contribute to their sense of self-worth (see Crocker & Wolfe, 2001).

The most recent wave of research on dyadic trust is starting to identify the specific brain structures and processes implicated in the psychological experience of trust. The amygdala, one of the primary centers for processing emotions, becomes more active when people view faces that they later rate as untrustworthy relative to faces they rate as more trustworthy (Winston, Strange, O'Doherty, & Dolan, 2002). Although little is known about the facial cues that signal untrustworthiness, some evidence suggests that expressions of anger and sadness or the absence of happiness might be systematically tied to perceptions of untrustworthiness, at least from photographs (Adolphs, 2002). In addition, bilateral damage to the amygdala tends to impair judgments of untrustworthy faces (Adolphs, Tranel, & Damasio, 1998). Adolphs (2002) suggests that seeing untrustworthy people may trigger automatic, largely unconscious processing in the amygdala, which launches more conscious processing in the orbitofrontal area of the brain. This secondary processing may then generate the cognitive and somatic changes that individuals sense as negative emotional responses.

Other recent research has revealed that people are more likely to trust others who resemble them in physical appearance (DeBruine, 2002), and they systematically reward partners who trust them and punish those who do not (Fehr & Rockenbach, 2003). Though preliminary, some research suggests that levels of the hormone oxytocin rise when individuals who are playing monetary games against others receive information that their interaction partners trust them (Zak, 2003). Individuals who have higher baseline levels of oxytocin are also more likely to reciprocate signs of trust, and women who are ovulating are *less* inclined to trust others compared to when they are not ovulating (Zak, 2003). When oxytocin is administered experimentally to people, they are more willing to trust strangers (but not machines) in trust games that involve real monetary stakes (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005).

In conclusion, a reasonable amount of research, much of which has been informed by interdependence theory or attachment theory, has investigated various aspects of trust in close relationships. The first wave of seminal work was structured around single-trial PDG experimental paradigms involving strangers (e.g., Deutsch, 1960). Largely in response to problems with this research, the second wave of research conceptualized trust as a fairly stable disposition relevant to others in general. The most recent wave of work, which was launched in the 1980s, has conceptualized trust in more dyadic (interpersonal) and process-oriented terms (e.g., Holmes & Rempel, 1989; Mikulincer, 1998; Wieselquist et al., 1999). Most of the extant research, however, has focused on individual differences in trust (e.g., the way in which individuals who differ in the degree of trust within a relationship differentially think, feel, and behave in trust-relevant situations) rather than the more basic normative processes that, over time, might promote or inhibit the development of trust in relationships. Future research needs to test and delve more deeply into the normative components of models that attempt to explain the growth and decline of trust in established relationships across time.

MAJOR PRINCIPLES AND A DYADIC MODEL OF TRUST IN RELATIONSHIPS

One of the goals of the current volume is to identify overarching principles that capture the core essence of a given psychological construct or set of processes. Based on the preceding review, at least six fundamental principles of trust are apparent.

- *Principle 1*: The desire to view partners positively and to want to trust them (i.e., enhancement motives) plays a larger role in guiding relationship and social interaction processes during the early phases of relationship development.
- *Principle 2*: The desire to view partners more accurately and objectively (i.e., accuracy motives) becomes more important in guiding relationship and social inter-

action processes in latter stages of relationship development.

- *Principle 3*: Individuals assess the degree to which they can trust their partners by observing whether partners display proper transformation of motivation in *trust diagnostic* situations (e.g., in potentially noncorrespondent "trust" or strain test situations in which partners make decisions that work against their own self-interest and support the best interests of the individual and/or the relationship).

- *Principle 4*: Individuals may at times intentionally enter or create trust diagnostic situations to test whether their current level of trust in a partner is warranted.

- *Principle 5*: Individual differences associated with attachment processes, self-differentiation, and/or self-esteem are likely to influence the growth or decline of trust at different normative stages over time in relationships (see the model presented below).

- *Principle 6*: Neither the level nor the trajectory of trust in relationships can be fully understood without considering the dispositions and actions of *both* relationship partners, especially in trust diagnostic situations.

These principles are by no means exhaustive. However, they capture six of the most central themes that anchor the major theoretical and empirical work that has focused on interpersonal trust.

Several theories and models of trust have been reviewed. How might their major features be integrated to highlight the most important situational and psychological processes through which trust is likely to develop and be maintained in close relationships? Figure 25.3 depicts one way of conceptualizing how some of the core constructs reviewed earlier in the chapter might be interconnected in trust-relevant social interactions. The dyadic model of trust in relationships contains both normative and individual difference components. The normative part of the model is shown in the five boxes (constructs) that form the middle of the figure. The individual difference portion of the model is displayed in the dispositions of *each* relationship partner and her or his respective links to each normative construct. To keep the figure simple, feedback loops from the terminal normative construct (*each partner's* perception of felt security) to the construct that should launch future trust-relevant interactions (*each partner's* decision about whether to enter the next trust-relevant situation) are not depicted, but they are presumed to exist. The model holds that, during interactions, each individual's perceptions of both his or her own *and* the partner's standing on each construct is necessary to explain what transpires for each partner in the downstream constructs of the model.

The model assumes that information about the dispositions of *both* partners (e.g., their attachment styles and working models, their typical relationship decision-making tendencies, and their attribution styles) is necessary to explain the growth of trust—or the lack thereof—in a given relationship across several interactions. Each
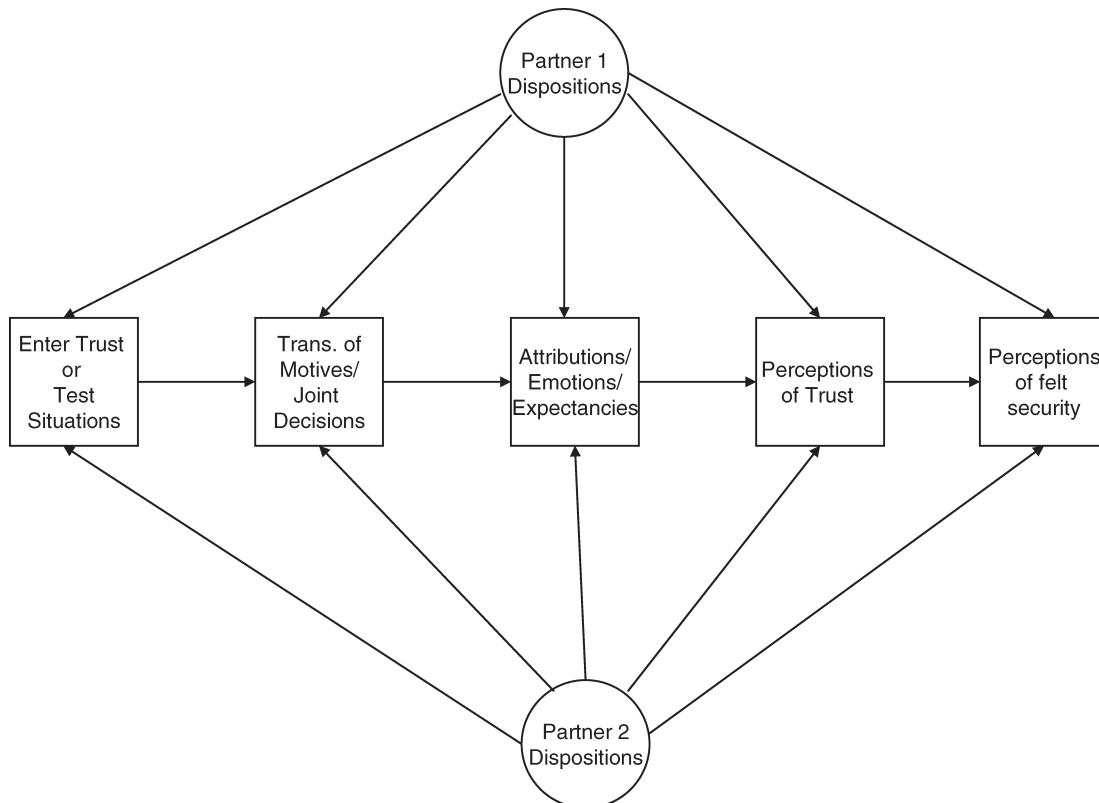
FIGURE 25.3. The dyadic model of trust in relationships.

partner enters the relationship with myriad personal and interpersonal experiences, some of which have shaped his or her attachment style and working models, his or her self-esteem, or the degree to which he or she has a differentiated self-concept. These dispositional attributes are important because they might affect what occurs in downstream components of the model, given each individual's own attributes (e.g., his or her own attachment style) and his or her *partner's* attributes (e.g., the partner's attachment style).

Dispositional tendencies such as greater attachment security, a more differentiated self-concept, and/or higher self-esteem may motivate and enable individuals to enter, elicit, or transform social interactions into ones that foster greater trust across time. Two types of situations should give individuals a particularly good opportunity to assess the degree of trust that is warranted in the partner/relationship at a specific point in time: (1) "trust situations" (Kelley et al., 2003) in which partners can repeatedly make or fail to make a1/b1 decisions, and (2) "strain test" situations (Holmes, 1981; Kelley, 1983) in which partners can demonstrate or fail to demonstrate their willingness to make personal sacrifices for the good of the partner or the relationship. For these *trust diagnostic* situations to be entered or created, however, one or sometimes both partners must have sufficient confidence to take the interpersonal risks required to confirm or reaffirm that the partner truly *can* be trusted. Individuals and partners who have more secure, more self-differentiated, and/or more positive self-views should be in a better position to take these risks and leaps of faith on a regular basis.

Once individuals find themselves in trust diagnostic situations, those who routinely display the transformation of motivation needed to carry out MaxOther or MaxJoint decisions in everyday trust situations or occasional strain test situations should be on the path toward experiencing greater trust and felt security. Moreover, individuals and partners who have more secure, more self-differentiated, and/or more positive self-views may engage in these transformations more often and more extensively. Given their stronger sense of security and greater promotion focus, these individuals should be more motivated to—and perhaps more capable of— steering trust-relevant social interactions toward a1/b1 decisions. Once such decisions have been made, each individual's working models may then influence how he or she interprets the extent of transformation that both the self *and* the partner have undergone. As a consequence, highly secure, differentiated, and/or self-confident individuals may grant themselves and their partners "fuller credit" for each partner's willingness to maximize partner or relationship outcomes instead of each partner's more egocentric self-interests.

This process, in turn, may trigger more benevolent attributions regarding one's own as well as the partner's core relationship motives (Miller & Rempel, 2004). These positive attributions may then facilitate more constructive problem solving, more adaptive emotion regulation, and more optimistic expectations about what is likely to happen in future trust diagnostic interactions.

These outcomes might also be affected by each partner's degree of attachment security, self-differentiation, and/or self-esteem. Patterns of attribution, emotion regulation, and situation-specific expectancies are displayed in a single box in the model because the temporal order of these processes is likely to differ depending on idiosyncratic features of the interaction (e.g., the specific topic being discussed), the dispositions of each partner, or unique emergent properties of the relationship itself. These positive outcomes should then increase perceptions of trust, which ought to make individuals feel more secure, at least temporarily (Holmes & Rempel, 1989). Once again, greater security, more self-differentiation, and/or higher self-esteem might facilitate perceptions of greater trust and felt security. Perceptions of felt security may then "set up" the next trust-relevant interaction, determining whether or when the next trust situation or strain test is entered.

Across time, individuals who repeatedly experience a1/b1 outcomes with their partners should begin to perceive greater "added value" following these outcomes, especially when they occur in trust diagnostic situations. These mutually beneficial outcomes may also encourage individuals and their partners to perform additional relationship-maintenance acts (e.g., disparagement of alternatives, accentuation of partner or relationship attributes, or heightened cognitive interdependence) that might facilitate or reinforce a1/b1 decisions in the future. These effects should diminish, of course, if one or both partners have dispositional vulnerabilities (e.g., an anxious attachment style; Wieselquist et al., 1999), fail to display proper transformation, repeatedly decide *not* to enter a1/b1 agreements, or make chronically negative attributions regarding their partner's basic relationship motives.

It is important to emphasize that this model does *not* presume that dispositions are chiefly responsible for generating, sustaining, or short-circuiting trust in relationships. Certain dispositions most likely amplify or qualify the basic normative processes outlined in the middle section of the model. Individuals who are insecurely attached, have weak self-differentiation, or have low self-esteem can and do form trusting relationships. To achieve higher levels of trust, however, such persons may need to either (1) be involved with partners who help them suppress or change their destructive perceptual, emotional, and behavioral tendencies or (2) have unusually committed or mutually rewarding relationships that enable these "vulnerable" individuals and their partners to overcome their liabilities.

There is an interesting paradox surrounding the transformation of motivation and trust as relationships grow and develop. Early in relationship development, large amounts of transformation by both partners should be a good barometer of the amount of trust that is warranted in a partner/relationship, particularly when partners are equally dependent on one another for unique or important outcomes (Holmes & Rempel, 1989) or have similar levels of vulnerability and commitment (Wieselquist et al., 1999). As relationships progress, however, most individuals incorporate their partners and relationships

more centrally into their self-concepts (Aron et al., 2001). This means that *less* transformation of motivation is likely to occur in well-established relationships. For this reason, partners in long-term relationships occasionally may need to identify *new* trust diagnostic issues or situations in which each partner's self-interests start out being slightly discrepant from what might be best for his or her partner or the relationship. If partners can jointly demonstrate that they are still willing to undergo large transformations in new trust diagnostic situations, this might reaffirm and sustain trust in well-established pairs.

As discussed at the beginning of the chapter, our knowledge of how trust is generated, sustained, and compromised in close relationships remains surprisingly limited. Future theory and research on dyadic trust might profit from addressing the following issues. First, we need to gain a better understanding of how the dispositions and behaviors of relationship *partners* affect how individuals (actors) think, feel, and behave in trust-relevant situations. Very few if any studies have investigated whether or how "partner effects" might affect these processes. Second, research should explore how and why certain within-dyad combinations of partner attributes (e.g., very different levels of power, dependence, or commitment in a relationship) promote or impede the development and maintenance of trust. Of particular importance will be learning whether, how, and why certain highly asymmetrical relationships (e.g., those in which partners have vastly different levels of power) occasionally generate very high levels of trust (see Drigotas et al., 1995). Third, research should examine whether the normative component of the model shown in Figure 25.3 holds with equal effect at different stages of relationship development (see Holmes & Rempel, 1989). Fourth, more needs to be understood about how normative processes and individual differences reciprocally influence one another across time in trust-relevant situations. For simplicity, the model depicted in Figure 25.3 does not show that normative processes may, over time, shape and revise dispositions within individuals, yet such reciprocal feedback is bound to take place. Fifth, greater attention should be devoted to understanding the events depicted in the middle box of the model (attributions/emotions/expectancies). Features critical to the development of intimacy in relationships, such as feeling understood, valued, and cared for by one's partner, may also promote trust and felt security (Deutsch, 1973; Mikulincer, 1998; Reis & Shaver, 1988), especially when these perceptions repeatedly occur in trust-relevant interactions. Sixth, more needs to be known about what "added value" entails and how it is generated. Though the prospect of receiving future rewards from one's partner may make joint a1/b1 decisions feel unusually good (Huesmann & Levinger, 1976), more than the prospect of rewards is likely to be involved in generating high "added value" perceptions.

In conclusion, there are few constructs in the field of interpersonal relationships that are more central or important to relationship functioning and outcomes than trust. Without trust, voluntary relationships are not likely to develop, let alone grow or be maintained. Experiencing trust in close relationships may be more important today than ever before when one considers that most Americans have fewer and less meaningful ties to friends, organizations, and community activities than was true 30 years ago (Putnam, 2000). This pattern of declining social capital may be exerting further pressure on people to establish stronger emotional connections, greater dependency, and deeper levels of trust with their intimate partners. Trust may, in fact, be a "cardinal construct" (cf. Asch, 1946) in relationships to the extent that it transforms the meaning of other accompanying attributes and descriptors of a relationship. The fact that so little is known about such a pivotal construct attests to the difficulty of operationalizing, measuring, and studying trust in close relationships. For the science of relationships to advance, however, we must develop a deeper and more sophisticated understanding of how trust emerges, operates, changes, and declines within close relationships.

ACKNOWLEDGMENT

REFERENCES

Adolphs, R. (2002). Trust in the brain. *Nature Neuroscience*, *5*, 192–193.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, *393*, 470–474.

Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine de Gruyter.

Allison, M. D., & Sabatelli, R. M. (1988). Differentiation and individuation as mediators of identity and intimacy in adolescence. *Journal of Adolescent Research*, *3*, 1–16.

Aron, A., Aron, E. N., & Norman, C. (2001). Self-expansion model of motivation and cognition in close relationships and beyond. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 478–501). Malden, MA: Blackwell.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Baldwin, M. W., Fehr, B. Keedian, E., Seidel, M., & Thompson, D. W. (1993). An exploration of the relational schemata underlying attachment styles: Self report and lexical decision approaches. *Personality and Social Psychology Bulletin*, *19*, 746–754.

Baldwin, M. W., Keelan, J. P. R., Fehr, B., Enns, V., & Koh-Rangarajoo, E. (1996). Social-cognitive conceptualization of attachment working models: Availability and accessibility effects. *Journal of Personality and Social Psychology*, *71*, 94–109.

Barber, B. (1983). *The logic and limits of trust*. New Brunswick, NJ: Rutgers University Press.

Bartle, S. E. (1996). Family of origin and interpersonal contributions to the interdependence of dating partners' trust. *Personal Relationships*, *3*, 197–209.

Bartle, S. E., & Sabatelli, R. M. (1995). The Behavioral and Emotional Reactivity Index: Preliminary evidence for construct validity from three studies. *Family Relations*, *44*, 267–277.

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, *62*, 647–670.

Block, J. H. (1983). Differential premises arising from differential socialization of the sexes. *Child Development*, *54*, 1335–1354.

Bowen, M. (1976). Theory and practice in psychotherapy. In P. J. Guerin (Ed.), *Family therapy: Theory and practice* (pp. 42–90). New York: Gardner Press.

Bowen, M. (1978). *Family therapy in clinical practice*. New York: Jason Aronson.

Bowlby, J. (1969). *Attachment and loss: Volume 1. Attachment*. New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss: Volume 2. Separation: Anxiety and anger*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and loss: Volume 3. Loss*. New York: Basic Books.

Bowlby, J. (1988). *A secure base: Parent–child attachment and healthy human development*. New York: Basic Books.

Brewer, M. B., & Caporael, L. R. (1990). Selfish genes versus selfish people: Sociobiology as origin myth. *Motivation and Emotion*, *14*, 237–243.

Butler, J. K. (1986). Reciprocity of dyadic trust in close male–female relationships. *Journal of Social Psychology*, *126*, 579–591.

Canary, D. J., & Cupach, W. R. (1988). Relational and episodic characteristics associated with conflict tactics. *Journal of Social and Personal Relationships*, *5*, 305–325.

Carson, R. C. (1979). Personality and exchange in developing relationships. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships* (pp. 247–269). New York: Academic Press.

Chisholm, J. S. (1993). Death, hope, and sex: Life-history theory and the development of reproductive strategies. *Current Anthropology*, *34*, 1–24.

Chisholm, J. S. (1996). The evolutionary ecology of attachment organization. *Human Nature*, *7*, 1–38.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York: Academic Press.

Clutton-Brock, T. H., & Parker, G. A. (1995). Punishment in animal societies. *Nature*, *373*, 209–216.

Comrey, A. L., & Schiebel, D. (1983). Personality test correlates of psychiatric outpatient status. *Journal of Consulting and Clinical Psychology*, *51*, 757–762.

Cook, K. S., & Cooper, R. M. (2003). Experimental studies of cooperation, trust, and social exchange. In E. Ostrom & J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 209–244). New York: Russell Sage.

Cooper, C. R., & Grotevant, H. D. (1987). Gender issues in the interface of family experience and adolescents' friendship and dating identity. *Journal of Youth and Adolescence*, *16*, 247–264.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163–228). New York: Oxford University Press.

Crittenden, P. M., & Ainsworth, M. (1989). Child maltreatment and attachment theory. In D. Cicchetti & V. Carlson (Eds.), *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect* (pp. 432–463). Cambridge, UK: Cambridge University Press.

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593–623.

DeBruine, L. M. (2002). Facial resemblance enhances trust. *Philosophical Transactions of the Royal Society of London, B* [Online]. Available at: doi:10.1098/rspb.2002.2034.

Deutsch, M. (1960). The effect of motivational orientation upon trust and suspicion. *Human Relations*, *13*, 123–139.

Deutsch, M. (1973). *The resolution of conflict*. New Haven, CT: Yale University Press.

Drigotas, S. M., Whitney, G. A., & Rusbult, C. E. (1995). On the peculiarities of loyalty: A diary study of responses to dissatisfaction in everyday life. *Personality and Social Psychology Bulletin*, *21*, 596–609.

Eibl-Eibesfelt, I. (1989). *Human ethology*. New York: Aldine de Gruyter.

Ellison, C. W., & Firestone, I. J. (1974). Development of interpersonal trust as a function of self-esteem, target status, and target style. *Journal of Personality and Social Psychology*, *29*, 655–663.

Erickson, E. (1963). *Childhood and society*. New York: Norton.

Fehr, B. (1988). Prototype analysis of the concepts of love and commitment. *Journal of Personality and Social Psychology*, *55*, 557–579.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–791.

Fehr, E., & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, *422*, 137–140.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, *114*, 817–868.

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). Ideals, perceptions, and evaluations in early relationship development. *Journal of Personality and Social Psychology*, *79*, 933–940.

Foley, R. (1987). *Another unique species: Patterns in human evolutionary ecology*. New York: Longman/Wiley.

Folkman, S., Lazarus, R. S., Gruen, R. J., & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of Personality and Social Psychology*, *50*, 571–579.

Gable, S. L., & Reis, H. T. (2001). Appetitive and aversive social interaction. In J. H. Harvey & A. E. Wenzel (Eds.), *Close romantic relationship maintenance and enhancement* (pp. 169–194). Mahwah, NJ: Erlbaum.

Gaines, S. O., Jr., Panter, A. T., Lyde, M. D., Steers, W. N, Rusbult, C. E., Cox, C. L., et al. (1997). Evaluating the circumplexity of interpersonal traits and the manifestation of interpersonal traits in interpersonal trust. *Journal of Personality and Social Psychology*, *73*, 610–623.

Gilligan, C. (1982). *In a different voice: Psychological theory of women's development*. Cambridge, MA: Harvard University Press.

Gintis, H. (2003). The hitchhiker's guide to altruism: Gene–culture co-evolution and the internalization of norms. *Journal of Theoretical Biology*, *220*, 407–418.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, *213*, 103–119.

Gurtman, M. B. (1992). Trust, distrust, and interpersonal problems: A circumplex analysis. *Journal of Personality and Social Psychology*, *62*, 989–1002.

Haas, D. R. (1981). Trust and symbolic exchange. *Social Psychology Quarterly*, *44*, 3–13.

Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology*, *7*, 1–52.

Hardin, R. (2003). Gaming trust. In E. Ostrom & J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 80–101). New York: Russell Sage.

Henrich, J., & Boyd, R. (2001). Why people punish defectors—Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, *208*, 79–89.

Hill, K. (2002). Altruistic cooperation during foraging by the Ache, and evolved human predisposition to cooperate. *Human Nature*, *13*, 105–128.

Hinde, R. A. (1986). Some implications of evolutionary theory and comparative data for the study of human prosocial and aggressive behavior. In D. Olweus, J. Block, & M Radke-Yarrow (Eds.), *Development of antisocial and prosocial behavior* (pp. 13–32). Orlando, FL: Academic Press.

Holmes, J. G. (1981). The exchange process in close relationships: Microbehavior and macromotives. In M. J. Lerner & S. C. Lerner (Eds.), *The justice motive in social behavior* (pp. 261–284). New York: Plenum Press.

Holmes, J. G. (1991). Trust and the appraisal process in close relationships. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 2, pp. 57–104). London: Jessica Kingsley.

Holmes, J. G., & Rempel, J. K. (1986). *Trust and conflict in close relationships*. Invited address at the meeting of the American Psychological Association, Washington, DC.

Holmes, J. G., & Rempel, J. K. (1989). Trust in close relationships. In C. Hendrick (Ed.), *Close relationships* (pp. 187–220). Newbury Park, CA: Sage.

Holtzworth-Munroe, A., & Jacobson, N. S. (1985). Causal attributions of married couples: When do they search for causes? What do they conclude when they do? *Journal of Personality and Social Psychology*, *48*, 1398–1412.

Huesmann, L. R., & Levinger, G. (1976). Incremental exchange theory: A formal model for progression in dyadic social interaction. In L. Berkowitz & E. Walster (Eds.), *Advances in experimental social psychology* (Vol. 9, pp. 151–193). New York: Academic Press.

Huston, T. (2000). The social ecology of marriage and other intimate unions. *Journal of Marriage and the Family*, *62*, 298–320.

Johnson-George, C., & Swap, W. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific order. *Journal of Personality and Social Psychology*, *43*, 1306–1317.

Keelan, J. P. R., Dion, K. L., & Dion, K. K. (1994). Attachment style and heterosexual relationships among young adults: A short term panel study. *Journal of Social and Personal Relationships*, *11*, 201–214.

Kelley, H. H. (1980). The situational origins of human tendencies: A further reason for the formal analysis of structures. *Personality and Social Psychology Bulletin*, *9*, 8–30.

Kelley, H. H. (1983). Love and commitment. In H. H. Kelley, E. Berscheid, A. Christensen, J. H. Harvey, T. L. Huston, G. Levinger, et al. (Eds.), *Close relationships* (pp. 265–314). New York: Freeman.

Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. M. (2003). *An atlas of interpersonal situations*. New York: Cambridge University Press.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*, 66–91.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relationships: A theory of interdependence*. New York: Wiley.

Kelly, R. L. (1995). *The foraging spectrum: Diversity in hunter–gatherer lifeways*. Washington, DC: Smithsonian Institution Press.

Konnor, M. (1982). *The tangled wing: Biological constraints on the human spirit*. New York: Holt, Rinehart & Winston.

Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*, 673–676.

Kramer, R. M., & Carnevale, P. J. (2001). Trust and intergroup negotiation. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp. 431–450). Malden, MA: Blackwell.

Kruglanski, A. W. (1971). Attributing trustworthiness in supervisor–worker relations. *Journal of Experimental Social Psychology*, *6*, 214–232.

Kurzban, R. (2003). Biological foundations of reciprocity. In E. Ostrom & J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 105–127). New York: Russell Sage.

Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family*, *42*, 595–604.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces*, *63*, 967–985.

MacDonald, A. P., Kessel, V. S., & Fuller, J. B. (1972). Self-disclosure and two kinds of trust. *Psychological Reports*, *30*, 143–148.

McAllister, H. A., & Bergman, N. J. (1986). Set size effect in self-disclosure. *Journal of Social Psychology*, *126*, 337–343.

McClintock, C. G. (1972). Social motivation: A set of propositions. *Behavioral Science*, *17*, 438–454.

McClintock, C. G. (1976). Social values: Their definition, measurement, and development. *Journal of Research and Development in Education*, *12*, 121–137.

Mikulincer, M. (1998). Attachment working models and the sense of trust: An exploration of interaction goals and affect regulation. *Journal of Personality and Social Psychology*, *74*, 1209–1224.

Miller, P. J. E., & Rempel, J. K. (2004). Trust and partner-enhancing attributions in close relationships. *Personality and Social Psychology Bulletin*, *30*, 695–705.

Murray, S. L., & Holmes, J. G. (1993). Seeing virtues in faults: Negativity and the transformation of interpersonal narratives in close relationships. *Journal of Personality and Social Psychology*, *65*, 707–722.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology*, *78*, 478–498.

Murray, S. L., Rose, P., Holmes, J. G., Derrick, J., Podchaski, E. J., Bellavia, G., et al. (2005). Putting the partner within reach: A dyadic perspective on felt security in close relationships. *Journal of Personality and Social Psychology*, *88*, 327–347.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573–577.

Ostrom, E., & Walker, J. (2003). Introduction. In E. Ostrom & J. Walker (Eds.), *Trust and reciprocity: Interdisciplinary lessons from experimental research* (pp. 3–19). New York: Russell Sage.

Pistole, M. C. (1994). Adult attachment styles: Some thoughts on closeness-distance struggles. *Family Process*, *33*, 147–159.

Pruitt, D. G. (1965). Definition of the situation as a determinant of international action. In H. C. Kelman (Ed.), *International behavior* (pp. 393–432). New York: Holt, Rinehart & Winston.

Pruitt, D. G., & Rubin, J. Z. (1986). *Social conflict: Escalation, statement, and settlement*. New York: Random House.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*, 61–79.

Reeves, P. C., & Johnson, M. E. (1992). Relationship between family-of-origin functioning and self-perceived correlates of eating disorders among female college students. *Journal of College Student Development*, *33*, 44–49.

Regan, P. C., Kocan, E. R., & Whitlock, T. (1998). Ain't love grand! A prototypic analysis of the content of romantic love. *Journal of Social and Personal Relationships*, *15*, 411–420.

Reis, H. T., & Shaver, P. R. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.), *Handbook of personal relationships* (pp. 367–389). Chichester, UK: Wiley.

Rempel, J. K. (1987). *Trust and attributions in close relationships*. Unpublished doctoral dissertation, University of Waterloo, Canada.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*, 95–112.

Rempel, J. K., Ross, M., & Holmes, J. G. (2001). Trust and communicated attributions in close relationships. *Journal of Personality and Social Psychology*, *81*, 57–64.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *35*, 651–665.

Rotter, J. B. (1971). Generalized expectancies of interpersonal trust. *American Psychologist*, *26*, 443–452.

Rotter, J. B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American Psychologist*, *35*, 1–7.

Rusbult, C. E. (1983). A longitudinal test of the investment model: The development (and deterioration) of satisfaction and commitment in heterosexual involvements. *Journal of Personality and Social Psychology*, *45*, 101–117.

Ryder, R. G., & Bartle, S. E. (1991). Boundaries as distance regulators in personal relationships. *Family Process*, *30*, 393–406.

Sato, K. (1988). Trust and group size in a social dilemma. *Japanese Psychological Research*, *30*, 88–93.

Shapiro, E. R. (1988). Individual change and family development: Individuation as a family process. In C. J. Falicov (Ed.), *Family transitions: Continuity and change over the family life cycle* (pp. 159–180). New York: Guilford Press.

Sherman, P. W. (1988). The levels of analysis. *Animal Behavior*, *36*, 616–619.

Simpson, J. A. (1990). Influence of attachment styles on romantic relationships. *Journal of Personality and Social Psychology*, *59*, 971–980.

Simpson, J. A. (1999). Attachment theory in modern evolutionary perspective. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 115–140). New York: Guilford Press.

Simpson, J. A., Ickes, W., & Grich, J. (1999). When accuracy hurts: Reactions of anxiously-attached dating partners to a relationship-threatening situation. *Journal of Personality and Social Psychology*, *76*, 754–769.

Sroufe, L. A., & Waters, E. (1977). Attachment as an organizational construct. *Child Development*, *48*, 1184–1199.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tinbergen, N. (1963). On the aims and methods of ethology. *Zeitschrift fur Tierpsychologie*, *20*, 410–433.

Tooby, J., & Cosmides, L. (1990). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality*, *58*, 17–68.

Tooby, J., & Cosmides, L. (1992). Psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). New York: Oxford University Press.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Tyler, T. R. (2001). Why do people rely on others? Social identity and the social aspects of trust. In K. S. Cook (Ed.), *Trust in society* (pp. 285–306). New York: Russell Sage.

Wieselquist, J., Rusbult, C. E., Foster, C. A., & Agnew, C. R. (1999). Commitment, prorelationship behavior, and trust in close relationships. *Journal of Personality and Social Psychology*, *77*, 942–966.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*, 277–283.

Wright, R. (1994). *The moral animal*. New York: Vintage.

Wrightsman, L. S. (1974). *Assumptions about human nature: A social–psychological analysis*. Monterey, CA: Brooks/Cole.

Wrightsman, L. S. (1991). Interpersonal trust and attitudes toward human nature. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 373–412). San Diego, CA: Academic Press.

Yovetich, N. A., & Rusbult, C. E. (1994). Accommodative behavior in close relationships: Exploring transformation of motivation. *Journal of Experimental Social Psychology*, *30*, 138–164.

Zahavi, A. (1995). Altruism as a handicap—The limitations of kin selection and reciprocity. *Journal of Avian Biology*, *26*, 1–3.

Zak, P. (2003, November). *Oxytocin may help humans bond*. Paper presented at the annual meeting of the Society for Neuroscience, New Orleans, LA.

The Psychology of Negotiation
Principles and Basic Processes

CARSTEN K. W. DE DREU
BIANCA BEERSMA
WOLFGANG STEINEL
GERBEN A. VAN KLEEF

The myriad agreements underlying close relationships, group decision making, or other forms of collective action often result from some type of negotiation—the communication between parties with perceived divergent interests to reach agreement on the distribution of scarce resources, work procedures, the interpretation of facts, or some commonly held opinion or belief (Pruitt, 1998). Negotiation occurs at all levels of society—between diplomats talking about arms reduction, between business leaders settling the conditions of a merger, between the owner of a house and a potential buyer, and even between travelers on an airplane deciding on who gets what part of the armrest. When individuals create high-quality agreements that meet both parties' needs and integrate both parties' aspirations, they create order and stability, foster social harmony, increase feelings of self-efficacy, reduce the probability of future conflict, and stimulate economic prosperity (Rubin, Pruitt, & Kim, 1994). When individuals create poor agreements, or fail to agree, they leave dissatisfied, create frustration and annoyance, and face continued conflict and disharmony.

This chapter is about negotiation and has three goals. First, we review recent developments in the social psychological study of negotiation. Such a review is needed because other reviews are 5–10 years old (e.g., Bazerman, Curhan, Moore, & Valley, 2000; Carnevale & Pruitt, 1992; Pruitt, 1998; Thompson, 1990) and do not cover recent work on, for example, motivated information processing and the social functions of emotion in negotiation. Second, we develop a set of basic principles that covers current insights into the negotiation process and

captures cognitive, motivational, and affective influences on the quality of agreements people reach. Third, we develop the idea that to make strategic decisions, individuals in negotiation need to make sense of their situation and their counterpart. That is, to understand negotiation we need to understand how people search and process information and use the emerging insights to make strategic decisions that, ultimately, affect their own as well as their counterpart's outcomes.

We begin this chapter with a brief discussion of the *structure* of negotiation and argue that individuals in negotiation face fuzzy situations that are full of uncertainties and ambiguities and require sense making on the part of the negotiators. We then discuss the *strategies and interaction patterns* that characterize negotiation and develop principles about strategic repertoires negotiators have, and about action–reaction patterns across different phases of negotiation. In the third section we discuss the *negotiator as motivated information processor*, concentrating on the (often detrimental) impact of cognitive heuristics, naive realism, and ego defensiveness. In this section we also discuss work showing that the influence of these information-processing barriers may be countered by the epistemic motivation to process information systematically and deliberately. In the fourth section we view the *negotiator as social animal*, and focus on impression management motives, and the wealth of research on proself versus prosocial motivation, questioning the rather popular assumption that individuals in conflict and negotiation are self-interested and ignorant of their counterpart's needs and desires. In the fifth section we consider

the *emotional negotiator* and discuss the intra- and interpersonal functions of affect and emotion in negotiation. In each of these five sections we identify one or more *basic principles of negotiation*. To examine the generality of these basic principles and processes, we review in the sixth section recent research on *cross-cultural differences* in negotiation. We conclude with a summary and integration of the 10 principles identified in this review and provide some general direction for future inquiry.

STRUCTURAL FEATURES OF NEGOTIATION

To understand negotiation we need to understand the structure of the interdependencies between negotiators. This structure provides the basis for strategizing, defines the boundaries on negotiator goals, and gives insight into the range of agreement options that exist. In this section we begin with a description of the interdependence structure that characterizes most, if not all, negotiation situations. We briefly discuss such notions as mixed-motive interdependence, power, and integrative agreements. We end this section with an analysis of the way negotiators themselves may perceive their situation, noting that most negotiations are not so clear-cut and simple but, instead, rather *fuzzy situations*. We conclude that an important task for individuals in negotiation is to make sense of their situation. The sense-making process, and its consequences for strategic choice and the quality of agreement, is further analyzed in the sections that follow.

Critical in what follows is that negotiation is about distributing and creating value. Value can be, first of all, positive or negative and negotiation can be about gains and rewards, about costs and losses, or about both (Larrick & Blount, 1997). Second, value can be realized immediately or delayed (e.g., Okhuysen, Galinsky, & Uptigrove, 2003) and, third, it can be certain or probable at best (e.g., Lax & Sebenius, 1986). These latter two aspects—delay and uncertainty—are important aspects of the fuzziness of negotiation situations, something we return to at the end of this section. Finally, it should be emphasized that value often is material and tangible, but this need not to be. Examples of nonmaterial value are relationship quality (e.g., in marriage), job satisfaction, and positive affective tone (in a work group). A good understanding of negotiation takes material as well as nonmaterial results into consideration (De Dreu & Beersma, 2005).

Interdependency, Power, and Control

In negotiation, parties depend on each other to acquire positive outcomes, to avoid negative outcomes, or both. This interdependence is usually mixed motive (Schelling, 1960): People start negotiations because they believe an agreement is potentially more beneficial than no agreement. They therefore have a *cooperative incentive* to work together with the other party. At the same time, they also have a *competitive incentive* to increase personal gain. Of course, the cooperative and competitive incentives are not always in balance. In some negotiations the cooperative incentive is relatively weak, for example, because negotiators have good alternatives to a negotiated agreement (Pinkley, 1995) or can walk away from the negotiation table and start new negotiations with someone else (Giebels, De Dreu, & Van de Vliert, 2000). In other negotiations the cooperative incentive is relatively strong, for example, because time in negotiation is costly or because one wishes to secure a valued relationship with the other.

The interdependence structure in negotiation provides the basis for influence and thus determines the participant's power (Bacharach & Lawler, 1981; Kelley & Thibaut, 1978). Power as the possibility to control and influence is different from any specific use of power, such as launching an attack, providing a reward, uttering a threat, or making a promise. Although power may derive from a variety of power bases such as someone's position within a group or one's expertise, it is the mutual dependence of individuals that allows power to occur. In negotiations between two persons A and B, B's dependence on A increases with the value of the benefits A can give B, and it decreases with B's access to alternative sources for those benefits (Emerson, 1972). Power and power asymmetries influence negotiation processes and outcomes primarily through their effects on the setting of aspirations and goals (Zetik & Stuhlmacher, 2002), and on the depth and deliberateness of information search and processing (Fiske, 1993). We return to these processes in the section on motivated information processing.

Negotiation research has been concerned primarily with the extent to which one can provide a negotiation partner with rewards and punishments (for reviews, see Bacharach & Lawler, 1981; Rubin & Brown, 1975). Some researchers have also considered the second aspect of dependence—the availability of alternatives (e.g., Giebels et al., 2000; Pinkley, 1995). In recent work, power has even been operationalized as the extent to which someone is central or peripheral in the group where negotiation takes place (Kim, 1997), or how easily coalitions with others can be formed (Beersma & De Dreu, 2002). Network positions and attractiveness as a coalition partner are relevant bases for power and influence in multiparty negotiations and less important (or even nonexistent) in dyadic negotiations (e.g., Polzer, Mannix, & Neale, 1998; Van Beest, Wilke, & Van Dijk, 2003).

When negotiation parties are groups rather than individuals, individuals experience mixed-motive interdependence within their ingroup, as well as with the outgroup. Bornstein (2003) provides examples of situations in which hostility toward the outgroup is perceived as cooperative, loyal behavior by members of one's ingroup. Vice versa, cooperative and conciliatory behavior toward an outgroup may be perceived, by one's fellow ingroup members, as disloyal and cowardly behavior that jeopardizes the ingroup's fate. As a result of this, groups tend to be more competitive and less cooperative with each other than individuals are in a interpersonal negotiation (Mikolic, Parker, & Pruitt, 1997; Robert & Carnevale, 1997; Wildschut, Pinter, Vevea, Insko, & Schopler, 2003).

Constituencies and Accountability

In intergroup conflict, it is typically neither feasible nor desirable for all those concerned to be present at the bargaining table, and usually representatives conduct the negotiation (Walton & McKersie, 1965). Representatives are often accountable to their constituents, either for the outcomes they achieve or for the process of making decisions (Tetlock, 1992). The former type is called *outcome accountability* and generally biases representatives' thinking and decision making toward the outcome (assumed to be) desired by the constituency. For instance, accountable representatives are motivated to impress constituents (Wall, 1991), and this often translates into a competitive stance toward their counterpart (e.g., Carnevale, Pruitt, & Seilheimer, 1981; O'Connor, 1997; but see Gelfand & Realo, 1999; Peterson & Thompson, 1997). The latter type is called *process accountability* and means that representatives have to account for the strategic decisions they made during negotiation. It generally increases the epistemic motivation to think deeply and systematically about one's reasons for (not) making certain decisions (Tetlock, 1992). We return to this when we discuss motivated information processing.

Single and Multiple Dimensions of Value

Negotiations can be about one single dimension of value or they can be multi-issue. In the case of the shared armrest in an airplane, a negotiation is about one single issue. However, when negotiating the division of household chores, partners talk not only about making dinner but also about the laundry, paperwork, gardening, and so on. When buying a car, we discuss the price as well as delivery and financing issues. These are multi-issue negotiations.

In the case of a single dimension of value the negotiation is about how to divide and distribute value. Such single-issue situations have been commonly studied in research using *prisoner's dilemmas* (PD; Dawes, 1980) or *ultimatum bargaining games* (UBG; Guth, Schmittberger, & Schwarze, 1982). In the UBG, for example, one participant divides a certain amount of value (e.g., $10) between him- or herself and a recipient. The recipient then decides whether to accept the proposed division or to reject it. When rejected, neither participant gets anything.

Negotiation often involves multiple issues. When the number of issues in a negotiation increases, it becomes less and less likely that parties (1) disagree on each and every issue, (2) see all issues as equally important, and (3) have the same rank order of issue importance (Raiffa, 1982). For example, a potential car buyer may be very concerned about the price of the car but indifferent about the delivery time. To the salesperson, however, the delivery time may be very important and getting the advertised price on the car may be a lesser concern. If preferences across issues are not diametrically opposed, the negotiation has *integrative potential* (Pruitt, 1981) and agreements exist that provide higher joint outcome than 50–50 compromises, or victory-for-one solutions. This is illustrated in Figure 26.1, which depicts the joint utility
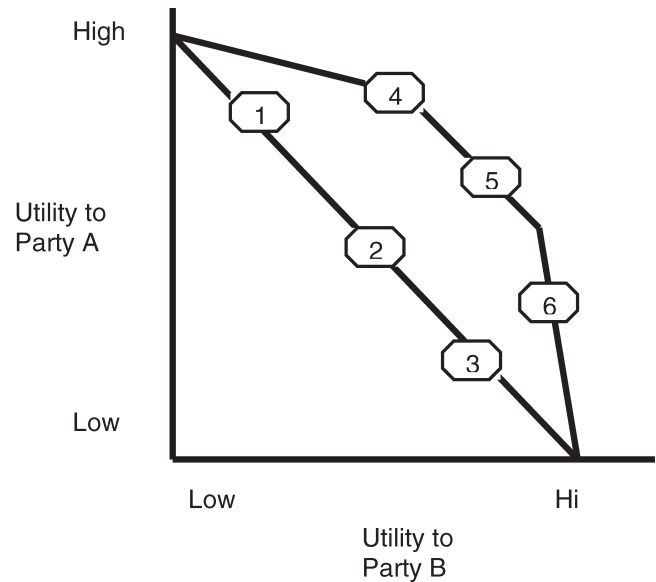
FIGURE 26.1. Integrative potential in a joint utility space.

space for a two-person negotiation. Agreements above the distributive border (numbers 4, 5, and 6) provide higher joint outcome than agreements at that that border (numbers 1, 2, and 3). Note also that agreements 1 and 4 provide Party A with a relative advantage, that agreements 3 and 6 provide Party B with a relative advantage, and that agreements 2 and 5 provide an equal distribution of outcomes between both negotiators. This illustrates that the level of joint outcome achieved by the negotiators does not necessarily mean there is an equal, or fair, distribution of outcomes between the two. Finally, it can be seen that Party A(B) does not get more outcomes in agreement 4(6) than in agreement 1(3), but that the former does provide the counterpart, and thus the collective, with higher outcomes than the latter. Agreements 4–6 are integrative agreements (note that agreement 5 also provides an equal distribution of outcomes between the two parties).

Many analysts have argued that integrative agreements are superior because they optimally exploit resources available to the negotiators, and because they provide high levels of relationship stability—both parties are happy with the agreement, feel self-efficacious, and are motivated to accurately implement the negotiated agreement (O'Connor & Arnold, 2001; Rubin et al., 1994). A key issue in much of the negotiation research and theory that we discuss in this chapter has been to understand when and why individuals (fail to) create and discover integrative potential and (fail to) reach integrative agreements. Accordingly, much of the negotiation research we review and discuss below has relied on (some variation of) a task developed by Siegel and Fouraker (1960), Kelley (1966), and Pruitt and Lewis (1975).

An example of this typical negotiation task is given in Table 26.1. The participants play the roles of management and union negotiators faced with the task of reach-

TABLE 26.1. Issue Charts for Management (Top) and Union (Bottom) Negotiators

Salary	Vacation days	Annual raise	Medical coverage
	Management issue chart		
E70.000 (00)	3 weeks (00)	15% (00)	100% (00)
E65.000 (15)	2.5 weeks (30)	12% (60)	80% (100)
E60.000 (30)	2 weeks (60)	9% (120)	60% (200)
E55.000 (45)	1.5 weeks (90)	6% (180)	40% (300)
E50.000 (60)	1 week (120)	3% (240)	20% (400)
	Union issue chart		
E70.000 (400)	3 weeks (120)	15% (240)	100% (60)
E65.000 (300)	2.5 weeks (90)	12% (180)	80% (45)
E60.000 (200)	2 weeks (60)	9% (120)	60% (30)
E55.000 (100)	1.5 weeks (30)	6% (60)	40% (15)
E50.000 (00)	1 week (00)	3% (00)	20% (00)

Note. In most studies individuals only see their own issue chart and are instructed not to show it to their counterpart.

ing agreement, here on four issues. The management negotiator is given the issue chart shown at the top of Table 26.1, and the union negotiator is given the issue chart shown on the bottom. These charts list five possible settlement points on each issue. Each side sees only its own chart, and typically both sides get the instruction that they can say anything to one another about it. As can be seen, the issue on the left has greater value to one side than to the other, and the opposite is true for the issue to the far right. This means that there are many possible agreements, and they vary in their value to the negotiators, both individually and collectively. The two issues in the center of the task have equal value for both parties and, in the parlance of negotiation research, are "distributive" issues; the other two issues that can be traded for one another are called "integrative." An integrative agreement in this task (a settlement of Euro 70 on "salary" and 20% on "medical") provides an outcome of greater value to each individual and to the pair together than a 50–50 agreement that split these two issues down the middle (joint gain of 800 vs. 460 on these two issues; see Table 26.1).

Fuzzy Situations and Sense Making in Negotiation

In many analyses of negotiation and strategic decision making, including those in behavioral economics and game theory, it is assumed that negotiators are fully informed players who perceive the structure as it is. We reject this assumption because it is unrealistic. Negotiators face and manage fuzzy, ambiguous, and messy situations. Parties to a negotiation do not have a full and accurate insight into the structure of the negotiation. They lack information about their partner's utility functions in that they do not know what is and what is not important to their partner, they do not know the amount of gain or loss their partner faces on specific issues, and they do not

know what goals their partner strives for. In terms of the integrative negotiation task in Table 26.1, each party only sees his or her own payoff chart and not the counterpart's (most studies using this task even explicitly forbid participants to show each other their charts).

Through communication, both verbal and nonverbal, negotiators may receive information about their counterpart's preferences and priorities. But even if they do, negotiators still have many good reasons to doubt the accuracy and trustworthiness of information about their partner's utility functions. One reason is that "organisms could employ functionally irrelevant external stimuli as seed values from which unpredictable responses could be generated by some internal and nonrandom process. . . . Animals must be able to produce behavior that is irreducibly uncertain from the point of view of their opponents" (Glimcher, 2003, pp. 293–294). In other words, the (human) brain has a fundamental capacity to confuse and mislead interaction partners. Another reason is the so-called information dilemma in which negotiators often find themselves. (Kelley & Thibaut, 1978; see also Murnighan, Babcock, Thompson, & Pillutla, 1999; Steinel & De Dreu, 2004). The exchange of honest and accurate information fosters the achievement of high joint gain and, at the same time, makes one vulnerable to exploitation by a (relatively more knowledgeable) partner. To prevent exploitation and foster personal gain, negotiators can use misrepresentation and deception. Individuals in negotiation thus constantly face the question of whether they should provide accurate information to achieve high collective outcomes, or rather strategically misrepresent their preferences to foster the achievement of good personal outcomes. Because of the information dilemma, negotiators have a fundamental reason to doubt any information from their partner. (This sometimes leads to quite ironic processes. For example, Moore [2004a, 2004b] showed that although revealing the deadline one faces speeds up concession making by one's counterpart, people believe such revelations are strategically unwise and thus forego personal and collective gain.)

The conclusion that negotiation situations are, in the eyes of the negotiators themselves, rather fuzzy situations directs us away from the use of simple, clear-cut games in which individuals have full information about both their own and their partner's outcomes and evaluations thereof, and in which individuals and their partners face a limited set of decision alternatives. We are not saying that using such situations to study negotiation is not useful and that insights gleaned from research into these situations are not interesting—we have, in fact, used this type of analysis quite often in our own research (e.g., De Dreu & McCusker, 1997; Steinel & De Dreu, 2004). What we are saying, however, is that to understand how individuals or groups interact and manage conflict-prone situations, we need to study how individuals *make sense* of their social environment: What types of information are negotiators looking for? How much information do individuals need before making a strategic decision? How thoroughly is new information that becomes available during negotiation processed and integrated into exist-

ing knowledge structures? And does all this influence the nature of the strategic interaction and the quality of the agreements people reach? It is these and similar questions about strategic decision making in fuzzy situations that drives the social psychological study of negotiation, setting it apart from analyses in game theory and behavioral economics that often are mathematically sophisticated and logically comprehensive but weak in truly accounting for human judgment and strategic choice.

NEGOTIATION STRATEGIES AND INTERACTION PATTERNS

The fuzziness of most negotiation situations, and the need to make sense of the situation, makes the communication processes and interaction patterns extremely important. Communication—verbal or nonverbal—contains information that allows negotiators to develop a more or less accurate understanding of their situation. As mentioned previously, however, the information dilemma in which negotiators find themselves may lower the extent to which communication is truthful and thus the degree to which negotiators develop an accurate understanding of their fuzzy situation. Furthermore, and because of the mixed-motive nature of negotiation, communication and interaction directly or indirectly serves the goal of safeguarding and promoting (1) personal interests and needs, (2) collective interests and needs including those of the counterpart(s), or (3) both. For example, making demands and refusing to concede serves personal interests, whereas giving up a position and making concessions fosters agreement and serves the counterpart's needs and interests. And as we see below, exchanging information about preferences and priorities often serves both personal and the partner's needs and interests.

In this section we review and categorize negotiation tactics and strategies, and we consider interaction patterns and processes. We develop our first three principles—the *strategic repertoire principle*, the *mindless matching–mindful mismatching principle*, and the *differentiation-before-integration principle*. We end with a brief discussion of the effectiveness of (combinations of) tactics and strategies in producing integrative agreements.

Strategic Repertoire Principle

Although an infinite number of strategies and tactical moves in negotiation may be conceived of, research and theory tend to converge on a five-way taxonomy that distinguishes between contending, compromising, conceding, problem solving, and inaction (Pruitt, 1998; Van de Vliert, 1997). Contending, compromising, and conceding are distributive strategies because they all focus on a particular distribution of outcomes that are predominantly in one's own advantage (contending), or in the counterpart's advantage (conceding). *Contending*, sometimes referred to as claiming value, focuses on imposing one's will on the other side. It involves the use of threats

and bluffs, of persuasive arguments, and of positional commitments. The flip side of contending is *conceding*, sometimes referred to as yielding. It is oriented toward accepting and incorporating the counterpart's will and involves unilateral concessions, unconditional promises, and offering help. *Compromising*, finally, involves the matching of other's concessions, making conditional promises and threats, and actively searching for a middle ground.

Problem solving, sometimes referred to as integration, involves a full and open exchange of information about priorities and preferences, the showing of and verification of insights, and logrolling—making trade-offs between important and unimportant issues (Pruitt & Rubin, 1986). Sometimes problem solving is implicit and indirect. Consider, for example, a "heuristic trial-and-error" strategy (Pruitt & Lewis, 1975; Siegel & Fouraker, 1960). Heuristic trial-and-error can be used in multi-issue negotiations and means that negotiators make, over time, a series of offers each having the same overall value to oneself but, perhaps, different value to one's partner. In Table 26.1, for example, some offers (e.g., E60,000, 2 weeks, 12%, and 40%) provide the same overall value of 450 to management as some other offers (e.g., E70,000, 1.5 weeks, 12%, and 40%) but more value to union (i.e., 455 and 625, respectively). By systematically varying offers in this way negotiators explore integrative potential without directly communicating about their preferences and priorities. Heuristic trial-and-error thus is a strategy that can be used when there is low trust and high fear of being exploited by one's counterpart.

In addition to these distributive (claiming) and integrative (creating) strategies, negotiators often engage in withdrawal behaviors, remain inactive, and avoid their counterpart. *Avoiding* is sometimes part of an overarching contending strategy and used strategically to get further concessions. For example, by (threatening to) withdraw from negotiation the counterpart may realize that making a concession is better than having no agreement whatsoever.

Often avoiding is genuine and the result of a decision dilemma. Contending is dangerous (it may escalate the conflict), conceding is contrary one's own interests, and problem solving is difficult, is time-consuming, and involves the risk of being exploited. To escape this difficult choice, individuals may opt for avoiding. Avoiding does not necessarily need to be inactive and passive. For instance, an interesting instance of avoiding that future research could further explore what we call *complexification*—making issues at the negotiation table more difficult, complex, fuzzy, and uncertain—to confuse the counterpart and to convince him or her that it is not (yet) the right time to make decisions.

There is quite some evidence for the above five-way taxonomy. Many studies have coded behavior from videotaped interactions into a number of categories that represent one of the aforementioned negotiation strategies, most often using an adapted version of a coding scheme developed by Pruitt and Lewis (1975; see, e.g., Carnevale & Lawler, 1986; De Dreu, Giebels, & Van de

Vliet, 1998; Weingart, Hyder, & Prietula, 1996). For example, contending is represented by "making positional commitments" (a strong statement that the speaker is unwilling to change his or her proposal) and by "persuasive arguments" (statements that bolster the importance of the speaker's case, or downplay the importance of the counterpart's case). Problem solving is often represented by "information exchange" (searching or providing information about preferences and priorities) and "cooperative statements" (expressing willingness to collaborate with the counterpart). Other work has used questionnaires in both field and laboratory settings and examined self-reports and peer reports of more global conflict management strategies. Psychometric analyses of these questionnaire data again reveal good support for the above five-way taxonomy (e.g., De Dreu, Evers, Beersma, Kluwer, & Nauta, 2001; Rahim & Magner, 1995).

Taken together, several decades of work on the types of negotiation tactics and strategies, in both the laboratory and the field, point to a *strategic repertoire principle*: Negotiator behavior takes one of five distinct forms—contending, conceding, inaction, compromising, and problem solving. Underlying these five behavioral strategies are a distributive (claiming value) and an integrative (creating value) dimension.

The Mindless-Matching–Mindful-Mismatching Principle

The strategic repertoire principle is mute about behavioral processes and interaction sequences. Some work, primarily using single-issue negotiation tasks, tracked over time the exchange of demands and concessions. In general, individuals have a strong and seemingly automatic tendency to mimic and reciprocate their counterpart's behavior (Axelrod, 1984; Gouldner, 1960). Pruitt and Carnevale (1993) refer to this as the matching of concession size and/or frequency.

Matching is a powerful tendency because it is simple (just do what the other did), because it prevents exploitation and at the same time rewards cooperation. Recent work provides further support. It has analyzed reciprocity and sequences of distributive and integrative behavior in multi-issue negotiations. Using Markov chain analyses, which describes the probability of an event occurring as dependent on the occurrence of a preceding event, Weingart, Prietula, Hyder, and Genovese (1999) found that negotiators tended to respond in kind to both distributive and integrative behaviors of their opponents. Interestingly however, the model could predict behaviors only when the previous two behaviors were known—rather than being a mere reaction to the behavior of the counterpart, a negotiator's behavior seems to depend on both parties' prior behaviors. Future cooperation, therefore, cannot be predicted from one prior cooperative move of the other party but only from prior reciprocated cooperation (see also Brett, Shapiro, & Lytle, 1998; Olekalns, Smith, & Walsh, 1996).

Matching and reciprocity are not limited to negotiation. They are very basic properties of social interaction.

In reviewing several relevant literatures, Ostrom (1998) noted: "while individuals vary in their propensity to use reciprocity, . . . a substantial proportion of the population . . . has sufficient trust that others are reciprocators to cooperate with them even in one-shot, no-communication experiments. Furthermore, a substantial proportion of the population is also willing to punish non-cooperators (or individuals who do not make fair offers) at a cost to themselves. . . . " (p. 12). Likewise, Komorita, Parks, and Hulbert (1992) note "the reciprocity norm is relevant and important in a social dilemma situation . . . [it] prescribes that we should help those who have helped us in the past, and retaliate against those who have injured us . . . the reciprocity norm is the underlying basis of stable relationships, and theorists have postulated that reciprocity is a basic norm of social interaction . . . " (p. 608).

In negotiation people may have incentives to not reciprocate their counterpart's behavior. For example, reciprocating other's competitive and tough strategy may lead to an impasse, providing neither party with positive outcomes. Vice versa, reciprocating other's cooperative behavior may be unwise when one's goal is to maximize personal gain, or when other's cooperative behavior was a strategic ploy and reciprocating other's cooperativeness leads to being exploited. That reciprocity is indeed not always the best option is reflected in the fact that negotiators often engage in *mismatching*—they place high demands when their counterpart appears soft and low demands when their counterpart appears tough (e.g., Liebert, Smith, Hill, & Keiffer, 1968). Mismatching occurs especially when there is time pressure and deadlines are looming (Smith, Pruitt, & Carnevale, 1982). Interestingly, mismatching also occurs when negotiators lack information about the partner's outcomes (Liebert et al., 1968) and limits (Pruitt & Syna, 1985). This suggests that from the counterpart's concessions negotiators derive insights into other's limit (Pruitt, 1981).

Pruitt and Carnevale (1993) propose that mismatching is more likely in the early and in the late phases of the negotiation, whereas matching is more likely in the middle. Early mismatching takes place because parties try to develop an understanding of their counterpart's bargaining strength and limit. Late mismatching occurs because a deadline is looming and agreement is better than impasse. In both cases, there is quite some strategic thinking underlying the tendency to mismatch, something that not necessarily precedes matching and reciprocity. This then allows us to advance the *mindless matching–mindful mismatching principle*: Whereas negotiators who do not engage in deep, deliberate, and strategic thinking match their counterpart's behavioral tendency, those who engage in such deep and strategic thinking will choose mismatching in early and late phases of the negotiation and matching in the middle phase.

The Differentiation-before-Integration Principle

Some work on interaction processes and behavioral sequences has looked at the different phases through

which negotiations cycle. For example, following an in-depth analysis of labor-management negotiations, Walton and McKersie (1965) suggested that negotiations often begin in quite a harsh and competitive way. After a series of competitive exchanges parties come to realize that this leads nowhere but to a costly impasse, that a change in behavior is needed, and that mutual problem solving is a viable alternative for safeguarding and promoting self-interest. This so-called differentiation-before-integration pattern is not limited to the ritual dance of labor-management negotiations. A laboratory study by Brett and colleagues (1998) showed that procedural remarks—statements that refer to the process of the negotiation itself—changed the focus from contentious, distributive communication to more constructive, integrative communication. Harinck and De Dreu (2004) coded temporary impasses—points in the negotiation where parties deadlocked on some issues and remained stuck for a while. Higher levels of contending early in the negotiation were related to temporary impasses, and temporary impasses were, in turn, related to problem solving late in the negotiation. Stepping back from and reflecting on the negotiation during a temporary impasse appears to facilitate a switch from competitive contending to more cooperative problem solving (see also Olekalns, Brett, & Weingart, 2003).

The work on interaction patterns and phases in negotiation points to the *differentiation-before-integration principle*: Negotiators engage in distributive contending early in the negotiation and in later phases switch to integrative problem solving.

Effectiveness of Strategies and Interaction Sequences

The research reviewed above often included some measure of dyadic or group performance, such as the extent to which agreements were integrative. As one might expect, problem-solving behavior, and especially the exchange of information about preferences and priorities, is an important positive predictor of integrative agreements (Pruitt, 1998). Studies that track contending and problem solving over time reveal that problem solving early in the process seems to have little impact, while problem solving late in the negotiation strongly fosters integrative agreements (Harinck & De Dreu, 2004). From the work on differentiation-before-integration it even follows that contending early in the negotiation facilitates, through temporary deadlocks, late problem solving and thus the discovery of integrative agreements.

Related to this is work on the effectiveness of the *good-cop/bad-cop strategy* showing that problem solving is more effective in creating integrative agreements when it is preceded by contending behavior (Hilty & Carnevale, 1992). Perhaps negotiators fear their counterpart returns to contending when the attempt at problem solving fails and thus cooperate without restraint (Brodt & Tuchinsky, 2000). Alternatively, the attempt at problem solving may loom particularly cooperative and constructive when it is embedded in, and contrasted with, a tough and contending strategy that preceded it.

THE NEGOTIATOR AS MOTIVATED INFORMATION PROCESSOR

The social psychological analysis of negotiation has traditionally considered the reasons for people's failure to uncover and exploit integrative potential. Although some answers derive from an analysis of the (sequences) of tactics and strategies that people employ, this type of work does not answer the question why it is that people do not engage in problem solving, and why it is that people more often than never try to exploit their counterpart. To understand this, we need to realize that even the simplest negotiations are both cognitively and emotionally taxing. Earlier we noted that most negotiation situations are fuzzy, and that sense making is a crucial aspect of negotiation. In this section, we review two decades of research that dealt with information processing in negotiation and revealed a number of information-processing barriers to constructive negotiation that prohibit people from reaching mutually beneficial, integrative agreements. These information-processing barriers can be placed under the rubric of cognitive heuristics, naive realism, and ego defensiveness (De Dreu & Carnevale, 2003; Ross & Ward, 1995). We discuss exemplars of each of these consecutively. Thereafter, we critically assess the assumptions underlying this work and propose that much of the influence of cognitive heuristics, naive realism, and ego defensiveness is reduced or even alleviated when individuals have high levels of epistemic motivation to process information deep and deliberately.

The Cognitive Heuristics Principle

Kahneman and Tversky (1973) identified three types of heuristics that people use to make sense of their complex environment and that sometimes lead to suboptimal, even downright incorrect, conclusions: (1) *anchoring*—the tendency to overly rely on an arbitrarily chosen reference point; (2) *availability*—the tendency to overly rely on information that is salient in memory; and (3) *representativeness*—the tendency to make judgments based on the most obvious features of the stimulus. Instances of these heuristics have been shown to be a barrier to constructive negotiation and dispute resolution.

Negotiation is all about relative values, and comparison standards have an important influence when it comes to setting one's aspiration or determining one's limit (Pruitt, 1981; Raiffa, 1982). Goals promote performance when they are difficult but feasible (Locke & Latham, 1990), and high aspirations should thus be useful when based on thorough and carefully calibrated information processing. Higher aspirations indeed produce higher demands and smaller concessions but also closer scrutiny of the issues and negotiation problems, and thus sometimes they promote both personal and joint gain (Huber & Neale, 1986; Zetik & Stuhlmacher, 2002). Unfortunately, however, humans have a tendency to base estimates on irrelevant anchor information and to adjust their estimates inappropriately (Kahneman & Tversky, 1973). Indeed, real estate agents insufficiently adjust for irrelevant anchor information when providing estimates

of real estate value (Northcraft & Neale, 1987). Negotiators also anchor on their counterpart's opening offer and inadequately adjust for this when making a counterproposal (Galinsky & Mussweiler, 2001; Kristensen & Gärling, 1997; Moran & Ritov, 2002). Put differently, anchoring leads to inadequately high or low aspirations, and thus negotiators unnecessarily deadlock or forego personal and joint gain. Interestingly, these anchoring effects hold for individuals as well as groups of negotiators and generalize across novices (e.g., students) and professionals (Whyte & Sebenius, 1997).

A related phenomenon is that people simplify the negotiation by coding prospective outcomes above a salient anchor or reference point as gains and outcomes below that anchor as losses (Kahneman & Tversky, 1979). When everything below one's aspiration is coded as a loss, the majority of the outcomes represent losses and concessions imply one is increasing the amount of loss one is willing to incur. When, in contrast, everything above the no-settlement outcome is coded as a gain, the majority of the outcomes represent gains and concessions imply that one is reducing the amount of gain one is striving for. Put differently, one and the same concession can be seen as an increase in one's losses or a decrease in one's gains. Because the negative utility of losses exceeds the positive utility of gains (Kahneman & Tversky, 1979), loss framing increases resistance to concession making (Pruitt, 1998), produces smaller concessions (De Dreu, Carnevale, Emans, & Van de Vliert, 1994), and increases the likelihood of an impasse (Bottom, 1998; Kristensen & Gärling, 1997). In multi-issue negotiation, loss framing tends to decrease integrative agreements (Bazerman, Magliozzi, & Neale, 1985), although this appears to be true especially when individuals are focused on personal rather than collective welfare (De Dreu & McCusker, 1997; Trötschel & Gollwitzer, in press).

The representativeness heuristic manifests itself in the use of stereotypes to interpret and predict the counterpart's behavior (Macrae, Bodenhausen, & Milne, 1994). Negotiators rely on stereotypic information when predicting their partner's tendency to compete, and concomitant expectations predict one's own behavior. De Dreu, Yzerbyt, and Leyens (1995) showed, for instance, that the decision to cooperate or not was influenced by information about one's partner's group membership. When the partner was a business major—a group seen as opportunistic and competitive by Dutch psychology students—participants made less cooperative choices than when the partner was a religion major—a group seen as moral and cooperative. Morris and colleagues likewise showed that individuals tend to interpret their partner's bargaining behavior in terms of his or her personality (e.g., low vs. high agreeableness) whereas the behavior was in fact due to situational constraints (i.e., high vs. low aspirations given by the experimenter). Importantly, this person attribution then formed the basis for subsequent action, leading negotiators into a self-fulfilling prophecy (Morris, Larrick, & Su, 1999).

Related to the above is recent work on *reputation*. Negotiation dyads involving individuals notorious for their competitiveness achieve lower joint outcomes than dyads without such an individual (Tinsley, O'Connor, & Sullivan, 2002). In iterated social dilemmas, competitive individuals get involved in fewer transactions and tend to end up with lower personal outcomes than do cooperative individuals (e.g., Axelrod, 1984). Like group membership, reputation serves as a heuristic cue that people use to make decisions about whether to start negotiations with this person and, if so decided, how to approach the counterpart.

A final example of representativeness is *fairness* as a heuristic device to settle disputes. An agreement that is deemed fair by both parties often serves as a focal point on which people coordinate their activities, and therefore it often is prominently reflected in the final settlement (Messick, 1993; Ohtsubo & Kameda, 1998). The prominence of the fairness heuristic can be seen in ultimatum bargaining where, from a utility-maximizing point of view, it is rational to propose the other an outcome slightly above zero (e.g., $9 – $1), and to accept any proposal that provides one with an outcome above zero. It is most common, however, that people propose a 50–50 split of the available value and reject proposals that deviate too much from such a fair division (e.g., Van Dijk, De Cremer, & Handgraaf, 2004). Pruitt (1981; see also Schelling, 1960) refers to these types of settlement as *prominent solutions* and discusses a variety of fairness heuristics including "equality" (both parties get the same amount), "equity" (each party receives a share proportional to their input), and "need" (the party that needs it the most gets the greatest share). Less well known, but highly relevant in some negotiation settings, are fairness heuristics like "opportunities," which state that the party who can make most use of a reward should therefore get the largest share of the reward, and the "historical precedent," which defines the parties' entitlements. Prominent solutions of this kind often develop into norms that have some aura of moral superiority and correctness and therefore strongly guide behavior (Thibaut & Kelley, 1959; see also Tyler, Huo, & Lind, 1999).

Using fairness heuristics can be efficient in single-issue negotiations. In multi-issue negotiations, however, they often lead parties away from integrative agreements because fairness principles focus people on (equally) distributing rather than creating value. Also, some fairness heuristics serve the personal interests of one party more than those of the counterpart. For example, in a divorce negotiation, the husband may prefer an equality rule for dividing the savings account when he contributed less salary to the household but an equity rule when he contributed more salary. What fairness rule to use may thus constitute a source of conflict in itself (Thompson & Loewenstein, 1992).

The above are only a few examples of the many cognitive heuristics negotiators may use to make sense of their situation and to reach judgments and decisions. This work on cognitive heuristics can be summarized in our *Cognitive Heuristics Principle*: Sense making in fuzzy negotiation situations leads individuals to rely on cognitive heuristics that provide simplified views of the negotiation. Reliance on these heuristics often speeds up effi-

cient decision making but also undermines the quality of negotiated agreement.

The Naive Realism Principle

Some bias in the individual's attention, encoding, and retrieval processes are due to naive realism—the individual's tendency to assume that he or she sees the world as it is, and that other rational perceivers will therefore share these perceptions. When others fail to see the world as it is, this reflects that others had different information, are lazy, or are biased by ideology or self-interest (Robinson, Keltner, Ward, & Ross, 1995; Ross & Ward, 1995).

Naive realism manifests itself in the *fixed-pie assumption* (Thompson & Hastie, 1990). When negotiators lack information about other's preferences and priorities they tend to assume that the other wants the same and values the same things in the same way as they do. The size of the pie thus is perceived to be fixed, and one's own and the other's preferences are perceived as diametrically opposed (Schelling, 1960). Making a fixed-pie assumption leads negotiators to engage in distributive bargaining and to forego possibilities for integrative agreement (Gelfand & Christakopoulou, 1999; Harinck, De Dreu, & Van Vianen, 2000; Pinkley, Griffith, & Northcraft, 1995; Thompson & Hrebec, 1996).

Naive realism is also reflected in *optimistic overconfidence*, which can lead disputants to forego attempts at settlement, "believing that time is on their side, and that complete, unilateral victory is just around the corner" (Ross & Ward, 1995, p. 266; see also Neale & Bazerman, 1985; Pronin, Gilovich, & Ross, 2004). Optimistic overconfidence may lead both parties to assume that an arbitrator will side with their case and therefore to prefer costly arbitration to constructive negotiation.

Finally, naive realism is reflected in the strong tendency for negotiators to engage in confirmatory information search. Disputants often seek confirmation rather than disconfirmation of their initial beliefs, plans, and strategies: "By framing questions in ways that could only confirm their hypotheses, both sides would be likely to discover only what they wanted to find" (Rubin et al., 1994, p. 105). Confirmatory search exacerbates the problems associated with building a strategy on inadequate and incorrect assumptions, beliefs, and cognitive structures (Neale & Bazerman, 1991). Diekmann, Tenbrunsel, and Galinsky (2003) found that negotiators who expect their counterpart to be very competitive set lower limits and began with less competitive claims than negotiators who expect their counterpart to be not so competitive. This tendency reflects the mismatching pattern we discussed in the section on strategies and interaction patterns. Obviously, a negotiator who sets low limits and makes modest claims invites his or her counterpart to set high limits and to respond with competitive claims. Indeed, Diekman et al. found strong evidence for such a self-fulfilling prophecy—the more competitive one expected the other to be, the lower one's demands, and the higher (i.e., more competitive) one's counterpart's demands became (for related work, see De Dreu & Van Kleef, 2004; Morris et al., 1999).

The aforementioned work can be summarized under the header of the *naive realism principle*: Negotiators make sense of their fuzzy situations by assuming others, including their counterpart, views and thinks like them. This tendency leads to inadequate assumptions and inaccurate conclusions about the true state of the negotiation situation; it drives toward distributive strategies including contending, compromising, and yielding and away from problem-solving behavior.

The Self-Threat Principle

Not all problems in negotiation can be traced to naive realism or the use of inadequate heuristics. Many problems stem from the fact that humans are motivated to develop, maintain, and protect a positive self-concept and that evaluations of the self are positively biased (e.g., Campbell & Sedikides, 1999). Ego defensiveness and the tendency to view oneself as better and more cooperative than average, including one's counterpart, hampers conflict resolution (e.g., Paese & Yonker, 2001). Consider, for example, reactive devaluation—the tendency to reduce the value of a concession or proposal received just because the concession was made or the proposal was offered. Mo'az, Ward, Katz, and Ross (2002) showed that Israeli Jews evaluate an actual Israeli-authored peace plan far less positively when it is attributed to the Palestinians than when it is attributed to their own government: What comes from the participant's ingroup is good and valuable, and what comes from the counterpart is bad, useless, and suspicious.

Self-serving bias extends to objects associated with the self in that people have greater liking for objects they possess compared to objects they do not own. Individuals indeed ask more money to give up something they own (e.g., a coffee mug) compared to what they are willing to pay to acquire that same object (Kahneman, Knetch, & Thaler, 1990). Although this may reflect loss aversion—giving up an object induces greater pain than receiving that same object induces pleasure—*mere ownership effects* occur without the possibility of loss (e.g., Nesselroade, Beggan, & Allison, 1999) and are stronger for people in high need of self-enhancement (Beggan, 1992).

Mere ownership has important implications for conflict management and dispute resolution. Kahneman and colleagues (1990) aptly noted that when opposing individuals overvalue what they own and undervalue what they do not own, trade becomes exceedingly difficult. Carnevale and Ledgerwood (2004) reported that endowment effects are moderated by social identity, with the effect being larger for group-oriented objects that reflect a person's group identity. De Dreu and Van Knippenberg (2005) showed that because people identify with arguments associated with their self-concept so easily, any attack of, or opposition to, that argument constitutes a self-threat. This in turn elicits ego-defensive responses, including competitive behavior, negative views of the interaction partner, and attitude polarization (see also Baumeister, Smart, & Boden, 1996).

The work on self-enhancement tendencies and ego defensiveness points to the *self-threat principle*: Because indi-

viduals desire a positive self-view, and opposition and conflict inherently form a threat to the self, negotiators develop ego defensive (i.e., hostile and competitive) reactions vis-à-vis their counterpart. Self-threat drives toward contending (and perhaps avoiding) behaviors and away from problem-solving and yielding. Self-threat thus is a basis for conflict escalation.

Epistemic Motivation and the Deep Thinking Principle

A fundamental assumption underlying much of the work reviewed thus far is that people have limited cognitive capacity and are bound to use cognitive heuristics, to engage in erroneous reasoning, and to perceive the world exclusively from their own perspective. Although social psychologists have embraced this "cognitive miser" perspective for quite some time (see, e.g., Fiske & Taylor, 1991), work on individual judgment and decision making and on social influence and persuasion uncovered that individuals can, within their limited cognitive capacity, choose from two alternative strategies for processing information. The first strategy is to solve logical problems, to evaluate persuasive arguments, or to form impressions of others through a quick, effortless, and heuristic processing of information that rests on well-learned prior associations. Alternatively, individuals may engage in more effortful, deliberate and systematic processing that involves rule-based inferences (Chaiken & Trope, 1999).

The extent to which individuals in negotiation search, encode, and retrieve information depends on their *epistemic motivation*—the desire to develop and hold accurate and well-informed conclusions about the world (De Dreu & Carnevale, 2003; Kruglanski & Webster, 1996; Webster & Kruglanski, 1994). Epistemic motivation can be dispositional. For example, individuals differ in their chronic need for cognitive closure. Individuals with high need for closure are characterized by cognitive impatience, a tendency to leap to judgments on the basis of inconclusive evidence, and rigidity of thought. Individuals with low need for closure, in contrast, may prefer to suspend judgment, engaging in extensive information search and generating multiple interpretations for known facts. However, situational cues may influence epistemic motivation as well. Epistemic motivation increases, for example, when there is *process accountability* and individuals expect to be observed and evaluated by others with unknown views about the process of judgment and decision making (Lerner & Tetlock, 1999). Epistemic motivation also depends on *power differences*. Fiske (1993) has argued, for example, that low-power individuals try to (re)gain control over their situation by paying close and careful attention to their powerful other, to accurately predict other's needs, desires, and possible actions. Finally, *time pressure and noise* reduce epistemic motivation and increase the motivation to reach (cognitive) closure (e.g., Webster, Richter, & Kruglanski, 1996).

Features of the negotiation process may also raise or undermine epistemic motivation. The negotiation process may be tiring, and fatigue has been shown to lower epistemic motivation (Kruglanski & Webster, 1996). Unexpected events, such as having one's first offer being accepted, provoke surprise and counterfactual thought (Galinsky, Seiden, Kim, & Medvec, 2002) and may raise epistemic motivation. Finally, there is growing evidence that critical events in the negotiation, such as temporary impasses (Harinck & De Dreu, 2004) and so-called turning points (Druckman, 1986), raise epistemic motivation and lead to thorough and open-minded information processing.

Epistemic motivation reduces the extent to which cognitive heuristics impact negotiation processes and outcomes. When negotiators have low need for cognitive closure, are placed under low time pressure, or are stimulated to think about counterfactuals, they are less influenced by irrelevant anchors, or by stereotypical information about their counterpart (e.g., Bar-Joseph, & Kruglanski, 2003; De Dreu, 2003; De Dreu, Koole, & Oldersma, 1999; Galinsky & Mussweiler, 2001; Golec & Federico, 2004). Epistemic motivation also reduces the impact of naive realism. When negotiators have less power than their opponent, they engage in more diagnostic and less confirmatory information search (De Dreu & Van Kleef, 2004), and when they are placed under low time pressure or under process accountability, they are more likely to revise their erroneous fixed-pie assumption (De Dreu, 2004; De Dreu, Koole, & Steinel, 2000). Also, when individuals adopt a third-party instead of partisan perspective, they consider the conflict from multiple instead of one perspective and come to more balanced judgments (Kemmelmeier & Winter, 2000). Finally, as we see in the section on affect and emotion, epistemic motivation influences the extent to which negotiators take their opponent's emotional state into consideration during the negotiation (Van Kleef, De Dreu, & Manstead, 2004b). In other words, raising epistemic motivation leads negotiators to engage in more systematic and thorough information processing, and this undermines the (often detrimental) impact of cognitive heuristics, of naive realism, and of self-enhancement and ego-defensive tendencies (De Dreu & Carnevale, 2003; De Dreu & Van Knippenberg, 2005).

The foregoing work points to the *deep thinking principle*: Negotiators may be more or less motivated to develop a rich and accurate understanding of their situation and therefore engage in more or less deep, systematic, and deliberate search for, and processing of, information. Higher levels of epistemic motivation are associated with less reliance on cognitive heuristics, faster and more rigorous correction of inadequate assumptions and perceptions, and less ego defensiveness following self-threat.

Summary

Constructive negotiation is often impeded by the use of cognitive heuristics, naive realism, and ego defensiveness. Part of the problem is related to the complexity and uncertainty that are inherent to negotiation and forces individuals to use suboptimal cognitive shortcuts. Part of the problem is, however, due to fundamental psychologi-

cal principles that operate inside as well as outside negotiation. Because of naive realism, negotiators make a fixed-pie assumption and forego possibilities of joint gain, or engage in confirmatory information search and get trapped into a self-fulfilling prophecy. Because of ego defensiveness, negotiators underestimate the value of their partner's proposal and respond overly hostile when their partner opposes their arguments or disputes their needs and desires. However, consistent with dual-process models (Chaiken & Trope, 1999) and lay epistemic theory (Kruglanski & Webster, 1996), growing evidence suggests that negotiators can switch from shallow and heuristic to deep and deliberate information processing. Systematic information processing, which is likely under high epistemic motivation, reduces the detrimental influence of using cognitive heuristics, naive realism, and ego defensiveness.

THE NEGOTIATOR AS SOCIAL ANIMAL

For quite some time, the study of the individual negotiator as a motivated information processor ignored the social context in which negotiation takes place (Kramer & Messick, 1995), and thus that much of what negotiators do is driven by their need to impress an audience or constituency, or by their need to build a particular reputation. In addition, the work reviewed above implicitly or explicitly proceeds on the assumption that negotiators are motivated by self-interest and are ignorant of, or indifferent about, their counterpart's concerns, needs, and interests.

In this section we review research and theory that view the negotiator as a social animal, one concerned with his or her social reputation, with the impressions he or she makes on others, including the counterpart, and with the needs and interests of his or her counterpart. We begin with work on impression management and reputation concerns and then discuss research and theory on social motives—the negotiator's preference for a particular distribution of outcomes between oneself and the counterpart (McClintock, 1977).

Impression Management and Reputation Concerns

Impression motivation is usually defined as the desire to make a good impression on the other and to get along (Snyder, 1992; Tetlock & Manstead, 1985). Individuals in negotiation may, however, also be highly motivated to convey the impression of a tough negotiator not to be trifled with (Chaiken, Gruenfeld, & Judd, 2001; Wall, 1991). For example, a manager may be motivated to convey the image of a tough negotiator, only to avoid his or her employees trying to negotiate about every single task assignment or allocation decision.

Impression management motives are particularly prominent in intergroup negotiations where individuals need to convey a certain impression not only toward their counterpart but also toward their fellow ingroup members. Brown (1977) showed, for instance, that negotiators are willing "to cut their nose to save their face"

(i.e., to incur substantial cost to impress an audience). Sometimes impressing an audience and a counterpart needs to be done at more or less the same time, for example, when constituents are surveying the negotiation (Carnevale et al., 1981). In some instances, this needs to be done sequentially, when a deal needs to be sold to one's constituents, when negotiation progress is discussed during caucuses, or when negotiation results are part of someone's annual performance appraisal. How the timing of impression management concerns influences its nature and effects is yet unknown, however, and could be addressed in future research.

Early work suggests that constituent surveillance increases competitiveness because negotiators believe being competitive would impress their constituents (Carnevale, Pruitt, & Britton, 1979). More recent studies suggest that surveillance reduces the negotiator's tendency to strategically misrepresent information (Steinel & De Dreu, 2004), and in collectivist cultures it leads to greater cooperation (Gelfand & Realo, 1999). Apparently, the expectations of the audience drive the kind of impressions negotiators wish to convey, and this steers them toward a cooperative or competitive stance (Enzle, Harvey, & Wright, 1992).

The issue of impression management motivation can also be approached from intergroup relations perspectives such as social identity theory (Turner, 1987). An example is a study by Van Kleef, Steinel, Van Knippenberg, Hogg, and Svensson (in press), who investigated how a representative's behavior in an intergroup negotiation is influenced by his or her standing within the group. They reasoned that peripheral group representatives, who occupy a marginal position within the group, should be motivated to gain the approval of their fellow group members, especially if group membership is seen as attractive. One way to do this would be to adopt a tough negotiation stance vis-à-vis the outgroup. Their study indeed showed that peripheral group representatives were tougher than prototypical representatives (who are at the core of the group), but only when the other group members could monitor their negotiation behavior and group membership was perceived as attractive.

Some studies on impression management in negotiation have considered cognitive processes. For example, Jordan and Roloff (1997) classified negotiators as being high or low in self-monitoring and showed that high self-monitors engaged more in planning of impression management (e.g., "be friendly so he'll think I'm giving him a good deal"). They also designed a greater variety of tactics and strategies than low self-monitors. The latter result may reflect greater flexibility of thought among high self-monitors. Ohbuchi and Fukushima (1997) indeed found self-monitoring to be associated with more integrative negotiation, albeit only when there was mild (rather than acute) time pressure.

Taken together, impression management and reputation concerns clearly are important in negotiation. Some promising findings have been reported, but clearly more and more systematic research is needed to uncover the forms impression management motivation can take, and to further substantiate the idea that impression manage-

ment motivation relates to flexibility of thought and strategic choice.

Social Motivation

In the section on motivated information processing we noted that individuals are predisposed to develop, maintain, and defend a positive view of themselves. This should, however, not be taken as an indication that individuals are only motivated to pursue their self-interests. In fact, a variety of social motives can be distinguished, including altruistic, competitive, individualistic, and cooperative (McClintock, 1977). Many studies on social dilemmas, conflict, and negotiation have relied on the more global distinction between proself and prosocial motivation (e.g., Beersma & De Dreu, 2002; Carnevale & Lawler, 1986; De Dreu & Van Lange, 1995; Weingart, Bennett, & Brett, 1993). Proself motivation comprises both competitive and purely individualistic goals, and prosocial motivation comprises both cooperative and purely altruistic goals. Individuals with a proself motivation desire to maximize their own outcomes, and they have no (or negative) regard for their partner's outcomes. They tend to see the negotiation as a competitive game in which power and personal success is important. Individuals with a prosocial motive desire a fair distribution that maximizes both own and other's outcomes, and they have a positive regard for their partner's outcomes. They tend to see the negotiation as a cooperative game in which fairness, morality, and harmonious social relations are key.

Although prosocial motivation sometimes reflects "enlightened self-interest," growing evidence indicates that it often is genuine and detached from immediate self-interest (Batson, 1998; Caporael, Dawes, Orbell, & Van de Kragt, 1989). Evidence also suggests that social motives have a strong dispositional component. Kelley and Thibaut (1978) have argued that individuals learn to adopt a specific social motive through direct experience, observation, and modeling (e.g., when one witnesses the behavior of important social models like parents, siblings, and teachers). Examples of dispositions that drive toward a prosocial motivation include dispositional trust (Parks, Henager, & Scamahorn, 1996), agreeableness (Barry & Friedman, 1998; Graziano, Jensen-Campbell, & Hair, 1996), need for affiliation (Langner & Winter, 2001), and prosocial value orientation (De Dreu & Van Lange, 1995).

Apart from individual dispositions, the tendency to adopt a prosocial versus proself motivation can also be triggered by features of the situation. For example, interaction partners can be instructed to see the other as a "partner" instead of as "opponent" (Burnham, McGabe, & Smith, 2000), or they can be rewarded for joint instead of personal success (e.g., Weingart et al., 1993). Finally, positive affect (Carnevale & Isen, 1986; Forgas, 1998) and the expectation of cooperative future interaction (Ben-Yoav & Pruitt, 1984) make a prosocial motivation more likely than a proself one.

Consistent with the notion that people encode and retrieve goal-consistent information easier and better

(Fiske & Taylor, 1991), proself negotiators selectively search and process information that corroborates their idea that the negotiation is a competitive game and that their counterpart cannot be trusted. Pro-socially motivated negotiators, in contrast, selectively search and process information that corroborates their idea that the negotiation is a collaborative game, that their counterpart can be trusted, and that cooperation is morally superior (e.g., Camac, 1992; De Dreu & Boles, 1998; Van Kleef & De Dreu, 2002). Finally, there is some evidence that prosocial individuals see the world in broader categories, and that proself individuals see the world in more discrete, narrow categories (Carnevale & Probst, 1998).

Most work on social motives has focused on strategic choice and negotiation outcomes. This work centers around two interrelated theories—the theory of cooperation and competition developed by Deutsch (1973), and dual-concern theory (Pruitt & Rubin, 1986). The *theory of cooperation and competition* argues, in brief, that proself negotiators develop distrust, hostile attitudes, and negative interpersonal perceptions. They use persuasive arguments, positional commitments, threats, bluffs, and coercive power to get their way. Prosocial negotiators, in contrast, develop trust, positive attitudes, and perceptions; engage in constructive exchange of information; listen; and seek to understand one another's perspective. Pro-socially motivated negotiators thus are more likely to reach integrative agreement.

Pruitt and Rubin's *dual-concern theory* postulates two kinds of concern, other concern and self-concern, each ranging in strength from weak to strong. Other concern is related to the concept of social motive, with proself negotiators having weak other concern and prosocial negotiators having strong other concern. Self-concern is related to *resistance to yielding*—the negotiator's intransigence about concession making. Negotiators have higher resistance when there is low rather than high time pressure, when they have been given a high rather than low limit or aspiration level, when outcomes are framed as losses rather than gains, or when parties have good rather than bad alternatives to an agreement (Druckman, 1994; Pruitt, 1998). The theory further postulates that social motives and resistance are independent—the former refers to outcome utility and desirable end states, and the latter refers to concession making as a means of achieving that end state. Thus, a prosocial negotiator with high resistance may truly value his partner's outcomes but may be unable to make any more concessions. Likewise, a proself negotiator with low resistance may not care at all about her partner's outcomes but may make yet another concession to avoid an impasse yielding no outcomes whatsoever.

A meta-analysis of 28 studies by De Dreu, Weingart and Kwon (2000) examined support for the theory of cooperation and competition and the dual-concern theory. Effects of social motive and resistance to yielding (high vs. low vs. unknown) on contending, problem solving, and joint outcomes were examined. Consistent with dual-concern theory, results showed that negotiators were less contentious, engaged in more problem solving, and achieved higher joint outcomes when they had a

prosocial rather than proself motive, but especially when resistance to yielding was high (or unknown). Although more complex than the theory of competition and cooperation, dual-concern theory thus better fits the data. In addition, the meta-analysis revealed no moderating role of the way social motivation was measured or manipulated. Dispositions and situational cues thus appear to have similar effects on negotiation processes and outcomes.

It is important to note that most of the research discussed above provided all participants with the same social motive—all members of the dyad or group had a prosocial, or a proself motivation (for an exception, see Rhoades & Carnevale, 1999). When in motivationally heterogeneous settings prosocial individuals discover that their counterpart has a proself motivation, they tend to switch from cooperative, problem-solving behavior to more competitive strategies. In fact, there is some evidence for an overassimilation effect—prosocial individuals behave even more competitively with a proself partner than proself people do (Kelley & Stahelski, 1970; Steinel & De Dreu, 2004; Van Lange, 1992). An explanation for this overassimilation effect is that prosocial individuals have a moralistic desire to teach the other a lesson (Van Lange, 1992). In general this work suggests a "pull" toward proselfness in dyads and groups where at least one member has a proself motivation.

In the section on motivated information processing we reviewed work showing that epistemic motivation increases systematic, effortful information search and processing. Combining this insight with work on social motives led to the formulation of a *motivated information-processing model* of negotiation (Carnevale & De Dreu, 2006; De Dreu, 2005; De Dreu & Carnevale, 2003; De Dreu, Weingart, et al., 2000; De Dreu et al., 1999; De Dreu, Koole, & Steinel, 2000; De Dreu & Van Kleef, 2004; Van Kleef et al., 2004b). The basic idea is that because negotiators often lack information about the task and about their counterpart, a motivated search for and provision of information provides new pieces of information on an almost continuous basis. Based on the research evidence discussed previously, the model further assumes that social motivation drives the kind of information (i.e., either cooperative or competitive) that negotiators seek, provide, and consider. Epistemic motivation, in contrast, primarily determines the extent to which this "cooperative" or "competitive" information is being processed, and thus the extent to which this information impacts strategic choice and the quality of agreement.

According to the motivated information-processing model, prosocial negotiators are more likely to develop trust, to engage in problem solving, and to reach high-quality agreements than are proself negotiators, but especially when epistemic motivation is high. Evidence for this central hypothesis was obtained in three experiments that showed that prosocial compared to proself negotiators had better recall of cooperative information, developed more trust, engaged in more problem solving, and reached more integrative

agreements when they had high but not when they had low levels of epistemic motivation (De Dreu, Beersma, Stroebe, & Euwema, 2006). The motivated information-processing model thus accounts for and integrates past research and makes new and valid predictions about the interplay between basic classes of motives in conflict and negotiation.

Summary: The Social Motive Principle

Although research has demonstrated the importance of impression management concerns, especially in the context of intergroup negotiations, more systematic work is needed to uncover its consequences for information processing and strategic choice. Much more developed is our insight into the determinants and consequences of having a proself versus prosocial motive, and this lends itself for proposing the *social motives principle*: (1) Negotiators have or adopt a prosocial or a proself motivation; that (2) drives for confirmatory sense-making processes so that prosocials are more likely to see the negotiation as a collaborative game, and their counterpart as a trustworthy person, whereas proselves are more likely to the negotiation as a competitive game, and their counterpart as untrustworthy; and (3) leads prosocials to engage in more problem solving when there is high resistance to concession making and more yielding when there is low resistance.

One important area for theory and research is to further our understanding of the interplay between impression management concerns and social motives. At first blush it may seem that prosocial negotiators desire and develop a softer and friendlier image than proself negotiators, who more easily develop a reputation of being cunning, shrewd, and tough. However, we believe that negotiators often use the impression they wish to convey in service of the prosocial or proself end goals they try to achieve. Prosocial negotiators may well believe that conveying a tough image serves their desire to get collectively beneficial outcomes better than conveying a soft image. Vice versa, proself negotiators may convey the impression of a weak, soft, and lenient negotiator to "trick" their counterparts in making more concessions. It is these possibilities that require further study, and such work will allow us to further integrate currently isolated lines of inquiry.

THE EMOTIONAL NEGOTIATOR

Perhaps because negotiation and conflict research has strong roots in game theory and models of rational decision making, social psychology has for a long time ignored emotions as part and parcel of the negotiation process. In the past decade this unfortunate situation has changed, with more and more articles incorporating emotions into the study of conflict and negotiation (Barry, Fulmer, & Van Kleef, 2004). In this section we review this work and demonstrate that much of the effects of emotions can be understood in terms of the motivated

information-processing principles outlined in the previous sections.

There are myriad definitions of *emotion*, most of which point to three distinct features of emotion: physiological reactions, action tendencies, and subjective experience (Lazarus, 1991). Emotions differ from *moods* in that they are discrete, of relatively high intensity and short duration, and intentional, that is, directed at an object, person or event (Frijda, 1993). We use the term "emotion" in the sense intended above, whereas *affect* is used as an umbrella concept encompassing both moods and emotions (cf. Barry et al., 2004).

Much of what happens in negotiation provokes emotions. Impasse in negotiations elicits frustration (O'Connor & Arnold, 2001), and perceptions of unfairness trigger anger and spite (Pillutla & Murnighan, 1996). Procedural justice in negotiations enhances positive feelings about the negotiation and attenuates negative feelings over outcomes, while distributive justice produces satisfaction and lowers disappointment and resentment (Hegtvedt & Killian, 1999). Other work shows that having your first offer accepted elicits not only counterfactual thoughts but also worry and disappointment (Galinsky, Mussweiler, & Medvec, 2002; Galinsky, Seiden, et al., 2002; see also Kwon & Weingart, 2004; Naquin, 2003). Likewise, receiving outcomes superior to the counterpart's may induce happiness, whereas receiving outcomes inferior to the counterpart's produces a variety of negative feelings (Thompson, Kramer, & Valley, 1995), although this seems to be particularly true among individuals with a proself rather than prosocial motivation (Gillespie, Brett, & Weingart, 2000).

The Moody Negotiator Principle and the Emotion-as-Strategic-Information Principle

Emotions not only follow from but also drive negotiation processes and outcomes. An early study was conducted by Carnevale and Isen (1986). They manipulated mood by means of humorous cartoons and a small gift and found that participants with a positive mood used fewer contentious tactics and obtained higher joint outcomes than did negotiators in a neutral mood (see also Baron, 1990; Forgas, 1998). In contrast, negative moods tend to produce less constructive behavior. Sadness stimulates evasiveness especially when conflict is intense (Forgas & Cromer, 2004), and anger increases the use of competitive strategies (Forgas, 1998) and reduces joint outcomes (Allred, Mallozzi, Matsui, & Raia, 1997). Together these works point to the *moody negotiator principle*: Angry negotiators are more likely to play tough and to make small concessions, than are sad negotiators who are more evasive and happy negotiators who easily make concessions.

Recognizing that negotiation is a social phenomenon, several scholars have emphasized the importance of the interpersonal or social effects of emotions in negotiations, arguing that emotions have important social functions and consequences (Barry et al., 2004; Morris & Keltner, 2000). Most notably, emotions convey information about how one feels about things, about one's social intentions, and about one's orientation toward other people. In this way, emotions can serve as incentives or deterrents for other people's behavior: Negative emotions serve as a call for mental or behavioral adjustment, whereas positive emotions serve as a cue to stay the course (Keltner & Haidt, 1999). Thus, emotions can influence our own behavior and the behavior of others.

To study the interpersonal effects of emotions in negotiation, Van Kleef and colleagues (Van Kleef, De Dreu, & Manstead, 2004a; Van Kleef et al., 2004b) used a computer-mediated negotiation task introducing a simulated opponent who appeared angry, nonemotional, or happy. Participants with an angry opponent made larger concessions than did participants with a nonemotional opponent, whereas those with a happy opponent made smaller concessions. Results further showed that these effects result from *tracking* (see Pruitt, 1981): Negotiators with an angry opponent estimated others' limit to be high, and to avoid costly impasse they made large concessions. Conversely, negotiators with a happy opponent judged the opponent's limit to be low, assumed agreement was not endangered, and accordingly made small concessions. These findings suggest that the epistemic motivation to consider and process other's emotion information should moderate the foregoing effects of anger and happiness. Van Kleef and colleagues (2004b) indeed showed that negotiators with high epistemic motivation (i.e., individuals with a low need for cognitive closure, placed under low time pressure, or being relatively independent of the counterpart) were strongly influenced by their counterpart's emotions, whereas those with low epistemic motivation (high need for closure, high time pressure, high dependence) were unaffected.

Whereas the work by Van Kleef and colleagues (2004a, 2004b) shows that angry negotiators elicit more concessions from their opponent than happy negotiators, recent work by Friedman, Anderson, Brett, Olekalns, and Lisco (2004) suggests that alternative patterns may be possible as well. In their study, Friedman et al. used data from online mediations in disputes arising over eBay transactions. They coded, from email messages parties send to each other and their mediator, anger words and correlated this with dispute settlement rate. Their results showed that expressing anger lowers settlement rate, in part because expressing anger elicits an angry response from one's counterpart. When respondents were particularly vulnerable, however, expressing anger did not lower settlement rates, suggesting that powerless negotiators are more likely to yield to an angry opponent than are powerful negotiators (see also Van Kleef, De Dreu, Pietroni, & Manstead, 2006).

As is clear from the foregoing, research on the interpersonal effects of emotions in negotiations so far has paid a disproportionate amount of attention to the effects of anger and happiness. At present, we are aware of only one study that addressed demands and concession making in negotiation as a function of other emotions expressed by the counterpart (Van Kleef, De Dreu, & Manstead, 2006). Extending previous work on supplica-

tion and appeasement (e.g., Clark, Pataki, & Carver, 1996), this study revealed that participants whose opponents expressed emotions of appeasement (guilt or regret) developed a positive impression of their opponents but were nonconciliatory in their demands. By contrast, participants whose opponents showed supplication emotions (disappointment or worry) rated their opponents less positively but made larger concessions and placed lower demands.

The study of the social functions of emotion in negotiation is still in its infancy. Nevertheless, there is evidence to the support the *emotions-as-strategic-information principle*, which holds that one's counterpart's emotions (e.g., anger, happiness, disappointment, and guilt) provide information that has strategic implications, and which is used as such provided that there is sufficient motivation.

A logical next step in the exploration of the role of emotions in negotiation would be to investigate the interplay between the intrapersonal and interpersonal effects of emotions. Do the positive effects of happiness at the intrapersonal level of analysis (e.g., more creative problem solving) outweigh the negative effects at the interpersonal level (e.g., risk of exploitation)? Does the strategic advantage of expressing anger outweigh the negative impact of anger on the negotiation climate? Aside from the interplay between the intra- and interpersonal effects of emotions, future research could investigate the effects of other emotions than anger and happiness. For instance, what are the effects of fear, sadness, disappointment, and worry? And how do people respond to others' expressions of guilt or regret? We believe that such questions constitute a fruitful starting point for future research on the role of emotion in negotiation.

NEGOTIATING WITHIN AND ACROSS CULTURES

The research we have reviewed so far has predominantly been conducted in Europe and the United States. More and more interactions take place, however, between people from different parts of the world, with different cultural backgrounds. Not surprisingly, therefore, scholars from around the world have questioned the validity of "Western" knowledge and insights and commenced to study the possible cultural influences on negotiation processes and outcomes (Gelfand & Brett, 2004; Tyler, Lind, & Huo, 2000). In this section we review this work, partly because it represents an exciting area for further inquiry, and partly because it allows us to see whether the principles and processes reviewed in the previous sections generalize across cultures.

Much of the work on culture and negotiation relies on Hofstede's (1991) distinction between individualism and collectivism. According to Hofstede, "individualism pertains to societies in which the ties between individuals are loose. . . . Collectivism as its opposite pertains to societies in which people from birth onwards are integrated into strong, cohesive ingroups, which throughout people's lifetime continue to protect in exchange for unquestion-

ing loyalty" (p. 5). The United States and the Western European countries are high on individualism. Collectivistic countries can be found in Southern Europe, Asia, Africa, and South America.

Culture and Information Processing

Culture influences information processing in a number of ways. It shapes and determines the norms people have for managing disputes and resolving conflicts (Tinsley & Brett, 2001), including whether or not lying and deception are morally appropriate (Triandis et al., 2001). Culture also drives the cognitive representations of conflict people have. Individuals with a collectivist background, for example, view conflict more as compromise focused whereas individuals with an individualist background view these very same conflicts more as win–lose focused (Gelfand et al., 2001). This finding maps onto our earlier suggestion that individuals scoring high on collectivism may have stronger prosocial motivation than individuals scoring high on individualism (see also Chen, Chen, & Meindl, 1998; Chen, Mannix, & Okumura, 2003; Hulbert, Correa da Silva, & Adegboyega, 2001; Probst, Carnevale, & Triandis, 1999). It also suggests that some heuristics—like the equality rule—may be more prominent in some culture than in others (e.g., Leung, 1987).

Culture also affects the *level* at which information processing bias emerges. For example, Carnevale (1995) found *mere ownership effects* in U.S. subjects, who held individualistic values, but not in Hong Kong subjects, who held collectivistic values. Among participants with collectivistic values, a group ownership effect occurred: Whereas it was relatively easy for these subjects to part with objects that were their own, it was difficult for them to give up an object that was owned by themselves and an ingroup member. Carnevale referred to this as a *group endowment effect*. This work points to the intriguing possibility that in negotiation, egocentric bias and self-enhancement tendencies emerge at the group rather than individual level among individuals with a collectivist background. Instead of ego defensiveness, individuals with a collectivist background may thus show ingroup defensiveness—as long as their negotiation counterpart is an ingroup member, they will be relatively cooperative, but their behavior is likely to become more competitive when their counterpart is an outgroup member. Future research needs to address this general question of cultural influences on information processing in conflict and negotiation.

Culture, Social Motives, and Strategic Choice

Given that individualism—collectivism has much in common with the proself–prosocial motive distinction discussed earlier, it comes as no surprise that this culture dimension strongly influences strategic choice in negotiation and conflict. Tjosvold, Hui, Ding, and Hu (2003) have shown that Chinese participants often approached conflicts in a problem-solving way, leading to enhanced team effectiveness. Other work also shows that collectiv-

ist individuals are more inclined to make concessions, to engage in problem solving, or to accept a 50–50 compromise. This tendency holds especially when the counterpart is not seen as outgroup—when this is the case, individuals from collectivist backgrounds tend to become more hostile and competitive than those with an individualist background (Carnevale & Leung, 2001; Gelfand & Brett, 2004).

Related to individualism–collectivism is the distinction between low- versus high-context cultures (Triandis, 1989). In high-context cultures, such as Japan and China, meaning is communicated not only by a person's words or acts but also by the context in which those words or acts are communicated. Many Western cultures (e.g., United States and Germany) are low-context cultures, in which meaning is carried by words or acts and communication is typically direct. Negotiators from the United States indeed rely on logic and reasoning to make their point, whereas negotiators from Taiwan (a high-context culture) relied more on normative statements, referrals to social roles, and relationships (Drake, 1995). Negotiators from low-context cultures engage in more direct communication (e.g., saying "no") than negotiators from high-context cultures (e.g., Adair, Okumura, & Brett, 2001). All in all, these studies thus show that culture influences social motives—the tendency to focus on own, or on own and other's outcomes—and impression management and reputation concerns.

Culture and Emotion

Culture, finally, may shape the intra- and interpersonal effects of emotion in negotiation. Thus far the interpersonal effects of emotions in negotiations have been investigated in individualistic cultures like the Netherlands (Van Kleef et al., 2004a, 2004b), the United States (Friedman et al., 2004), and Italy (Van Kleef, De Dreu, Pietroni, et al., 2006). Although a considerable body of research has investigated cultural variations in antecedents, appraisal, regulation, expression, and perception of emotion (e.g., Mesquita & Frijda, 1992), little is known about cultural differences in the ways people respond *behaviorally* to others' emotions. An exception is a study by Rodriguez Mosquera, Manstead, and Fischer (2000), who showed that expressions of shame may be more appropriate in collectivistic cultures, whereas expressions of pride may be more positively received in individualistic cultures. It would be interesting to see how cultural differences influence the ways in which negotiators respond to their opponent's emotions. Do people in collectivistic cultures respond more empathically to another's worry or disappointment, or do they find these emotions inappropriate because they communicate a preoccupation with the self rather than the collective? Will collectivists exploit others' expressions of guilt or regret to the same extent as individualists, or will they be more lenient because guilt and regret signal interpersonal concern? Future research could address these and other questions regarding possible cultural influences on reactions to other people's emotional experiences and expressions.

PRINCIPLES OF NEGOTIATION: SUMMARY AND CONCLUDING THOUGHTS

We began this review by noting that much of what individuals in negotiation do is making sense of their situation. This sense-making process is closely intertwined with strategic considerations, and these in turn depend heavily on motivational goals. We organized the relevant literatures under the headers of three interrelated rubrics—the negotiator as *motivated information processor*, the negotiator as *social animal*, and the *emotional negotiator*. We examined cognitive, motivational, and affective influences on behavioral strategy and interaction processes that were captured in three basic principles:

1. *Strategic repertoire principle*: At any given point in time negotiator behavior takes one of five distinct forms—contending, conceding, inaction, compromising, and problem solving. Underlying these five behavioral strategies are two dimensions—distributive (claiming value) and integrative (creating value) strategies.
2. *Mindless matching–mindful mismatching principle*: Whereas negotiators who do not engage in deep, deliberate and strategic thinking match their counterpart's behavioral tendency, those who engage in such deep and strategic thinking choose mismatching in early and late phases of the negotiation, and matching in the middle phase.
3. *Differentiation-before-integration principle*: Negotiators engage in distributive contending early in the negotiation and in later phases switch to integrative problem solving.

The motivated information-processing perspective has resulted in a variety of cognitive and motivational barriers to constructive negotiation, including the use of inadequate cognitive heuristics, naive realism, and ego defensiveness. The detrimental effects of these processes can be reduced when negotiators have epistemic motivation and engage in deep and deliberate information processing. This happens when time pressure is mild rather than strong, when there is a power balance rather than power asymmetry, when negotiators are process accountable, or when they have low chronic need for cognitive closure. In short, this section gave rise to four basic principles:

4. *Cognitive heuristics principle*: Sense making in fuzzy negotiation situations leads individuals to rely on cognitive heuristics that provide simplified views of the negotiation. Reliance on these heuristics often speeds up efficient decision making but also undermines the quality of negotiated agreement.
5. *Naive realism principle*: Negotiators make sense of their fuzzy situations by assuming others, including their counterpart, view and think like them. This tendency leads to inadequate assumptions and inaccurate conclusions about the true state of the negotiation situation, it drives negotiators toward distributive strategies, including contending, compro-

mising, and yielding, and away from problem-solving behavior.

6. *Self-threat principle*: Because individuals desire a positive self-view, and opposition and conflict inherently form threat to the self, negotiators develop ego-defensive (i.e., hostile and competitive) reactions vis-à-vis their counterpart. Self-threat drives toward contending (and perhaps avoiding) behaviors and away from problem solving and yielding. Self-threat thus is a basis for conflict escalation.

7. *Deep thinking principle*: Negotiators may be more or less motivated to develop a rich and accurate understanding of their situation and therefore engage in more or less deep, systematic, and deliberate search for, and processing of, information. Higher levels of epistemic motivation are associated with less reliance on cognitive heuristics, faster and more rigorous correction of inadequate assumptions and perceptions, and less ego defensiveness following self-threat.

Whereas the motivated information-processor perspective focused on the individual negotiator, the social animal perspective emphasizes the social context within which negotiations take place. In our review we distinguished between impression management concerns and proself versus prosocial motives. We advanced the (8) *social motives principle*: (a) Negotiators have or adopt a prosocial or a proself motivation; that (b) drives for confirmatory sense-making processes so that prosocials are more likely to see the negotiation as a collaborative game and their counterpart as a trustworthy person, whereas proselves are more likely to view the negotiation as a competitive game and their counterpart as untrustworthy; and (c) leads prosocials to engage in more problem solving when there is high resistance to concession making and more yielding when there is low resistance.

In the section on the emotional negotiator we returned to this individual–social divide when we showed that the individuals' emotions have an effect on their own cognitive and motivational processes, as well as on their counterpart's cognition, motivation, and behavioral moves. The work on interpersonal effects of emotions in negotiations suggests the intriguing possibility that information about one's counterpart's emotions is often used strategically, and that these strategic considerations are more important than, and perhaps even overrule, implicit and immediate contagion-like processes. Specifically, we formulated two basic principles, namely, the (9) *moody negotiator principle*, which holds that angry negotiators are more likely to play tough and to make small concessions than are sad negotiators, who are more evasive, and happy negotiators, who easily make concessions and are likely to collaborate; and (10) the *emotions-as-strategic-information principle*, which holds that a counterpart's emotions provide information that may serve as input in the negotiator's strategic decision-making process, but only when the negotiator is sufficiently motivated to consider the strategic implications of the other's emotion.

Stepping back from these 10 basic principles reveals that our decision to organize the literature around moti-vated information processing, social motives, and affect and emotion should not be taken as suggesting that these three components are unrelated to each other and can be understood in isolation from each other. In fact, the 3 perspectives and 10 principles are mutually related in a variety of ways: Affect and emotion influence information processing, and epistemic motivation determines to some extent whether affect and emotion can exert effects on negotiator demand and concession making. Social motives may explain why negotiators experience certain emotions, and much of the information processing in negotiation is social in nature. Given that we have developed a quite thorough understanding of the cognitive, motivational, and affective underpinnings of negotiation, future work needs to continue recent work on the interplay among these three core components (cf. De Dreu & Carnevale, 2003; Kramer & Messick, 1995; Morris & Gelfand, 2004; O'Connor, 1997).

Before concluding we wish to highlight two main questions that have thus far been ignored in negotiation research:

1. What makes people believe negotiation would work for them and solve their problem—why do people engage in negotiation?
2. What happens once the negotiation is over, and life retakes its course?

The first question basically points to *initiation motives*, which we define as those motives that people have to (not) begin negotiations with someone else (Carnevale & De Dreu, 2006). *Initiation motivation* guides the start of negotiation and has three basic forms that reflect the notion that the situation is *ripe for negotiation* (Zartman, 1991): The first is a mutually hurting stalemate, where the parties are not winning and costs are mounting. The second is an impending or narrowly avoided catastrophe, and the third is a mutually enticing opportunity, where the divergent interests stand in the way of accomplishing important aims on both sides. All three forms reflect the notion of negotiation as an adaptation to problems of interdependence, an adaptation that reflects sense making prior to rather than during negotiation.

The second main question basically points to the aftermath of negotiation and dispute resolution. Surprisingly little is known about what happens once agreement is reached, and surprisingly strong assumptions are made about the aftermath of conflict and negotiation. For example, integrative agreements are often assumed to be the best that could happen to negotiators and are recommended because they stimulate parties' self-efficacy, increase their self-esteem, renders them satisfied, and fosters economic prosperity and political stability (Rubin et al., 1994). That this thinking is wrong is suggested by a recent study showing that groups with proself-motivated negotiators showed greater dedication, functioned more effectively, and performed better than groups with prosocially motivated members in a creativity task, whereas prosocial groups outperformed proself groups on all these dimensions when the task involved planning (Beersma & De Dreu, 2005). This work thus shows that

affective, cognitive, and motivational processes during negotiation may have surprising effects on interpersonal and group processes that take shape once the negotiation is over and life retakes is normal course.

We began this chapter with the notion that negotiation is omnipresent and shapes much of our day-to-day interactions along with the more encompassing social events that affect the lives of many. The analysis of negotiation in terms of its motivational, cognitive, and affective underpinnings has generated a thorough understanding of when and why individuals (fail to) achieve mutually beneficial agreement. These insights into the psychology of negotiation may provide a fruitful basis for continued integration with work on negotiation and dispute resolution being conducted in adjacent disciplines, including economics and political science. Doing so will provide calibrated advice on negotiation and dispute resolution, move us toward more comprehensive theory, and stimulate new research in this important and exciting area.

ACKNOWLEDGMENTS

REFERENCES

Adair, W. L., Okumura, T., & Brett, J. M. (2001). Negotiation behavior when cultures collide: The United States and Japan. *Journal of Applied Psychology*, *86*(3), 371–385.

Allred, K., Mallozzi, J., Matsui, F., & Raia, C. (1997). The influence of anger and compassion on negotiation performance. *Organizational Behavior and Human Decision Processes*, *70*, 175–187.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Bacharach, S. B., & Lawler, E. J. (1981). *Bargaining: Power, tactics and outcomes*. Greenwich, CT: JAI Press.

Bar-Joseph, U., & Kruglanski, A. W. (2003). Intelligence failure and need for cognitive closure: On the psychology of the Yom Kippur surprise. *Political Psychology*, *24*, 75–99.

Baron, R. A. (1990). Environmentally induced positive affect: It's impact on self-efficacy, task performance, negotiation, and conflict. *Journal of Applied Social Psychology*, *20*, 368–384.

Barry, B., & Friedman, R. A. (1998). Bargainer characteristics in distributive and integrative negotiation. *Journal of Personality and Social Psychology*, *74*, 345–359.

Barry, B., Fulmer, I. S., & Van Kleef, G. A. (2004). I laughed, I cried, I settled: The role of emotion in negotiation. In M. J. Gelfand & J. M. Brett (Eds.), *The handbook of negotiation and culture* (pp. 71–94). Palo Alto, CA: Stanford University Press.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., vol. 2 pp. 282–316). Boston: McGraw-Hill.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, *103*, 5–33.

Bazerman, M. H., Curhan, J. R., Moore, D. A., & Valley, K. L. (2000). Negotiation. *Annual Review of Psychology*, *51*, 279–314.

Bazerman, M. H., Magliozzi, T., & Neale, M. A. (1985). Integrative bargaining in a competitive market. *Organizational Behavior and Human Decision Processes*, *35*, 294–313.

Beersma, B., & De Dreu, C. K. W. (2002). Integrative and distribu-

tive negotiation in small groups: Effects of task structure, decision rule, and social motive. *Organizational Behavior and Human Decision Processes*, *87*, 227–252.

Beersma, B., & De Dreu, C. K. W. (2005). Conflict's consequences: Effects of social motivation on post-negotiation convergent and divergent group processes and performance. *Journal of Personality and Social Psychology*, *89*, 358–374.

Beggan, J. K. (1992). On the social nature of non-social perception: The mere ownership effect. *Journal of Personality and Social Psychology*, *62*, 229–237.

Ben-Yoav, O., & Pruitt, D. (1984). Resistance to yielding and the expectation of cooperative future interaction in negotiation. *Journal of Experimental Social Psychology*, *34*, 323–335.

Bornstein, G. (2003). Intergroup conflict: Individual, group, and collective interests. *Personality and Social Psychology Review*, *7*, 129–145.

Bottom, W. (1998). Negotiator risk: Sources of uncertainty and the impact of reference points on negotiated agreements. *Organizational Behavior and Human Decision Processes*, *76*, 89–112.

Brett, J. M., Shapiro, D. L., & Lytle, A. L. (1998). Breaking the bonds of reciprocity in negotiations. *Academy of Management Journal*, *41*, 410–424.

Brodt, S. E., & Tuchinsky, M. (2000). Working together but in opposition: An examination of the "good cop/bad cop" negotiating team tactic. *Organizational Behavior and Human Decision Processes*, *81*, 155–177.

Brown, B. R. (1977). Face-saving and face-restoration in negotiation. In D. Druckman (Ed.), *Negotiations* (pp. 275–300). Beverly Hills, CA: Sage.

Burnham, T., McGabe, K., & Smith, V. L. (2000). Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior and Organization*, *43*, 57–73.

Camac, C. (1992). Information preferences in a two-person social dilemma. In W. Liebrand, D. Messick, & H. Wilke (Eds.), *Social dilemmas* (pp. 147–161). Berlin: Springer-Verlag.

Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, *3*, 23–43.

Caporael, L. R., Dawes, R., Orbell, J., & Van de Kragt, A. (1989). Pro-selfness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, *12*, 683–739.

Carnevale, P. J. (1995). Property, culture, and negotiation. In R. Kramer & D. M. Messick (Eds.), *Negotiation as a social process* (pp. 309–323). Newbury Park: CA: Sage.

Carnevale, P. J., & De Dreu, C. K. W. (2006). Motive: The negotiator's raison d'etre. In L. L. Thompson & J. Brett (Eds.), *The social psychology of negotiation* (pp. 55–76). New York: Psychology Press.

Carnevale, P. J. D., & Isen, A. M. (1986). The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. *Organizational Behavior and Human Decision Processes*, *37*, 1–13.

Carnevale, P. J., & Lawler, E. J. (1986). Time pressure and the development of integrative agreements in bilateral negotiations. *Journal of Conflict Resolution*, *30*, 636–659.

Carnevale, P. J., & Ledgerwood, S. (2004). *Endowment effects and social identity concerns*. Paper presented at the Society for Experimental Social Psychology Groups Pre-conference, Fort Worth, TX.

Carnevale, P. J., & Leung, K. (2001). Cultural dimensions of negotiation. In M. A. Hogg & Tindale, R. S. (Eds.), *Blackwell handbook of social psychology: Vol. 3. Group processes* (pp. 482–496). Oxford, UK: Blackwell.

Carnevale, P. J., & Probst, T. M. (1998). Social values and social conflict in creative problem solving and categorization. *Journal of Personality and Social Psychology*, *74*, 1300–1309.

Carnevale, P. J., & Pruitt, D. G. (1992). Negotiation and mediation. *Annual Review of Psychology*, *43*, 531–582.

Carnevale, P. J. D., Pruitt, D. G., & Britton S. D. (1979). Looking tough: The negotiator under constituent surveillance. *Personality and Social Psychology Bulletin*, *5*, 118–121.

Carnevale, P. J., Pruitt, D. G., & Seilheimer, S. (1981). Looking and competing: Accountability and visual access in integrative bargaining. *Journal of Personality and Social Psychology*, *40*, 111–120.

Chaiken, S., Gruenfeld, D. H., & Judd, C. M. (2001). Persuasion in negotiations and conflict situations. In M. Deutsch & P. Coleman (Eds.), *Conflict resolution*. (pp. 67–89). New York: Guilford Press.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chen, C. C., Chen, X., & Meindl, J. R. (1998). How can cooperation be fostered? The cultural effects of individualism-collectivism. *Academy of Management Review*, *23*, 285–304.

Chen, Y. R., Mannix, E. A., & Okumura, T. (2003). The importance of who you meet: Effects of self- versus other-concerns among negotiators in the United States, the People's Republic of China, and Japan. *Journal of Experimental Social Psychology*, *39*, 1–15.

Clark, M. S., Pataki, S. P., & Carver, V. H. (1996). Some thoughts and findings on self-presentation of emotions in relationships. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 247–274). Hillsdale, NJ: Erlbaum.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*, 169–193.

De Dreu, C. K. W. (2003). Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes*, *91*, 280–295.

De Dreu, C. K. W. (2004). Motivation in negotiation: A social psychological analysis. In M. J. Gelfand & J. M. Brett (Eds.), *The handbook of negotiation and culture* (pp. 114–135). Stanford, CA: Stanford University Press.

De Dreu, C. K. W. (2005). A PACT against conflict escalation in negotiation and dispute resolution. *Current Directions in Psychological Science*, *14*, 149–152.

De Dreu, C. K. W., & Beersma, B. (2005). Conflict in organizations: Beyond effectiveness and performance. *European Journal of Work and Organizational Psychology*, *14*, 105–117.

De Dreu, C. K. W., Beersma, B., Stroebe, K., & Euwema, M. C. (2006). Motivated information processing, strategic choice, and the quality of negotiated agreement. *Journal of Personality and Social Psychology*, *90*, 927–943.

De Dreu, C. K. W., & Boles, T. (1998). Share and share alike or winner takes all? Impact of social value orientation on the choice and recall of decision heuristics in negotiation. *Organizational Behavior and Human Decision Processes*, *76*, 253–267.

De Dreu, C. K. W., & Carnevale, P. J. (2003). Motivational bases of information processing and strategy in conflict and negotiation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 235–291). New York: Academic Press.

De Dreu, C. K. W., Carnevale, P. J. D., Emans, B. J. M., & Van de Vliert, E. (1994). Effects of gain loss frames in negotiation: Loss aversion, mismatching, and frame adoption. *Organizational Behavior and Human Decision Processes*, *60*, 90–107.

De Dreu, C. K. W., Evers, A., Beersma, B., Kluwer, E. S., & Nauta, A. (2001). A theory-based measure of conflict management strategies in the work place. *Journal of Organizational Behavior*, *22*, 645–668.

De Dreu, C. K. W., Giebels, E., & Van de Vliert, E. (1998). Social motives and trust in integrative negotiation: The disruptive effects of punitive capability. *Journal of Applied Psychology*, *83*, 408–422.

De Dreu, C. K. W., Koole, S., & Oldersma, F. L. (1999). On the seizing and freezing of negotiator inferences: Need for cognitive closure moderates the use of heuristics in negotiation. *Personality and Social Psychology Bulletin*, *25*, 348–362.

De Dreu, C. K. W., Koole, S., & Steinel, W. (2000). Unfixing the fixed-pie: A motivated information processing account of integrative negotiation. *Journal of Personality and Social Psychology*, *79*, 975–987.

De Dreu, C. K. W., & McCusker, C. (1997). Gain-loss frames in two-person social dilemmas: A transformational analysis. *Journal of Personality and Social Psychology*, *72*, 1093–1106.

De Dreu, C. K. W., & Van Kleef, G. A. (2004). The influence of power on the information search, impression formation, and demands in negotiation. *Journal of Experimental Social Psychology*, *40*, 303–319.

De Dreu, C. K. W., & Van Knippenberg, D. (2005). The possessive self as a barrier to constructive conflict management: Effects of mere ownership, process accountability, and self-concept clarity on competitive cognitions and behavior. *Journal of Personality and Social Psychology*, *89*, 345–357.

De Dreu, C. K. W., & Van Lange, P. (1995). The impact of social value orientations on negotiator cognition and behavior. *Personality and Social Psychology Bulletin*, *21*, 1178–1188.

De Dreu, C. K. W., Weingart, L. R., & Kwon, S. (2000). Influence of social motives on integrative negotiation: A meta-analytic review and test of two theories. *Journal of Personality and Social Psychology*, *78*, 889–905.

De Dreu, C. K. W., Yzerbyt, V. Y., & Leyens, J.-P. (1995). Dilution of stereotype-based cooperation in mixed-motive interdependence. *Journal of Experimental Social Psychology*, *31*, 575–593.

Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven, CT: Yale University Press.

Diekmann, K. A., Tenbrunsel, A. E., & Galinsky, A. D. (2003). From self-prediction to self-defeat: Behavioral forecasting, self-fulfilling prophecies, and the effect of competitive expectations. *Journal of Personality and Social Psychology*, *85*, 672–683.

Drake, L. E. (1995). Negotiation styles in intercultural communication. *International Journal of Conflict Management*, *6*, 72–90.

Druckman, D. (1986). Stages, turning points, and crises: Negotiating military base rights, Spain and the United States. *Journal of Conflict Resolution*, *30*, 327–360.

Druckman, D. (1994). Determinants of compromising behavior in negotiation: A meta-analysis. *Journal of Conflict Resolution*, *38*, 507–556.

Emerson, R. M. (1972). Exchange theory, Part I: A psychological basis for social exchange. In J. Berger, M. Zelditch, & B. Anderson (Eds.), *Sociological theories in progress* (Vol. 2, pp. 38–57). Boston: Houghton-Mifflin.

Enzle, M., Harvey, M., & Wright, E. (1992). Implicit role obligations versus social responsibility in constituency representation. *Journal of Personality and Social Psychology*, *62*, 238–245.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Forgas, J. P. (1998). On feeling good and getting your way: Mood effects on negotiator cognition and behavior. *Journal of Personality and Social Psychology*, *74*, 565–577.

Forgas, J. P., & Cromer, M. (2004). On being sad and evasive: Affective influences on verbal communication strategies in conflict situations. *Journal of Experimental Social Psychology*, *40*, 511–518.

Friedman, R., Anderson, C., Brett, J. M., Olekalns, M., & Lisco, C. C. (2004). The positive and negative effects of anger on dispute resolution: Evidence from electronically mediated disputes. *Journal of Applied Psychology*, *89*, 369–376.

Frijda, N. H. (1993). Moods, emotion episodes, and emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 381–403). New York: Guilford Press.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, *81*, 657–669.

Galinsky, A. D., Mussweiler, T., & Medvec, V. (2002) Disconnecting outcomes and evaluations: The role of negotiator focus. *Journal of Personality and Social Psychology*, *83*, 1131–1140.

Galinsky, A. D., Seiden, V. L., Kim, P. H., & Medvec, V. (2002). The dissatisfaction of having your first offer accepted: The role of counterfactual thinking in negotiations. *Personality and Social Psychology Bulletin*, *28*, 271–283.

Gelfand, M. J., & Brett, J. M. (Eds.). (2004). *The handbook of negotiation and culture*. Stanford, CA: Stanford University Press.

Gelfand, M. J., & Christakopoulou, S. (1999). Culture and negotiator cognition: Judgment accuracy and negotiation processes in

individualistic and collectivistic cultures. *Organizational Behavior and Human Decision Processes*, 79, 248–269.

Gelfand, M. J., Higgins, M., Nishii, L., Raver, J., Dominguez, A., Yamaguchi, S., et al. (2001). Culture and egocentric biases of fairness in conflict and negotiation. *Journal of Applied Psychology*, 86, 1059–1074.

Gelfand, M. J., & Morris, M. M. (2004). Cultural differences and cognitive dynamics. In M. J. Gelfand & J. M. Brett (Eds.), *The handbook of negotiation and culture* (pp. 45–70). Stanford, CA: Stanford University Press.

Gelfand, M. J., & Realo, A. (1999). Individualism–collectivism and accountability in intergroup negotiations. *Journal of Applied Psychology*, 84, 721–736.

Giebels, E., De Dreu, C. K. W., & Van de Vliert, E. (2000). Interdependence in negotiation: Effects of exit options and social motive on distributive and integrative negotiation. *European Journal of Social Psychology*, 30, 255–272.

Gillespie, J. J., Brett, J., & Weingart, L . (2000). Interdependence, social motives, and outcome satisfaction in multiparty negotiation. *European Journal of Social Psychology*, 30, 779–797.

Glimcher, P. W. (2003). *Decisions, uncertainty, and the brain: The science of neuroeconomics*. Cambridge, MA: MIT Press.

Golec, A., & Federico, C. (2004). Understanding responses to political conflict: Interactive effects of need for cognitive closure and salient conflict schemas. *Journal of Personality and Social Psychology*, 87, 750–762.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25, 161–178.

Graziano, W. G., Jensen-Campbell, L., & Hair, E. (1996). Perceiving interpersonal conflict and reacting to it: The case for agreeableness. *Journal of Personality and Social Psychology*, 70, 820–835.

Guth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining, *Journal of Economic Behavior and Organization*, 3, 367–388.

Harinck, F., & De Dreu, C. K. W. (2004). Negotiating values or resources: The moderating impact of time pressure. *European Journal of Social Psychology*, 34, 595–612.

Harinck, F., De Dreu, C. K. W., & Van Vianen, A. E. M. (2000). The impact of issue content on fixed-pie perceptions, problem solving and integrative outcomes in negotiation. *Organizational Behavior and Human Decision Processes*, 81, 329–358.

Hegtvedt, K. A., & Killian, C. (1999). Fairness and emotions: Reactions to the process and outcomes of negotiations. *Social Forces*, 78, 269–302.

Hilty, J. A., & Carnevale, P. J. (1992). Black-hat/white-hat strategy in bilateral negotiation. *Organizational Behavior and Human Decision Processes*, 55, 444–469.

Hofstede, G. (1991). *Culture's consequences*. London: Sage.

Huber, V., & Neale, M. (1986). Effects of cognitive heuristics and goals on negotiator performance and subsequent goal setting. *Organizational Behavior and Human Decision Processes*, 38, 342–365.

Hulbert, L. G., Correa da Silva, M., & Adegboyega, G. (2001). Cooperation in social dilemmas and allocentrism: A social values approach. *European Journal of Social Psychology*, 31, 641–658.

Jordan, J. M., & Roloff, M. E. (1997). Planning skills and negotiator goal accomplishment. *Communication Research*, 24, 31–63.

Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98, 1325–1348.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.

Kelley, H. H. (1966). A classroom study of the dilemmas in interpersonal negotiations. In K. Archibald (Ed.), *Strategic interaction and conflict: Original papers and discussion*. Berkeley, CA: Institute of International Studies.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, 16, 66–91.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion*, 13, 505–521.

Kim, P. H. (1997). Strategic timing in group negotiations: The implications of forced entry and forced exit for negotiators with unequal power. *Organizational Behavior and Human Decision Processes*, 71, 263–286.

Komorita, S. S., Parks, C., & Hulbert, L. (1992). Reciprocity and the induction of cooperation in social dilemmas. *Journal of Personality and Social Psychology*, 62, 607–617.

Kramer, R. M., & Messick, D. M. (1995). *Negotiation as a social process*. London: Sage.

Kristensen, H., & Gärling, T. (1997). The effects of anchor points and reference points on negotiation process and outcome. *Organizational Behavior and Human Decision Processes*, 71, 85–94.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, 103, 263–283.

Kwon, S., & Weingart, L. R. (2004). Unilateral concessions from the other party: Concession behavior, attributions, and negotiation judgments. *Journal of Applied Psychology*, 89, 263–278.

Langner, C. A., & Winter, D. G. (2001). The motivational basis of concessions and compromise: Archival and laboratory studies. *Journal of Personality and Social Psychology*, 81, 711–727.

Larrick, R. P., & Blount, S. (1997). The claiming effect: Why players are more generous in social dilemmas than in ultimatum games. *Journal of Personality and Social Psychology*, 72, 810–825.

Lax, D. A., & Sebenius, J. K. (1986). *The manager as negotiator: Bargaining for cooperation and competitive gain*. New York: Free Press.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125, 255–275.

Leung, K. (1987). Some determinants of reactions to procedural models for conflict resolution: A cross-national study. *Journal of Personality and Social Psychology*, 53, 898–908.

Liebert, R. M., Smith, W. P., Hill, J. H., & Keiffer, M. (1968). The effects of information and magnitude of initial offer on interpersonal negotiation. *Journal of Experimental Social Psychology*, 4, 431–441.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Upper Saddle River, NJ: Prentice-Hall.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, 66, 37–47.

Mo'az, I., Ward, A., Katz, M., & Ross, L. (2002). Reactive devaluation of an "Israeli" vs. "Palestinian" peace proposal. *Journal of Conflict Resolution*, 46(4), 515–546.

McClintock, C. (1977). Social motives in settings of outcome interdependence. In D. Druckman (Ed.), *Negotiations: Social psychological perspective* (pp. 49–77). Beverly Hills, CA: Sage.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin*, 112, 179–204.

Messick, D. M. (1993). Equality as a decision heuristic. In B. A. Mellers & J. Baron (Eds.), *Psychological perspectives on justice: Theory and applications* (pp. 11–31). New York: Cambridge University Press.

Mikolic, J. M., Parker, J. C., & Pruitt, D. G. (1997). Escalation in response to persistent annoyance: Groups vs. individuals and gender effects. *Journal of Personality and Social Psychology*, 72, 151–163.

Moore, D. A. (2004a). Myopic prediction, self-destructive secrecy, and the unexpected benefits of revealing final deadlines in negotiation. *Organizational Behavior and Human Decision Processes*, 94, 125–139.

Moore, D. A. (2004b). The unexpected benefits of final deadlines in negotiation. *Journal of Experimental Social Psychology*, 40, 121–127.

Moran, S., & Ritov, I. (2002). Initial perceptions in negotiations:

Evaluation and response to "logrolling" offers. *Journal of Behavioral Decision Making*, *15*, 101–124.

Morris, M. W., & Keltner, D. (2000). How emotions work: An analysis of the social functions of emotional expression in negotiations. *Research in Organizational Behavior*, *22*, 1–50.

Morris, M. W., Larrick, R. P., & Su, S. K. (1999). Misperceiving negotiation counterparts: When situationally determined bargaining behaviors are attributed to personality traits. *Journal of Personality and Social Psychology*, *77*, 52–67.

Murnighan, J. K., Babcock, L., Thompson, L., & Pillutla, M. (1999). The information dilemma in negotiations: Effects of experience, incentives and integrative potential. *International Journal of Conflict Management*, *10*, 313–339.

Naquin, C. E. (2003). The agony of opportunity in negotiation: Number of negotiable issues, counterfactual thinking, and feelings of satisfaction. *Organizational Behavior and Human Decision Processes*, *91*, 97–107.

Neale, M. A., & Bazerman, M. H. (1985). The effects of framing and negotiator overconfidence on bargaining behaviors and outcomes. *Academy of Management Journal*, *28*(1), 34–49.

Neale, M. A., & Bazerman, M. H. (1991). *Rationality and cognition in negotiation*. New York: Free Press.

Nesselroade, K. P., Jr., Beggan, J. K., & Allison, S. T. (1999). Possession enhancement in an interpersonal context: An extension of the mere ownership effect. *Psychology and Marketing*, *16*, 21–34.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, *39*, 84–97.

O'Connor, K. (1997). Motives and cognitions in negotiation: A theoretical integration and empirical test. *International Journal of Conflict Management*, *8*, 114–131.

O'Connor, K. M., & Arnold, J. A. (2001). Distributive spirals: Negotiation impasses and the moderating role of disputant self-efficacy. *Organizational Behavior and Human Decision Processes*, *84*, 148–176.

Ohbuchi, K. I., & Fukushima, O. (1997). Personality and interpersonal conflict: Aggressiveness, self-monitoring, and situational variables. *International Journal of Conflict Management*, *8*, 99–113.

Ohtsubo, Y., & Kameda, T. (1998). The function of equality heuristic in distributive bargaining: Negotiated allocation of costs and benefits in a demand revelation context. *Journal of Experimental Social Psychology*, *34*, 90–108.

Okhuysen, G. A., Galinsky, A. D., & Uptigrove, T. A. (2003). Saving the worst for last: The effect of time horizon on the efficiency of negotiating benefits and burdens. *Organizational Behavior and Human Decision Processes*, *91*(2), 269–279.

Olekalns, M., Brett, J. M., & Weingart, L. R. (2003). Phases, transitions, and interruptions. Modelling processes in multi-party negotiations. *International Journal of Conflict Management*, *14*, 191–211.

Olekalns, M., Smith, P., & Walsh, T. (1996). The process of negotiating: strategy and timing as predictors of outcomes. *Organizational Behavior and Human Decision Processes*, *68*, 68–77.

Ostrom, E. (1998). A behavioral approach to the rational choice theory of collective action. *American Political Science Review*, *92*, 1–22.

Paese, P. W., & Yonker, R. D. (2001). Toward a better understanding of egocentric fairness judgments in negotiation. *International Journal of Conflict Management*, *12*, 97–186.

Parks, C. D., Henager, R. F., & Scamahorn, S. D. (1996). Trust and reactions to messages of intent in social dilemmas. *Journal of Conflict Resolution*, *40*, 134–151.

Peterson, E., & Thompson, L. (1997). Negotiation teamwork: The impact of information distribution and accountability on performance depends on the relationship among team members. *Organizational Behavior and Human Decision Processes*, *72*, 364–383.

Pillutla, M. M., & Murnighan, J. K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, *68*, 208–224.

Pinkley, R. L. (1995). Impact of knowledge regarding alternatives to settlement in dyadic negotiations: Whose knowledge counts? *Journal of Applied Psychology*, *80*, 403–417.

Pinkley, R. L., Griffith, T. L., & Northcraft, G. B. (1995). "Fixed pie" al mode: Information availability, information processing, and the negotiation of suboptimal agreements. *Organizational Behavior and Human Decision Processes*, *62*, 101–112.

Polzer, J. T., Mannix, E. A., & Neale, M. A. (1998). Interest alignment and coalitions in multiparty negotiation. *Academy of Management Journal*, *41*, 42–54.

Probst, T., Carnevale, P. J., & Triandis, H. (1999). Cultural values in intergroup and single-group social dilemmas. *Organizational Behavior and Human Decision Processes*, *77*, 171–191.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*, 781–799.

Pruitt, D. G. (1981). *Negotiation behavior*. New York: Academic Press.

Pruitt, D. G. (1998). Social conflict. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., vol. 2, pp. 89–150). New York: McGraw-Hill.

Pruitt, D. G., & Carnevale, P. J. (1993). *Negotiation in social conflict*. Buckingham, UK: Open University Press.

Pruitt, D. G., & Lewis, S. A. (1975). Development of integrative solutions in bilateral negotiation. *Journal of Personality and Social Psychology*, *31*, 621–633.

Pruitt, D. G., & Rubin, J. Z. (1986). *Social conflict: escalation, stalemate, and settlement*. New York: McGraw-Hill.

Pruitt, D. G., & Syna, H. (1985). Mismatching the opponent's offers in negotiation. *Journal of Experimental Social Psychology*, *21*, 103–113.

Rahim, A., & Magner, N. R. (1995). Confirmatory factor analysis of the styles of handling interpersonal conflict: First-order factor model and its invariance across groups. *Journal of Applied Psychology*, *80*, 122–132.

Raiffa, H. (1982). *The art and science of negotiation*. Cambridge, MA: Belknap.

Rhoades, J. A., & Carnevale, P. J. (1999). The behavioral context of strategic choice in negotiation: A test of the dual concern model. *Journal of Applied Social Psychology*, *29*, 1777–1802.

Robert, C., & Carnevale, P. J. (1997). Group choice in ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, *72*, 256–279.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naïve realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology*, *68*, 404–417.

Rodriguez, P. M., Manstead, A. S. R., & Fischer, A. H. (2002). The role of honor concerns in emotional reactions to offences. *Cognition and Emotion*, *16*, 143–163.

Ross, L., & Ward, A. (1995). Psychological barriers to dispute resolution. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 255–304). San Diego, CA: Academic Press.

Rubin, J. Z., & Brown, R. (1975). *Social psychology of bargaining and negotiation*. New York: Academic Press.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict; escalation, stalemate, and settlement*. New York: McGraw-Hill.

Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Siegel, A. E., & Fouraker, L. E. (1960). *Bargaining and group decision making*. New York: McGraw-Hill.

Smith, D. L., Pruitt, D. G., & Carnevale, P. J. (1982). Matching and mismatching: The effect of own limit, other's toughness, and time pressure on concession rate in negotiation. *Journal of Personality and Social Psychology*, *42*, 876–883.

Snyder, M. (1992). Motivational foundations of behavioral confirmation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 67–114). New York: Academic Press.

Steinel, W., & De Dreu, C. K. W. (2004). Social motives and strategic misrepresentation in social decision making. *Journal of Personality and Social Psychology*, *86*, 419–434.

Tetlock, P. E. (1992). The impact of accountability on judgment

and choice: Toward a social contingency model. L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331–376). New York: Academic Press.

Tetlock, P. E., & Manstead, A. S. (1985) Impression management versus intrapsychic explanations in social psychology: A useful dichotomy? *Psychological Review*, *92*, 59–77.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Thompson, L. L. (1990). Negotiation behavior and outcomes: Empirical evidence and theoretical issues. *Psychological Bulletin*, *108*, 515–532.

Thompson, L. L., & Hastie, R. (1990). Social perception in negotiation. *Organizational Behavior and Human Decision Processes*, *47*, 98–123.

Thompson, L. L., & Hrebec, D. (1996). Lose-lose agreements in interdependent decision making. *Psychological Bulletin*, *120*, 396–409.

Thompson, L. L., Kramer, R. M., & Valley, K. L. (1995). The bittersweet feeling of success: An examination of social perception in negotiation. *Journal of Experimental Social Psychology*, *21*, 1–26.

Thompson, L. L., & Loewenstein, G. (1992). Egocentric interpretations of fairness in interpersonal conflict. *Organizational Behavior and Human Decision Processes*, *51*, 176–197.

Tinsley, C. H., & Brett, J. M. (2001). Managing workplace conflict in the United States and Hong Kong. *Organizational Behavior and Human Decision Processes*, *85*, 360–381.

Tinsley, C. H., O'Connor, K. M., & Sullivan, B. (2002). Tough guys finish last: The perils of a distributive reputation. *Organizational Behavior and Human Decision Processes*, *88*, 621–642.

Tjosvold, D., Hui, C., Ding, D. Z., & Hu, J. (2003). Conflict values and team relationships: Conflict's contribution to team effectiveness and citizenship in China. *Journal of Organizational Behavior*, *24*, 69–88.

Triandis, H. C. (1989). The self and social behavior in different cultural contexts. *Psychological Review*, *96*, 506–520.

Triandis, H. C., Carnevale, P. J., Gelfand, M., Robert, C., Wasti, A., Probst, T. M., et al. (2001). Culture, personality, and deception in intercultural management negotiations. *International Journal of Cross Cultural Management*, *1*, 73–90.

Trötschel, R., & Gollwitzer, P. M. (in press). Implementation intentions and the willful pursuit of prosocial goals in negotiations. *Journal of Experimental Social Psychology*.

Turner, J. C. (1987). *Rediscovering the social group: Self-categorization theory*. Oxford, UK: Blackwell.

Tyler, T., Huo, Y. J., & Lind, E. A. (1999). The two psychologies of conflict resolution: Differing antecedents of pre-experience choices and post-experience evaluations. *Group Processes and Intergroup Relations*, *2*, 99–118.

Tyler, T., Lind, E. A., & Huo, Y. J. (2000). Cultural values and authority relations. *Psychology, Public Policy, and Law*, *6*, 1138–1163.

Van Beest, I., Wilke, H., & Van Dijk, E. (2003). The excluded player in coalition formation. *Personality and Social Psychology Bulletin*, *29*, 237–247.

Van de Vliert, E. (1997). *Complex interpersonal conflict behavior*. London: Psychology Press.

Van Dijk, E., De Cremer, D., & Handgraaf, M. J. J. (2004). Social value orientations and the strategic use of fairness in ultimatum bargaining. *Journal of Experimental Social Psychology*, *40*, 697–707.

Van Kleef, G. A., & De Dreu, C.K.W. (2002). Social value orientation and impression formation: A test of two competing hypotheses about information search in negotiation. *International Journal of Conflict Management*, *13*, 59–77.

Van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. (2004a). The interpersonal effects of anger and happiness in negotiations. *Journal of Personality and Social Psychology*, *86*, 57–76.

Van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. (2004b). The interpersonal effects of emotions in negotiations: A motivated information processing approach. *Journal of Personality and Social Psychology*, *87*(4), 510–528.

Van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. S. R. (2006). Supplication and appeasement in conflict and negotiations: The interpersonal effects of disappointment, worry, guilt, and regret. *Journal of Personality and Social Psychology*, *91*, 124–142.

Van Kleef, G. A., De Dreu, C. K. W., Pietroni, D., & Manstead, A. S. R. (2006). Power and emotion in negotiation: Power moderates the interpersonal effects of anger and happiness in concession making. *European Journal of Social Psychology*, *36*, 557–581.

Van Kleef, G. A., Steinel, W., van Knippenberg, D., Hogg, M. A., & Svensson, A. (in press). Group member prototypicality and intergroup negotiation: How one's standing in the group affects negotiation behavior. *British Journal of Social Psychology*.

Van Lange, P. A. (1992). Confidence in expectations: A test of the triangle hypothesis. *European Journal of Personality*, *6*, 371–379.

Wall, J. A., Jr. (1991). Impression management in negotiations. In R. A. Giacalone & P. Rosenfeld (Eds.), *Applied impression management: How image-making affects managerial decisions* (pp. 133–156). Thousand Oaks, CA: Sage.

Walton, R. E., & McKersie, R. (1965). *A behavioral theory of labor negotiations: An analysis of a social interaction system*. New York: McGraw-Hill.

Webster, D., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*, 1049–1062.

Webster, D. M., Richter, L., & Kruglanski, A. W. (1996). On leaping to conclusions when feeling tired: Mental fatigue effects on impressional primacy. *Journal of Experimental Social Psychology*, *32*, 181–195.

Weingart, L. R., Bennett, R. J., & Brett, J. M. (1993). The impact of consideration of issues and motivational orientation on group negotiation process and outcome. *Journal of Applied Psychology*, *78*, 504–517.

Weingart, L. R., Hyder, E., & Prietula, M. J. (1996). Knowledge matters: The effect of tactical descriptions on negotiation behavior and outcome. *Journal of Personality and Social Psychology*, *70*, 1205–1217.

Weingart, L. R., Prietula, M. J., Hyder, E. B., & Genovese, C. R. (1999). Knowledge and the sequential processes of negotiation: A Markov chain analysis of response-in-kind. *Journal of Experimental Social Psychology*, *35*, 366–393.

Whyte, G., & Sebenius, J. K. (1997). The effect of multiple anchors on anchoring in individual and group judgment. *Organizational Behavior and Human Decision Processes*, *69*, 75–85.

Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A., & Schopler, J. (2003). Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect. *Psychological Bulletin*, *129*, 698–722.

Zartman, I. W. (1991). Regional conflict resolution. In V. A. Kremenyuk (Ed.), *International negotiation: Analysis, approaches, issues* (pp. 302–314). San Francisco: Jossey-Bass.

Zetik, D. C., & Stuhlmacher, A. F. (2002). Goal setting and negotiation performance: A meta-analysis. *Group Processes and Intergroup Relations*, *5*, 35–52.

Grounding Communication

Synchrony

GÜN R. SEMIN

FOCUS

Any close inspection of how face-to-face human communication is possible reveals that it is not short of a miracle unique to the species. It is a species-specific endowment that in its diverse instantiations can be the source of diversity, conflict, and strife or merriment, harmony, and creativity. Communication and its potential pitfalls occupy considerable space in popular books, gossip columns, and recommendations about relationship maintenance or political advancement. In social psychology, analyses of diverse socially relevant phenomena such as intergroup communication, discrimination, and gender differences constitute but a few of the compelling and inviting themes that have attracted considerable empirical investigation. While these types of investigations have uncovered interesting phenomena a scientific journey that takes us from the roots of communication to its remarkable manifestations reveals the extraordinary biological, neural, and social composition of what makes human communication possible. Although communication plays second fiddle to a great number of issues that are central to current social psychological research, it is undoubtedly the case that communication constitutes the most fundamental question that social psychology can address because human communication constitutes the core to social life as we know it.

Communication is not a biologically unique phenomenon. It is an endowment that a great number of species have and in each case it has its unique specialization (cf. Seyfarth & Cheney, 2003). This can be seen in the distinc-

tive types of dedicated media (e.g., gestural, acoustic, chemical, optical, mechanical, or even electrical as in the case of electrical fish) serving communication and that are reportedly idiosyncratic for different species. For example, vervet monkeys have three types of alarm cries (Seyfarth, Cheney, & Marler, 1980) that are specialized for three different kinds of predators (snakes, leopards, and eagles), which they call out when they see one of the predators. The respective cries engage the vervets to look down for snakes, run into trees (leopard alarms), or look up (eagle alarms). Bees use chemicals and movement for communication purposes. Honeybees (*Apis Melifera*) communicate via a number of different movements, grouped into dances (von Frisch, 1947, 1967).[1] Termites appear to use a sophisticated chemical (pheromone) communication system regulating social relations in the nest (e.g., Pasteels & Bordereau, 1998) and nest-building behavior (Beckers, Holland, & Deneubourg, 1994), while male stickleback fish use visual displays to initiate reproduction (Tinbergen, 1952). The particular specialization by means of which communication is achieved takes the form of a specific *shared medium* in each species, may this be acoustic, chemical, mechanical, optical, or some combination thereof. A large number of communication systems exist that are unique to the respective species, recruit unique resources, and remain largely incomprehensible between different species.

Explaining communication requires answers to two questions. The first is: "What makes communication possible?" The answer to this involves explaining how two (or more) organisms—adapted to group living—are con-

nected because successful group living requires that a member has to be "coupled" with another one in a two-way interaction (Semin, 2000). This means two (or more) members are linked together by some *common* or *symmetrical base*, namely, a base that puts them on comparable footing, that gives partial mutual access to internal states between them, thereby making it possible for the organisms to "communicate." Whatever specialization is deployed to achieve communication, and whatever the unique signal patterns that a species utilizes, *the chief function of these signals is to establish equivalence or parity* between *a sender and receiver*. The common denominator to any communication is the recruitment of processes and a medium shared by producer and perceiver (production and perception) by means of which producer and receiver is put on the same footing (cf. Liberman & Mattingly, 1985; Liberman & Whalen, 2000). This is a general requirement for any successful communication.

The second question is: "What are the biological, neural, psychological, and social foundations of *human* communication?" While the answer to the first question is about a necessary requirement for communication in general—irrespective of species—the second is about the specific processes that ground human communication. Research attempting to answer these questions in the social psychology of communication has been driven by a predominant focus on language. The issues driving such research have been how parity is established by means of language or how parity efforts are manifested in language. A significant proportion of this work has relied on the transmission of "representations" in interpersonal communication (see Krauss & Chiu, 1998; Krauss & Fussell, 1996, for reviews). A recurrent theme is what can be referred to as coordination in communication and has to do with the different ways in which representational correspondence can be established, for instance, in joint action. This theme has been approached from a variety of angles. Thus, audience design (see Krauss & Chiu, 1998; Krauss & Fussell, 1996), referential communication (Clark & Brennan, 1991; Fussell & Krauss, 1989a, 1989b), and grounding (Clark, 1996a, 1996b; Clark & Brennan, 1991; Clark & Wilkes-Gibbs, 1986; Keysar, 1997) are some of the converging approaches about how representational correspondence between the members to a dialogue is achieved. The actions and processes servicing coordination address the explicit or manifest face of how equivalence is obtained in interpersonal communication. The medium by which these actions and processes are realized is language and is therefore also a central feature of research on interpersonal communication. However, language and representational correspondence are not all there is to human communication.

A somewhat neglected aspect of communication is the embodied grounding of this process. Communication is a process drawing on experiences that originate from bodily interactions (e.g., emotional expressions, articulatory gestures, and bodily movements) and is grounded by the socially tuned architectures of our body and brain. It is a *multimodal* process involving a diversity of *neural and motor* processes (perceptible, coordinated bodily *productions* such as articulatory gestures, facial expressions,

and emotional displays, as well as "imperceptible" neural states) aside from language. These processes are involved in both the production and *perception* sides of communication. If we are to understand this miracle then we need to examine the biological, neural, psychological and social foundations of communication and this is the central theme of this chapter. Thus, this chapter provides a framework of the embodied grounding (Smith & Semin, 2004) of communication rather than reviewing the diverse classic thematic foci in the social psychology of communication, which constitute deep research pockets in their own right (see Krauss & Fussell, 1996, for a review).

FROM PARITY TO SYNCHRONIZATION

As I argued previously, the parity requirement to communication necessitates the presence of processes that are recruited symmetrically by both the sender and receiver of a communicative act. In this chapter, I argue that the psychological and neural mechanisms that constitute an embodied grounding (Smith & Semin, 2004) of communication rely on a unique knowledge system. The unique features of this knowledge system are that it is specialized for social cognition and is grounded on processes that give privileged and immediate mutual access by *coupling* two agents. It is only through *jointly* recruited processes that it is possible for two agents to be coupled and put on the same footing. In other words, such mechanisms must facilitate obtaining a state of equivalence between two parties: What counts for one member must also count for the other. I refer to this process as *synchronization*. Synchronization gives simultaneous partial mutual access to internal states, thereby establishing a type of knowledge that is different from knowledge about manipulable objects, spatial orientation, or numerosities. I regard *a jointly and simultaneously recruited process* by means of which correspondence is established at neural, perceptual, affective, and behavioral levels between producer and perceiver as synchronization (cf. Semin & Cacioppo, 2005). Moreover, I posit that synchronization precedes communication by means of language both evolutionarily and ontogenetically. It is a process that occurs without the presence of explicit communicative intent.

What evidence is there for synchronization? There is an unusually rich literature and substantial empirical research across diverse areas that have a direct bearing to synchronization processes at neural, emotional, and behavioral levels. This research literature has not been conceptualized as integral processes in the production of synchronization. A surprising facet of the research I review below is that its foundational significance for communication has not been realized. This is largely due to the fact that the terminological framework within which these diverse areas were developed has a perspective that is committed to the individual as the analytic unit and consequently focuses on individual driven processes rather than social, namely, jointly recruited, processes. This applies to broad research fields driven with con-

structs such as mirroring, simulating, and resonance upon perceiving another's action (e.g., Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996); the extensive social psychological work on mimicking, matching another's behavior (e.g., Bavelas, Black, Chovil, Lemery, & Mullett, 1988; Bernieri, 1988; Bernieri, Reznick, & Rosenthal, 1988; Chartrand & Bargh, 1999); and the voluminous work on imitation, empathy, sympathy (e.g., Heyes, 2001; Iacobini, 2005; Meltzoff & Moore, 1997; Preston & de Waal, 2002; Tomasello, 1998), or emotion contagion (Hatfield, Cacioppo, & Rapson, 1994). As the central constructs in these converging fields all indicate, the analytic anchor is located at the level of the *individual*, who is imitating somebody else, or being emotionally taken over by the feelings of the other, simulating, mimicking, or putting him- or herself in the shoes of the other, emphatically or otherwise. Undoubtedly, the processes that are observed are located in the individual; however, the phenomena are all examined in *co-presence* situations. Looking at the same research from a *communication perspective*, namely, as *jointly recruited processes*, gives these processes a different functionality, as compared to looking at these processes from an *individual-centered perspective*. The former perspective suggests that jointly recruited processes serve the function of establishing parity between producer and perceiver. An individual-focused analysis suggests that the perceiver is "emulating" the producer, but it is not necessarily clear what such emulation serves.[2]

In the following section I review research on synchronization, focusing first on neurophysiological evidence obtained with nonhuman primates and humans. Next, I turn to some developmental evidence obtained in infant studies and related comparative evidence with nonhuman primates. Finally, I provide an overview of the voluminous behavioral evidence, which has a long history in social psychology. In closing the next section I circumscribe what synchronization as a core knowledge system entails.

NEURAL FOUNDATIONS OF SYNCHRONIZATION

The origins of an intraspecies "synchronization" process were discovered in single neuron studies with macaque monkeys. These studies revealed that a particular class of visuomotor neurons (mirror neurons) discharge when the monkey *engages* in a particular *action* and when it *observes* another monkey engaging in the *same action* (e.g., Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti et al., 1996). These studies have given rise to a burgeoning field that can be seen as a major contribution toward uncovering the neural foundations of communication.[3] There are at least three distinctive reasons that this research is important to understand the neural foundations of communication. The first is that the logic of the design underlying this research field is typically a *social* (copresence) paradigm. It involves recording *and* comparing neural activities that are recruited jointly in

two agents, of whom one is engaged in executing an action and the other in observing the very same action. Despite its apparent simplicity, the design relies on the logic of communication by recording the type of neural activities recruited in the execution (producer) and observation (perceiver) of action. Second, "mirror neurons" constitute a *direct link* between an executor and perceiver in the absence of (1) any mediating symbols or—in fact—(2) apparent motive, drive, or intention to communicate. It is an automatic phenomenon that escapes conscious access. Finally, such jointly recruited and shared neural processes represent a means of achieving synchronization between or among individuals and thus providing an embodied grounding (Smith & Semin, 2004) for the essential parity requirement for communication, namely, the joint recruitment of neural processes by executor and perceiver (production and perception) by means of which producer and receiver are put on the same footing.

Evidence from Monkeys

A substantial body of neurophysiological evidence with human and nonhuman primates has shed light on the neural mechanisms underlying the representation of intentional action (e.g., Rizzollati & Craighero, 2004, for a recent review). This evidence suggests intentional action, and its perception has a shared neural notation. That is, intentional action is represented or "embodied" in the form of (nonsymbolic) cognition and is distributed across actor and observer. Gallese, Keysers, and Rizzolatti (2004) suggest that the function of this neural notation is to permit direct experiential access to the "mind" of the other (direct matching hypothesis). Observing the action can trigger the same neurophysiological mechanisms as executing the action, with the consequence being that the observer has implicit access to the mental contents and feelings of the actor.

The early studies showed that a particular class of visuomotor neurons (in the F5 area of the macaque monkey premotor cortex) discharge when the monkey engages in a particular action (e.g., grasping a peanut) *and* when it observes another monkey engaging in the same action (Rizzolatti et al., 1996). These mirror neurons do not respond to seeing a hand merely mimicking an action or for that matter observing the object alone (Gallese et al., 1996; Rizzolatti et al., 1996). Note that mirror neurons are not limited to the F5 area in the monkey brain but several interconnected areas (inferior parietal lobule area PF and STS) that are responsive to biological movement (Rizzolatti, Fogassi, & Gallese, 2001, p. 662).

Recent studies have shown that mirror neurons in the F5 area also become active even when the final part of the action (e.g., gasping the peanut) is hidden (Umilta et al., 2001). Critically, in the hidden condition, the monkey must "know" that an object (e.g., a peanut) is behind the occluder and must observe the experimenter's hand going behind the occluder without seeing the completion of the action. More recently, Keysers and his colleagues (2003) have reported that specific populations of neurons ("audiovisual mirror neurons") in the ventral pre-

motor cortex of the monkey discharge not only when a monkey performs a specific action but also when it sees or *hears* another monkey perform the same action. These neurons therefore represent actions independently of whether these actions are performed, heard, or seen. The adaptive nature of the mirror neuron system is illustrated in a recent study (Kohler et al., 2002) where monkeys were trained to rip paper. It was shown that once trained the mirror neurons involved in the execution of this action were recruited in response to the action sound only. These studies suggest not only that single neurons in the premotor cortex synchronize to the actions the other is executing but also that the action "goal" is represented and inferred across different modalities. The correlation between mirror neuron activation in partially observed or merely heard conditions corresponds largely to the pattern of neurons recruited in the full performance of the actions, which result in a sound (peanut cracking) or complete action (grasping food). It is argued that mirror neurons in the F5 area can *retain* abstract representations such as action goals and intentions. However, it is equally plausible to suggest that what is retained is the production of an event (e.g., with sound) as both seen and heard, rather than an abstract representation.

The important point is that single neurons are recruited for some actions irrespective of the modality (sound, vision) by means of which they are inferred. Moreover, it is not merely the action but the consequence of the action that is represented. Rizzolatti and his colleagues (e.g., Rizzolatti & Craighero, 2004) suggest that mirror neuron activity mediates meaning. It is argued that the brain capacity to link first- and third-hand experience is a "simulation" mechanism by the third party that is central to the understanding of action. The proposed mechanism is based on the *direct matching hypothesis*, which suggests that understanding an action results from mapping an observed action onto motor representations of that action. The mirror system translates visual information (action) into knowledge (e.g., Rizzolatti et al., 2001).

It is not entirely clear what knowledge means within the direct matching hypothesis, because an observer recruits only about 20% of the neurons that are recruited in the performance of an action (e.g., Adolphs & Spezio, in press; Gallese et al., 1996). Furthermore, a number of issues require closer inspection before we obtain a clearer picture of synchronization processes. One is that all the evidence on mirror neurons with adult monkeys must be the result of learned associations. Some support for the generalization of synchronization comes from studies involving learned auditory associations, where specific acquired sounds associated with the action goal were sufficient to activate action-related motor neurons (e.g., Kohler et al., 2002). The important question is the ontogeny of synchronization, and to my knowledge there is no evidence on this. However, some behavioral evidence with nonhuman primates and neonates, reviewed below, is suggestive of innateness.

The evidence to date suggests that actions involving biological movement that are goal directed (transitive), as well as the observation of the entire action or some aspect of it (e.g., goal attainment alone), activate a corresponding neural representation (i.e., namely executor and perceiver synchronize). The process takes place without any explicit communicative intent. The evidence also suggests that synchronization is a jointly recruited process that is highly adaptive to novel actions (e.g., ripping a paper) as well as novel features of actions or their consequences (e.g., sound of ripping paper or cracking peanut). What evidence is there for humans?

Evidence from Humans

Action–Bodily Movement

The findings regarding neural synchronization in the case of humans converge with those described in the earlier section, except that with humans synchronization processes have been revealed for a broader range of actions, significantly when observing and executing *intransitive* (e.g., biological movement with no apparent goal) and mimed actions as well as transitive actions (e.g., Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). Recent work (e.g., Cochin, Barthélémy, Roux, & Martineau, 1999) provides further support. For instance, Cochin and colleagues (1998) showed their participants video film consisting of 20-second sequences with still shots and moving shots that were either human movements or object movements. Their findings show the participation of the sensorimotor cortex during visual observation of human motion.

Research with functional magnetic resonance imaging (fMRI) suggests that the human premotor cortex (involved in voluntary movements) is active during the observation of actions made by another individual (Buccino et al., 2001). It seems that similar to monkeys, humans automatically generate a neural replica of the action they observe in their premotor cortex—that is, the regions activated correspond to those that are active during the execution of the action. Grèzes, Tucker, Armony, Ellis, and Passingham (2003), using event-related fMRI, also show neural synchronization when subjects execute movements or observe these movements (intraparietal and ventral limbs of the precentral sulcus). Convergent evidence using magnetoencephalography (MEG) indicates that the primary motor cortex is activated during both the execution of different manipulatory actions (e.g., using chopsticks) and their observation (e.g., Hari et al., 1998; Jarvelainen, Schürmann, & Hari, 2004; Nishitani & Hari, 2000).

The time course of synchronization is an important issue and there is some recent work on addressing this. Flanagan and Johansson (2003) used a block-stacking task and recorded the eye gaze and action–execution (hand coordination) for the executor and eye-gaze action–observation for the observer. Their results reveal that for observers eye gaze is predictive rather than reproductive or reactive of the executor's movements. An important contribution to this question is to be found in a study by Gangitano, Mottaghy, and Pascual-Leone (2001). They recorded the motor-evoked potentials of individuals at different moments while they were observ-

ing another person perform grasping movements and showed that motor cortical excitability followed the phases of the observed action.[4] As a more recent review by Fadiga, Craighero, and Olivier (2005) suggests, the "motor system simulates underthreshold the observed action in a strictly congruent fashion. The involved muscles are the same as those in the observed action and their activation is temporally strictly coupled with the dynamics of the observed action" (p. 213). What is important is that it is not merely observing an action but also listening to words that have to do with actions (verbs such as lick, kick, pick) that is revealed to be accompanied by motor facilitation (e.g., Tettamanti et al., 2005; Wilson, Saygun, Sereno, & Iacobini, 2004).

An interesting question in this context is, Why does observation of action and neural synchronization not lead to action in the observer? There are two answers. A preliminary one is provided by Baldissera, Cavallari, Craighero, and Fadiga (2001), who examined whether the observation of action involves not only cortical motor areas but also low-level motor structures that resonate the observed actions as if they were performed by the observer. They had participants observe hand opening and closure and measured H-reflex size (to examine spinal cord excitability) evoked in the flexor and extensor muscles. The findings reveal that observing participants' H-reflex for flexors increased for hand opening and decreased for hand closing. The reverse was noted for extensors. These results suggest that there is an inhibitory mechanism *in the spinal cord* that stops the execution of the observed hand action. While this may apply in their experimental setting when the execution of a simple movement is observed by explicit instruction, social psychological research on so-called mimicry reviewed below suggests that complete inhibition of behavioral synchronization in complex social interaction may not be the case. Thus, the conditions for inhibition and lack thereof present an interesting conceptual and empirical issue. The second answer to why action observation does not lead to action in the observer is perhaps obvious. What I have termed "synchronization" does not suggest complete equivalence or identity between producer and receiver but correspondence or "parity." For instance, in the research on mirror neurons with monkeys observing and performing an action, approximately 20% of the neurons in the F5 are overlapping (Gallese et al., 1996). Thus, it is important to qualify that synchronization does not produce *identity* between producer and perceiver but parity or correspondence. Obviously, if synchronization were to lead to equivalence between a producer's and perceiver's actions and emotions rather than an overlap, then there would be confusion between producer and perceiver (Adolphs & Spezio, in press).

Emotion

Emotions play a central regulatory role in social interaction, and there is substantial behavioral evidence on the synchronization of emotions, which is generally referred to as emotional contagion (Hatfield et al.,

1994). These authors suggest that unintentional mimicry of nonverbal and verbal behavior is responsible for the transfer of emotion in social interaction. There is also growing evidence with humans suggesting that the observation of perceptually opaque movements such as the facial expression of emotions induce the recruitment of the same neural activities that are involved in the execution of such expressions. A number of brain imaging studies examined experimentally induced disgust (e.g., Phillips et al., 1997, 1998) showing that the amplitude of the insular response was a function of the degree of expressed disgust. Indeed, impairment to the insula has been shown to impair the recognition of disgust (and other emotional expressions; Adolphs, Tranel, & Damasio, 2003). Wicker and colleagues (2003) address the question of whether the same sector of the insula is activated when experiencing disgust. In this fMRI study, participants were shown brief video clips of faces expressing disgust induced by inhaling odorants. The same participants were used for both the expression and the observation of disgust. Both conditions led to the activation of the same sites in the anterior insula (see also review by Singer & Frith, 2005).

Although single neuron recording experiments with humans show that the observation of pain and the experience of pain activate the same neurons (Hutchison, Davis, Lozano, Tasker, & Dostrovsky, 1999), generally speaking the regions of the brain important for imitation are not the regions important in emotion, but instead there is substantial overlap in the brain regions important in imitation and observation. By mimicking the observed action, however, individuals are in a better position to know by feeling themselves what another person is feeling. This reasoning implies a mechanism through which imitation produces *emotional* contagion. Evidence for this reasoning is provided by Carr, Iacobini, Dubeau, Mazziotta, and Lenzi (2003). Using fMRI, Carr and colleagues found that the brain regions important for action representation and imitation, such as the superior temporal sulcus, are connected to the insula and amygdala—regions in the limbic lobe that are involved in emotions.

Synchronization of perceptually opaque facial expressions (i.e., emotions) comes from a number of studies that investigate *electromyographic* (*EMG*) activity, when individuals were exposed to pictures of emotional expressions. For instance, Dimberg, Thunberg, and Elmehed (2000) presented happy, angry and neutral faces subliminally (30 msec, with backward masking) and revealed that participants reacted with distinct facial muscle reactions corresponding to the happy and angry stimulus faces and that this muscular synchronization occurs within a window of 500 msecs, automatically and without conscious access to the emotion expressed. Similar findings are reported by Sonnby-Borgstrom, Jonsson, and Svensson (2003), both at the subliminal level as well as supraliminally, moderated by differences in the degree of emotional empathy (see also Sonnby-Borgstrom, 2002). Walbott (1991) asked participants to classify photographs of people's expressions. While participants were

engaged in doing this task, they were surreptitiously videotaped. The results showed that participants synchronized to the expressions of the faces they were classifying—thus, while classifying a happy face they were more likely to smile. Indeed, accuracy of classification was positively correlated with the extent of synchronization between expression on the photograph and participants' expression.

BEHAVIORAL EVIDENCE FOR SYNCHRONIZATION

Evidence from Neonates

If synchronization is a jointly recruited process that provides an embodied grounding to communication, then the question arises as to whether or not this is an innate capacity. The neurophysiological evidence is obtained with adult subjects and it is difficult to disentangle whether synchronization is an innate process or an acquired one. It is self-evident that the diverse actions (e.g., cracking peanuts, ripping paper, and using chopsticks) used in human and nonhuman primate studies are acquired skills. If, however, synchronization is an innate social process then such jointly recruited processes should be evidenced early on in development. The question is therefore whether the evidence reviewed previously has an innate base or whether it is a capability that requires substantial environmental input before there is a manifest phenotypical performance. There is in fact behavioral evidence with humans and nonhuman primates that speaks to this issue and runs under the terminology of "imitation." The conceptual framework is thus defined in terms of a model performing some bodily movement and an observer copying it, whereby the copying is an instance of imitation if there is a clear causal relationship between movement observation and copying.[5] The issue is a so-called correspondence problem. That is, it is possible for observed actions to be matched with imitated actions and not entail the transfer of a novel skill by observation.

The research reported in this section uses the imitation terminology and therefore anchors the analytic focus at the individual level. Furthermore, it relies on the minimal co-presence social paradigm similar to those in the neurophysiological studies reported earlier. In this case, the dependent variables are behavioral and are concerned with whether an observer synchronizes (or reproduces) a feature of the body movement that is executed by a "model."

Research with primates suggests that chimpanzees can imitate (Preston & de Waal, 2002), whereas other studies suggest that extensive training (e.g., Custance et al., 1995) or experience and contact with humans (e.g., Whiten, 1998; Whiten, Custance, Gomez, Teixidor, & Bard, 1996) is necessary for imitation. For instance, Myowa (2004) examined a nursery-reared infant chimpanzee (Pan troglodytes) between 5 and 15 weeks of age demonstrating that the infant imitated tongue protrusion (during 5–10 weeks) and mouth opening (5–11

weeks). Both types of imitation ceased after 12 weeks. In another study with two chimpanzees reared from birth by their biological mothers, Myowa-Yamakoshi, Tomonaga, Tanaka, and Matsuzawaz (2004) observed discrimination between and imitation of tongue protrusion and mouth opening at less then 7 days of age. This differential neonatal imitative behavior ceased after 2 months of age and was replaced by indiscriminate mouth opening responses. What these results suggest is that the disappearance of neonatal imitation also marks the onset of social communicative behavior such as a decrease in neonatal spontaneous smiling and an increase in social smiling (e.g., Tomonaga et al., 2004). They also reveal evidence for a very early but limited and short-lived repertoire (tongue protrusion and mouth opening) of synchronization processes with primates.

Whether human infants have an innate ability to imitate specific facial gestures were stimulated by two experiments reported by Meltzoff and Moore (1977) suggesting that 2- to 3-week-old infants can imitate the specific facial gestures (e.g., tongue protrusion and mouth opening). The research was followed up by Meltzoff and colleagues (e.g., Meltzoff & Moore, 1983, 1992; see Meltzoff & Moore, 1997, for a review) as well as by other research groups (e.g., Heimann, Nelson, & Schaller, 1989; Legerstee, 1991). Although there are claims for neonate imitation across a range of gestures in the literature, a careful analysis of the data by Anisfield (1991, 1996) shows that there is only conclusive and reliable evidence for *tongue protrusion*—the most widely studied gesture (see also Heyes, 2001). The likelihood of neonates performing a tongue protrusion gesture increases significantly if they have observed the gesture being performed. Although there are parallels drawn between the research on mirror neurons and the research with neonates (e.g., Meltzoff & Decety, 2003), it is not clear to which extent the evidence with neonates warrants such sweeping interpretation. As Heyes (2001) points out:

> If tongue protrusion is the only body movement that newborns can imitate, it is plausible that the mediating process is an innate releasing mechanism; an inborn stimulus response link, wherein the response coincidentally resembles the stimulus from a third party perspective. (p. 253)

A less negative way of looking at the nature of the supportive and nonsupportive evidence is possible if one takes the current perspective that synchronization is constitutive for communication. Indeed, the finer analysis afforded by Anisfield can be seen as providing support for the contention that synchronization has an innate basis. The fact that the supportive evidence is only reliable in the case of movements of the mouth and face regions may not be surprising given that this part of the human body has the highest overall communicative significance. Thus, ironically, the critical assessment of the reliability of human neonate imitation may not be due to a coincidental resemblance, as Heyes (2001) has suggested, but may be regarded as constituting evidence for the innate

bases of synchronization processes that ground communication.

Evidence from Adults

The research literature on behavioral synchronization is extremely rich and varied in terms of both its history and the facets of behavior for which synchronization has been demonstrated. The types of behaviors that have been investigated range from paralinguistic features of speech like pauses or hesitations (see, e.g., Capella, 1981; Giles & Coupland, 1991, for different reviews) to arbitrary movements such as nose rubbing or foot shaking (Chartrand & Bargh, 1999) to experimental demonstrations of a particular syntactic structure that has been used by one person to be used again by the other dialogical partner (e.g., Bock, 1986, 1989; Bock & Loebell, 1990) to emotion contagion (Hatfield et al., 1994).

An early research pocket on synchronization examines a "thread of the communication tapestry" and has focused on specific measurable features of speech such as pauses, vocalization, and latency (see Capella, 1981, for a review). For instance, Jaffe and Feldstein (1970) assessed speech and silence at every 300-msec interval, thus recording vocalization and pauses across different dialogical settings (interviews, resolution of attitudinal differences, open-ended discussions over eight sessions). They reported substantial and consistent interclass correlations of average pauses between speakers across dialogical settings. In a set of conceptually convergent but methodologically different studies Matarazzo and Wiens (1967, 1972) experimentally manipulated interviewers' switch pauses. Their results showed that across the different switch pause conditions interviewees' pauses increased when the interviewers' pause increased and decreased when the interviewers' pause decreased. Capella and Planalp (1981) have reported convergent results for pauses and switch pause durations in informal 20-minute conversations by means of time series analyses.

A pattern similar to pauses-and-switch pause duration has been found for speech rate (Street, 1984; Webb, 1969, 1972; however, see Koomen, 1976) and vocal intensity. For instance, Natale (1975) found across 20 interviews that interviewees adapted their voice intensity to that of interviewers. In a second study, two-person dialogs lasting for 1 hour each and repeated over three occasions showed that the discrepancy in intensity between the speakers decreased linearly over time. These findings regarding voice intensity "convergence" appear to be stable across different studies (see Cappella, 1981). There are also numerous studies that have examined utterance length, number of words spoken, and other measures of verbal productivity across a number of dialogical settings and as Capella (1981) has summarized, there is no consistent pattern comparable to the ones obtained for pauses and vocal intensity, certainly not in informal settings. More recently, McGarva and Warner (2003) report coordination of vocal activity rhythms showing

speech accommodation. There is also a substantial amount of converging evidence from research on speech accommodation theory (see Giles & Coupland, 1991, for a review).

A substantial amount of psycholinguistic evidence on syntactical priming provides converging evidence on synchronization processes in dialogue or experimentally induced dialogical settings. The general gist of the findings is that if a particular syntactic structure has been used, the likelihood that the same grammatical structure will be used again increases significantly. The classic examples are a series of studies by Bock (1986, 1989; Bock & Loebell, 1990). Under the pretext of a memory test, Bock (1986) gave participants a prime sentence and subsequently a picture that they had to describe. Participants' task was to repeat the sentences (primes) they saw and then construct a sentence describing the picture they were presented next. The prime sentences varied in terms of their syntactic construction and the pictures were chosen such that they could be described by using one of two possible forms. For instance, in one condition participants were given an active sentence (e.g., "One of the fans punched the referee") and in the other a passive sentence (e.g., "The referee was punched by one of the fans"). Alternatively, a prime sentence was constructed with a propositional object form of an alternating dative verb in one condition (e.g., "A rock star sold some cocaine to an undercover agent") and in another condition the prime sentence used the double object form (e.g., "The rock star sold an undercover agent some cocaine"). Participants were much more likely to use the syntactic form of the prime when describing the picture they saw next, thus active or passive form, and so on, depending on the prime.

Furthermore, Bock was able to show that *prepositional object* sentences were much more likely to prime prepositional object sentences, although the prepositions differed, Thus, "The secretary took a cake to her boss" or "The secretary baked a cake for her boss" both elicited "The girl handed a paintbrush to the man." Similarly, sentences with prepositional phrases specifying locations (e.g., "The wealthy widow drove her Mercedes to the church") primed prepositional object descriptions even when the preposition did not specify location (e.g., "A rock star sold some cocaine to an undercover agent"). Bock and her colleagues argue that the production of a sentence activates *procedures* that are peculiar to the specific syntactic form. Thus, a passive sentence may be based on a specific procedure that is different from the production of a passive sentence. The procedure that is entailed in the production of a particular syntactic form thus persists and facilitates its subsequent use. Moreover, they argued against a view that suggests a phonetic or episodic trace of a sentence, as the diverse syntactic priming experiments show that syntactic priming occurs even when the prime and target sentences are very different (however, see Pickering & Brannigan, 1999).

Bernieri (1988) investigated *perceived movement synchrony* by having participant judges code the postures of 38 high school juniors and seniors in 19 interactive teach-

ing dyads presented on video recordings and compared them with pseudo-interaction control clips. Ratings of perceived movement synchrony were higher for real interactions (see also Bernieri et al., 1988; Bernieri, Davis, Rosenthal, & Knee, 1994, for synchrony measurement method and similar results). Moreover, there was a strong relationship between participants' rapport and the degree of perceived synchrony by raters—providing evidence for the interpersonal coordination hypothesis advanced by Tickle-Degnen and Rosenthal (1987). A recent study by Shockley, Santana, and Fowler (2003) showed interpersonal coordination of postural sway in the context of a cooperative verbal task.

Bavelas and her colleagues (e.g., Bavelas et al., 1988; Bavelas, Black, Lemery, & Mullett, 1986a, 1986b) report a number of studies on what they refer to as motor mimicry, defined as behavior by an observer appropriate to the situation of another (e.g., wincing at the other's injury or ducking when the other does). They argue that the function of such motor mimicking is primarily communicative (Bavelas et al., 1988b). They created a situation with a victim of an apparently painful injury and controlled the amount of eye contact that the victim had with a participant observing the victim. They were able to reveal that motor mimicry displayed by the observing participant was significantly shaped by the visual availability of the victim, thus suggesting that mimicry serves a communicative function.

Chartrand and Bargh's (1999) influential contribution initiated the currently active field on behavioral synchronization. They had a trained experimenter rub her nose or shake her foot while interacting with a participant. Their findings showed that when the experimenter rubbed her nose, participants were more likely to do so as well rather than shake their foot, and in the experimental condition when the experimenter shook her foot, the likelihood of participants shaking their foot was higher than rubbing their nose. This particular synchronization termed the "chameleon effect" (Chartrand & Bargh, 1999) refers to the tendency to adopt the postures, gestures, and mannerisms of interaction partners. The research evidence to date suggests that his type of mimicry occurs outside conscious awareness and without any intent to mimic or imitate. There are (in the interim) a number of studies that illustrate the diverse moderating conditions and consequences of the so-called action–perception link, namely, the assumed direct link between perceiving a behavior and performing that same behavior (Dijksterhuis & Bargh, 2001). Nonconscious synchronization has been shown to be moderated by self-monitoring (Snyder, 1974, 1987), with high self-monitors being more likely to mimic in situations involving affiliative cues than low self-monitors (Cheng & Chartrand, 2003). Similarly, inducing a self-construal that is more interdependent (or collective) either by experimentally priming or by recruiting participants differing on their chronic self-construal has been shown to give rise to stronger mimicry effects compared to persons who were chronically more individualistic or independent (van Baaren, Maddux, Chartrand, de Bouter, & van

Knippenberg, 2003; see van Baaren, Horgan, Chartrand, & Dijkmans, 2004, for a similar set of studies).

There is a substantial literature on emotion contagion (see Hatfield et al., 1994). Neumann and Strack (2000) report nonintentional mood contagion for participants who, under the pretext of a text comprehension study, listened to an affectively neutral speech spoken in a slightly sad or happy voice. They showed that the emotional tone induced a congruent mood state, leading the authors to conclude an emotion contagion (Hatfield et al., 1994) account of their results. This led them to interpret their findings in terms of a perception–behavior link (Chartrand & Bargh, 1999).

As can be seen from the foregoing compilation of studies from social psychology, the history of social behavioral research on what is termed here "synchronization processes" is extremely rich. Some of the earlier work is acknowledged *in passim* in the current social psychological literature on mimicry but is not fully recognized in terms of the conceptual foundations that these early researchers were trying to construct and establish. Indeed, the classic volume by Martha Davies (1982) was an important theoretical and methodological landmark, which somehow got buried by fashionable currency winning over cumulative knowledge. Indeed, one of the concerns central to this research was treating communicative behavior in terms of social interaction as the unit of analysis rather than the individual (e.g., Scheflen, 1982). The early research did deploy terminology referring to the jointly manifested and attuned feature of behaviors. Thus, one comes across a variety of terminological anchors such as "imitation," "mimicry," "empathy," "sympathy," "matching" (e.g., Capella, 1981), "entrainment" (Condon & Ogston, 1966), "interpersonal synchrony" (Bernieri et al., 1999; Condon & Ogston, 1966; Kendon, 1970), "periodicity" (Davies, 1982), and "contagion" (Hatfield et al., 1994).

CONCLUSIONS ON SYNCHRONIZATION AND ITS CONSEQUENCES

The research I have reviewed so far shows considerable consistency across disciplinary and thematic foci, ranging from neurophysiological to neuroscientific to developmental and social psychological traditions. First, there is an overall paradigmatic consistency. The experimental setups entail typically *co-presence* paradigms—this is a defining characteristic of the entire literature that has been reviewed so far. Obviously, the disciplinary visions and questions involve different but complementary definitions and manipulations of variables. In the neuroscientific tradition, behavior constitutes the tightly controlled independent variable and neural activity of the dependent variable. Oftentimes, the behavior is a very simple well-defined bodily movement (e.g., picking up a peanut) with the exception of perceptually opaque movements (e.g., facial expression of emotions). In the developmental work both independent and dependent variables are very limited and specific ranges of behaviors

with reliable evidence only for tongue protrusion. In the social psychological work, behavior is the dependent variable with changing controlled co-presence conditions. The general behavior in such settings is complex within which a specific experimentally controlled component (e.g., nose rubbing and foot shaking) is embedded. It is possible that in such experimental conditions other types of behaviors synchronize (e.g., speech rate, pauses, accent, bodily sway, and facial expressions) as revealed by a number of earlier studies that have investigated synchronization of motor behavior in natural interaction settings (Capella, 1981).

The second and more important consistency for the foundations of communication is simply that across all the reported research there is evidence of jointly and simultaneously recruited processes by means of which parity is established between producer and perceiver at neural, perceptual, affective, and behavioral levels. These processes are shown to be automatic, to escape conscious access, and to occur without the presence of explicit communicative intent. It is noteworthy that in the social psychological work the conceptual focus has been on phenotypical features of synchronization (e.g., affiliation and prosocial behavior) rather than the underlying communicative function of synchronization (discussed in the next section).

Obviously, the difference in research foci (e.g., neural vs. behavioral processes) brings with itself systematic differences in the wealth of observed and recorded data that are informative and can lead to cross-fertilization between the disciplinary fields, particularly as they throw a new light on communication and its embodied foundations. In the following section I proceed through these diverse implications and consequences that emerge from the work so far.

The Sameness Principle: "Genotypical" Basis of Communication

The social dimension of our lives plays a crucial role in the way we are shaped; the way we form images of our world; how we interact with and within it. Obviously, this is not only peculiar to primates or cetaceans but also to a number of species as different as graylag geese, termites, and bees. The nature and functions of social behavior across species differ, as do the mechanisms that drive them. However, what does not differ as a general principle for any social species is the necessity of mutual *recognition* and *intelligibility*. It is important for any social species to have evolved mechanisms that permit them to be coupled by being put on the same footing. Such mechanisms must facilitate obtaining a state of equivalence—what counts for one member must count for the other and such mechanisms, by giving privileged access to each other, ground mutual recognition and intelligibility.

Mutual recognition is by definition a collaborative process and it takes two to collaborate. In other words, "sameness recognition" cannot come about without interacting, or without exchanging information while navigating in a joint environment. While the particular ways in which mutual recognition and intelligibility are achieved vary between species, the general characteristics of the mechanism or process by which two parties communicate and by which mutual recognition and intelligibility are achieved remain the same.

For members of any social species to be able to communicate they must have mechanisms in place that permit them to be coupled by being put on the same footing. In other words, such mechanisms must facilitate obtaining a state of equivalence between two parties: What counts for one member also must count for the other and thereby also must ground mutual recognition and intelligibility. Imagine how short-lived the existence of vervet monkeys would be if vervets started to look down for snakes or look up for an eagle when they heard a leopard alarm. Communication can only succeed when two parties have a common understanding. This parity assumption (Liberman & Whalen, 2000) also constitutes the basis of mutual recognition and intelligibility. Thus, species sameness and its recognition is a biological necessity for any social species. *If there was no within-species parity, one would not be able to classify "same" and discriminate it from the "other." It is therefore significant for adaptive purposes not only to be able to recognize another as "same" or different but also to recognize that what another "same" does is intelligible.* I refer to this as the *sameness principle.* An important implication of this principle is that recognition of sameness goes hand in hand with the recognition of difference—an issue that will be further addressed at the conclusion of this chapter.

We have reviewed research that has to do with the neural and behavioral synchronization processes. This research reveals evidence supporting the notion of jointly recruited processes by two parties allowing for an embodied grounding of a coupling process. Notably, synchronization is not a process leading to equivalence between a producer and perceiver's neural activity, actions, and emotions. Such a state would lead to complete confusion between producer and perceiver. Rather, it is a *multimodal* process that provides a partial overlap.

One question that can be raised is whether the particular neural coupling process is species specific? Obviously, neural synchronization as a process is not specific to humans but has been observed in nonhuman primates. However, is it possible that observing the movements of the member of another species recruits corresponding neural structures? Indeed, if synchronization would be noted between species then it would be impossible to perceive sameness and difference. The research evidence available to date suggests that neural synchronization is not a process that is recruited across species and is not observed in the case of bodily movements of other species. In a fMRI study, Buccino and colleagues (2004) had participants observe mouth actions performed by humans and by individuals belonging to other species (monkey and dog). They were able to show that actions that belong to the motor repertoire of a human observer (e.g., biting and speech reading) were mapped on the observer's motor system. However, actions that do not belong to this repertoire (e.g., barking) are recognized on

the basis of their visual properties. Similarly, Martineau and Cochin (2004) have examined the relation between visual perception of motion and cortical activity by presenting children ranging from 29 months to 5 years with a still image (a lake) and animated images with human, animal, and virtual movement. Their study revealed that stimuli representing human, animal, and virtual movement activate different cortical areas suggesting that real human movement is processed neurally in a distinctive manner. It would therefore seem to be reasonable to assume that the architecture of the human perceptual and neural system is specifically designed for the recognition of species-specific movements in a privileged way, thereby *establishing a type of knowledge that has an entirely different ontological status than knowledge about the world in general*. This is the genotypical foundation of communication and the *embodied building block* of social cognition. Synchronization is therefore very much like a species-specific password that gives simultaneous privileged multimodal access to "each other," thereby establishing the basis for mutual recognition and the intelligibility.

"Phenotypical" Consequences of Synchronization and the Embodied Foundation of Communication

The social psychological research has long since converged on the positive consequences of behavioral synchronization along with the observation that these consequences are mediated automatically and unconsciously. The positive interpersonal consequences of unconscious behavioral synchronization postulated (e.g., Tickle-Degnen & Rosenthal, 1987) and reported early on along with supportive data (e.g., Bernieri, 1988) have revealed a strong relationship between participants' rapport and the degree of perceived synchrony. There is substantial evidence showing that the phenotypical manifestation of behavioral synchronization finds expression in experiences of greater liking for the other and reports of a more cohesive relationship, perceptions of shared identity, and smooth interaction *inter alia* (e.g., Bernieri, 1988; Hatfield et al., 1994; LaFrance, 1985; LaFrance & Broadbent, 1976; LaFrance & Ickes, 1981; Neuman & Strack, 2000). More recent work has examined the consequences of synchronization in terms of affiliation and prosocial behavior (Gump & Kulik, 1997; Lakin & Chartrand, 2003; Lakin, Jefferis, Cheng, & Chartrand, 2003; van Baaren, Holland, Kawakami, & van Knippenberg, 2004; van Baaren, Holland, Steenaert, & van Knippenberg, 2003). These consequences may also be what some refer to as the fundamental need to belong (e.g., Baumeister & Leary, 1995; Caporael, Dawes, Orbell, & van der Kragt, 1989; Leary & Baumeister, 2000).

It should be noted at this point that as much as synchronization is an automatic and nonconscious process the investigations of behavioral synchronization have focused predominantly on contexts that are not adversarial. Research within a speech accommodation theory (SAT; Giles & Coupland, 1991) framework has long ago demonstrated synchronization and desynchroni-

zation processes as a function of cooperative and competitive relationships. The behavioral focus in these studies was accent and changes in the interlocutor's accent or language (e.g., Giles, Bourhis, & Taylor, 1977; Giles & Smith, 1979; Giles, Taylor, & Bourhis, 1973). This research suggests that in multilingual contexts, a cooperative or competitive relationship between interlocutors influences accent and language use, originally taken as indicators of attitudes. Desynchronization (i.e., divergence in SAT) is found to be pronounced in intergroup contexts when the speaker expects competitive interactions with outgroup members (e.g., Doise, Sinclair, & Bourhis, 1976; Taylor & Royer, 1980) or when outgroup members are known to hold negative attitudes toward the transmitter's group (Bourhis, Giles, Leyens, & Tajfel, 1979). Bourhis and Giles (1977) provide a classic demonstration in which an adversarial conversational context is shown to induce participants to introduce a variety of strategies, such as accentuating differences in speech as well as nonverbal behavior to emphasize the difference between themselves and the adversarial "other" signaling divergence from the "other." In contrast to desynchronization (i.e., divergence), a number of communicative acts and styles are used to reduce differences, such as speech rate, pausal phenomena, utterance length, but also smiling, gaze, and so on. These are strategies by which individuals adapt to each other's communicative behaviors. Levin and Lin (1988) provide an interesting example of convergence. They showed that during the Watergate trials John Dean converged in terms of his median word frequency (an index of formality) to his different Senate interrogators. Similarly, Coupland (1984) conducted a phonological analysis of a travel agent and showed that she converged to her clients as a function of her clients' socioeconomic status and education. In short, situated goals influence whether synchronization or desynchronization is likely to be manifested.

There are two sets of implications of the foregoing findings:

1. The goals that are pursued shape whether *behavioral* synchrony or desynchrony will emerge. The point is that desynchronization is decoupling. Decoupling leads to the recognition of the other as different, and the phenotypically manifested form of it will be expressions of social distance. Thus, while synchronization leads to parity and the classification of the other as the same, desynchronization leads to disparity and discrimination of the other as different. It is therefore significant for adaptive purposes not only to be able to recognize another as "same" or different but also to recognize that what another "same" does is intelligible. Notably, this process is bidirectional. Thus, the process of behavioral synchronization or desynchronization can lead to positive or negative phenomenal expressions and experiences respectively as well as the reverse.

2. Phenomenal expressions and experiences rely on underlying synchronization processes and grounded by them. Thus, any experienced sense of "belongingness" or "motivation to engage in prosocial behavior" would

have to be based on and result from an exchange of information occurring at a modality different from speech—the behavioral evidence reviewed earlier repeatedly reveals the automatic and unconscious nature of the process. Moreover, there is certainly no access to neural synchronization. The evidence strongly suggests that synchronization is a process that serves a communicative function and is embodied. Thus, jointly recruited multimodal processes precede the phenotypically experienced social consequences of synchronization. Communication must have taken place for any other implication to be processed in the first place. This takes the form of jointly experienced bodily states—synchronization—and can be manifested in different phenotypical expressions—but the distinctive feature of the diverse phenotypical manifestations is a positive "oneness" of two individuals.

Synchronization: The Building Block of Communication

As the evidence reviewed so far shows, synchronization is a process fundamental for the embodied grounding of communication. It entails a specific architecture involving psychological and neural mechanisms that are designed for the joint processing of information about conspecifics and is a central *scaffold* to human communication, as we know it. The distinctive feature of this scaffolding architecture is that it is designed for the neural representation of knowledge that is jointly recruited in *coaction*, thus establishing correspondence between multimodal production and perception processes by automatically aligning processes of production and perception. Compared to other communication modalities, such as modern language, which are, in evolutionary terms, recent modes of communication and are specifically human, synchronization is prelinguistic and ancient, thus displaying some continuity with primates. Furthermore, synchronization is a heritable foundation albeit with considerable plasticity.[6] Moreover, as an architecture, it is specific to and constitutional for communication and by implication a uniquely defining feature of the domain of social cognition.

I would like to argue that synchronization is a specialized building block of communication very much akin to what Spelke (e.g., 2000, 2003; Hauser & Spelke, 2004) refers to as core knowledge systems. These knowledge systems rely on

> specialized perceptual systems for detecting particular kinds of sensory information and specialized motor systems guiding particular types of actions, infant animals have specialized, task-specific cognitive systems; systems for representing material objects, navigating through the spatial layout. . . . (Spelke, 2003, p. 278)

The type of knowledge systems she refers to are primarily nonsocial, such as numerosity and spatial orientation, and are grounded on a relational ontology in that features of these realities exist as a function of the interaction between the physical capabilities and properties of the observer and the physical properties of the environment (Gibson, 1977, 1979). Synchronization is, however, different, because it refers to a process underlying a knowledge system that is jointly recruited by a producer and perceiver.

The particular characteristics of core knowledge systems are that they are domain specific, task specific, and encapsulated and constitute an isolated system (Spelke, 2003). Synchronization shares these very same features, which are summarized by Hauser and Spelke (in press) as follows:

> a system of core knowledge is characterized by four properties. . . . First, it is domain specific: each system functions to represent particular kinds of entities such as conspecific agents, manipulable objects, places in the environmental layout, and numerosities. Second it is task-specific: each system uses its representation to address specific questions about the world, such as "who is this?" (face recognition), "what does this do?" (categorization of artifacts), "where am I" (spatial orientation), and "how many are there?" (numeration). Third, it is relatively encapsulated: each uses only a subset of the information delivered by the animal's input systems and send information only to a subset of the animal's output systems. Finally, the system is relatively automatic and impervious to explicitly held beliefs and goals.

Synchronization can therefore be regarded as constituting a separate core knowledge system because it shares the four general features but is distinctive from the types of knowledge systems investigated by Spelke and her colleagues. It has a uniquely designed architecture that joins two conspecifics and furnishes the possibility of socially shared representations by means of establishing neural parity between perceptual and production processes.

Much like core knowledge systems for navigating the nonsocial environment (or the physical features of the social environment), synchronization is a process that has undergone major transformations in human evolution—in particular with the acquisition of language as the powerful communication tool. The emergence of language as a tool to implement action gives rise to substantially sophisticated forms of communication—aside from dramatically extending our cognitive powers by introducing the uniquely human combinatorial capacity and thereby contributing to the creation of new systems of knowledge. While synchronization constitutes the fundamental building block for communication in general (and may also be responsible for the grounding of the uniquely human skill, namely, language[7]) the shape of communication and its uses underwent a dramatic transformation with the introduction of language. While *synchronization is a distinctly situation-specific process* regulating communication in terms of jointly recruited processes that establish parity between producer and perceiver, language is a means by which it becomes possible to detach from the immediacy of the situation both temporally and spatially. This facility enhances the attainment of increasingly complex individual and collective goals that are removed from the here and now by means of language- and convention-driven processes of coordination to which we turn in the next and final part of this chapter.

FROM SYNCHRONIZATION
TO COORDINATION

Communication as purposeful social interaction takes place in a social context and is regulated by social rules and conventions that are deployed to establish a shared reality and to attain individual or group goals (e.g., Austin, 1962; Grice, 1975; Higgins, 1981, 1992; Krauss & Fussell, 1996; Searle, 1969). When we move on to complex and purposeful communicative events such as a mundane dialogue or other joint activities, as in the case of recurrent social or task situations such as executing a heart operation, then the successful achievement of any of these activities relies on skillful *coordination* between two or more people. For instance, a heart surgery team consists of diverse experts (the surgeon, the anesthetist, etc.) who have expertise and have to coordinate unshared knowledge bases relevant to the performance of the operation. While each member of such a composite team of experts has a unique knowledge base they also have to have shared knowledge about how to coordinate their knowledge bases in order to be able to organize and execute the sequential joint activities properly. The coordinated communication between the different members of the team is intended to structure collective action such that a continuously monitored and publicly shared cognitive state is achieved (Hutchins, 1995). The socially distributed knowledge between team members in its public coordination, as well as physical tools such as heart–lung machine and electrocardiograph in which knowledge is literally embodied, become scaffolds for the successful performance of the operation. Instances ranging from dialogue to the performance of tasks that include external entities (e.g., other people or physical objects) to perform a task such as heart surgery require coordination and rely on communication by language to achieve such coordination.

Coordination in conversation or dialogue involves establishing parity at multiple levels of sequential processes. First of all, speakers and addressees rely on a number of conventions that have evolved to regulate the speaker–addressee relationship in conversation. Conversationalists concurrently access a set of "tacit" conventions or maxims. These maxims are derived from the unspoken *principle of cooperation*, which is important in conveying intended meaning in communication, and involve giving the right amount of information, making an informative, true, relevant and unambiguous contribution that is orderly and brief (Grice, 1975, 1978). The roles of speaker and addressee reverse in a turn-taking process regulated by conversational conventions signaling turns (Sacks, Schegloff, & Jefferson, 1974), opening and closing conversation, *inter alia* (Mehan & Wood, 1975).

For conversation to be successful interlocutors have to share information (i.e., mutual knowledge, beliefs, and assumptions). This shared information is referred to as common ground, and conversational partners continuously coordinate with each other to ground the content of their conversation (Clark & Brennan, 1991). Common ground refers to the set of knowledge and beliefs that are shared and assumed to be shared between the interlocutors. The process of grounding entails seeking and providing evidence of understanding in conversation. Accomplishing common ground is achieved by keeping track of the knowledge state of the interlocutor (e.g., Clark, 1996a; Clark, Schreuder, & Buttrick, 1983; Clark & Schober, 1992). To this end a number of—typically linguistic—strategies are employed to coordinate joint reference to objects and events in a communicative setting (e.g., Clark, 1992, 1996; Krauss & Fussell, 1996). Conversational coordination between speaker and addressee involves monitoring the *perspective of the addressee* (e.g., Fussell & Krauss, 1989a, 1989b; Schober, 1998), which contributes to the shape of the message production process. "Common ground" can be inferred from past conversations and supported by immediate surroundings and shared cultural background. Studies by Krauss and his colleagues (e.g., Fussell & Krauss, 1989a, 1989b) using a referential communication paradigm examine message design as a function of addressee characteristics. These types of studies illustrate how perspective taking influences the linguistic features of messages and how these in turn influence their communicative accuracy. It is not only the perspective of the addressee that is critical to successful communication. Considerable groundwork has to be done to achieve parity or intersubjectivity (Rommetweit, 1974; Schütz, 1962) in terms of the background knowledge shared by the interlocutors without which successful conversation cannot take place. In research terms this has meant a focus on processes that contribute to sharing a common ground (Clark & Wilkes-Gibbs, 1986), which has been regarded as a critical condition for successful communication.

In any purposeful complex task, it is not only the conventions noted for the coordination of a conversation that are sufficient but the additional coordination of joint activities that are given by the goal of the task. These have to do with who does what, when. Thus, what is required is not merely coordination at the dialogical level but also coordination of joint activities to achieve the goal of successful surgery. These are also driven by dialogue with a variety of correction practices to get back on course when events take unexpected turns.

The central theme driving this chapter is the necessity to establish equivalence or *parity* (cf. Liberman & Mattingly, 1985; Liberman & Whalen, 2000) between a sender and receiver for communication to be possible in the first place. So far, I have presented two sets of different processes by which parity is established. The first addressed in some length jointly recruited automatic processes underlying communication (i.e., *synchronization*) as an embodied grounding aspect of this process. The second, presented here briefly, is *coordination*, namely, the processes involved in purposeful human communication that establish equivalence in meaning and drive joint actions. Two final questions are addressed in this part. The first arises from the question of the differences between synchronization and coordination. The second has to do with how synchronization and coordination processes are interfaced. These are examined in the next two sections.

The Synchronization–Coordination Interface

How do synchronization and coordination interface? Obviously, human communication functions as an integral whole and not as a loosely connected set of processes. I turn to some features of dialogue and speech to illustrate this interface.

It takes the brain only a few seconds to put speech rate, accent, and message together for communication to occur. Speech is an incredibly fast process with an average production rate of 180–200 words per minute (approximately 333 msec per word); the upper range can go from fast (300 words per minute) to very fast (500 words per minute). We speak by accessing a lexicon with a volume between 20,000 and 60,000 (or more) words. Moreover, talk does not simply involve producing words. It requires choosing words from a lexicon to create sentences that are also linguistically structured. As Liberman and Whalen (2000) point out, an absolute requirement for particles in speech to be produced and perceived they have to be discrete, invariant and categorical—a consequence of the particulate principle (Abler, 1989, see below). A further constraint is inherent to the characteristics of the vocal tract and ear. Given the limitations of the vocal tract, the number of particles has to be small and put together in string form and these strings can "run considerable lengths. It is essential therefore, that the particles be produced and perceived rather rapidly if they are to be organized into the larger units of the language hierarchy. In fact, the consonants and vowels that are formed by the particles are delivered in speech at about 10 or 12 per second on average, and for short stretches the rate raises even higher" (Liberman & Whalen, 2000, p. 191). The point is that *perception* at these speeds is only possible because speech is *coarticulated*, thus production and perception are *synchronized* between producer and perceiver at the gesture level and not in the acoustic signal. Indeed, there is growing neuropsychological evidence providing provisional support for the notion of synchronization between producer and perceiver in speech. The evidence—however preliminary—relies on the notion that humans possess an "echo-neuron" system (cf. Rizzolatti & Craighero, 2004). An echo-neuron system refers to a neural system that synchronizes with the speech of another, in line with the motor theory of parity in speech perception (Liberman & Mattingly 1985; Liberman & Wahlen, 2000). For instance, Fadiga, Craighero, Buccino, and Rizzolatti (2002) presented participants with verbal and nonverbal stimuli (i.e., words, regular pseudowords, and bitonal sounds). In the middle of the words and pseudowords they had inserted either a double "f," which, when pronounced, requires slight tongue mobilization, or a double "r" requiring a tongue movement to be pronounced. When listening to words and pseudowords with "r" it was revealed that there was a significant increase of MEPs recorded from tongue muscles compared to the same stimuli with a double "f" as well as when listening to bitonal sounds. Watkins, Strafella, and Paus (2003) report similar results. Of course, as Rizzolatti and Craighero (2004) point out, these experiments—albeit less likely—do not exclude the alternative possibility that the echo-neuron system mediates the imitation of sounds. These findings lend support to the notion that an embodied grounding of speech and comprehension takes place via synchronization in a *coupled unit* of producer and perceiver and entails simultaneously and jointly recruited processes between producer and perceiver. This provides an anchor on the route to coordination, but it is still a long way away from purposeful, goal-driven communication via language use.

What is special about the evolutionary step introduced by symbolic vocalizations—speech—is that it furnished a novel possibility to human communication, namely, a combinatorial system. As Abler (1989) has pointed out, there is a distinct property of language, one that is also common to a number of natural systems. Abler draws attention to "Humbold's (1836) characterization of language, i.e. all systems 'make infinite use of finite media" whose "synthesis creates something that is not present *per se* in any of the associated constituents" (p. 1). Abler refers to this property as the *particulate principle*. "The creation of 'something that is not present per se in any of the associated constituents' involves the formation of at least a second level of organization in any system that conforms to Humbold's criteria . . . Human language involves phonemes at the first level, morphemes at the second level and phrase structures at the third level" (p. 3). This is the distinctive characteristic of human language. It is recursive. Importantly, as he points out, *organization at higher levels* (e.g., phrase structure) *tends to obscure lower-level constituents* (e.g., phonemes and phoneme composition).

Consequently, levels of organization have a propensity to act as shells, which enclose or hide their constituents. Translated into psychological terms, our attention and awareness during speech is driven by features regulating the dialogue, the communication goals that are pursued, and what is said and to be said, whereas "lower levels of organization" that act as scaffolds (phrase structure, morphemes, phonemes) for speech are used automatically and are not accessible and therefore outside consciousness.

The highest level of organization (at the dialogical level), "hide" the constituent levels, thus obscuring conscious access to them. These scaffolds entail particulars such as vocal gestures that make up speech (phrase structure, morphemes, phonemes) but also bodily movements, gestures, facial expressions, and their corresponding neural substrates, all of which are integrally matched to each other in communicative acts. All these modalities that are associated constituents, as suggested by the "particulate principle" (Abler, 1989), are recruited jointly and simultaneously by both producer and perceiver. It is not merely a matter of attention but the fact that the end product (synthesis) has a quality that is entirely different from the constituent parts. The direct analogy that illustrates this is, for instance, chemical compounds such as sand or water that are combinations of specific particulates (hydrogen, silicon, oxygen) that conceal their constituents. The constituents do not change character but retain their properties. However, the dif-

ferent syntheses produce compounds with distinct properties that are unique. Nevertheless, the constituent particles retain their qualities of distinctiveness and invariance and are categorical. Thus, it is inherent to the organization of communication that constituent levels of organization are not accessible. They are outside consciousness. The other reason is the one I mentioned earlier and has to do with the speed at which communication takes place. In contrast, the particular syntheses of these recursive constituents are a unique, nonreiterative, and consciously accessed "meaning."

Thus, while specific synchronization processes involved in speech facilitate the creative and unique syntheses manifested at the dialogical level, communication at the dialogical level (coordination) has also got to follow the parity principle. But now we are talking about parity at a consciously monitored level.

As I concluded in an earlier section, communication comes about in a *coupled unit* of producer and perceiver and entails simultaneously and jointly recruited processes between producer and perceiver. The processes are multimodal and occur by synchronization at the constituent levels of organization in communication, namely, recursive scaffolds of speech, but also in the case of the evolutionarily more ancient modalities of emotional and behavioral motor responses integral to communication.

Synchronization versus Coordination

Synchronization and coordination recruit distinctly different processes in terms of the *temporal frames* within which they operate. Synchronization fulfills an adaptive function by establishing neural, affective, and behavioral parity. With adaptive function I refer to the *continuous joint* behavioral and affective *monitoring* of a coupled unit of producer and perceiver that unfolds in continuous interaction. This is evident in the neural synchrony of motor movements and affect, as well as in the example of speech production and perception. If such monitoring would not be sensitive to immediate changes in the interaction then it would lose its adaptive function. Therefore, a significant feature of synchronization processes is that they operate in a *continuously situated* mode to monitor and establish parity.

An implication of the continuously situated operation mode is that synchronization processes have to be responsive to changes in very short temporal windows—otherwise they would lose their adaptive parity monitoring function. Indeed, one could argue that synchronization processes take place substantially below a 1,000-millisecond window,[8] which is one of the reasons (but not the only one) why synchronization processes are automatic and escape conscious access. A further distinctive feature of synchronization processes is that they are symmetrical in producer and perceiver.

Coordination in contrast is purposeful and goal driven and operates within much wider range of temporal windows. When cognitive activities are distributed across space and time and people, as they are with purposeful human communication, then language-driven coordination becomes an indispensable resource to structure and control actions. Any type of purposeful social interaction has a sequentially organized structure, which means that the coordination of the activities takes place over a temporal horizon. For a smooth and relatively flawless execution of joint action during purposeful social interaction it is crucial to ascertain that activities are coordinated and matched on at least two levels. One is within immediate communicative exchange but also across broader time intervals the range of which is defined by jointly pursued goals. Indeed, coordination runs over a substantial temporal horizon, as in the example of the heart surgery team or a variety of joint activities ranging from a dialogue to maintaining a relationship. The coordinating function of speech in dialogue is also very different. Compared to the symmetrical nature of synchronization processes coordination processes can take a wider range of dyadic formats, which have to do with the successful navigation of joint tasks as in the case of a dialogue or the example of the aforementioned heart surgery team. A further related feature of coordination processes is that they are purposeful and goal driven and consequently reflexive and consciously accessible. Coordinating activities frequently involves exchanging requests, instructions, corrections, and so on in order to maintain the smooth flow of joint activities, whereby such *correction* processes and deviations for coordinated action are consciously processed and instructions for corrections consciously produced. These processes are also situated processes; however, they operate within broader temporal windows.

CONCLUSIONS

Obviously, human communication, its nature, and its reach have been radically transformed over time. The crucial step for human communication was the evolution of gestural communication and speech as a result of two major changes in hominid vocal anatomy and physiology (e.g., Fitch, 2000; see note 7). The invention of writing in response for the need to record commercial transactions and administrative procedures in Mesopotamia approximately 6,000 years ago introduced a revolution in communication. This invention also meant that external devices could be used as aids for memory, cutting out intermediaries and distortions. The creation of different notations (alphabets, mathematics, music) and the introduction of simple devices, such as the space between words, radically changed the nature of the services that communication could be put into from advancing science to the reproduction and preservation of art to silent reading (Seanger, 1997). The invention of further technologies not only extended the communication reach but also transformed the nature of communication from the real and immediate (conversations, disputes, negotiations, etc.) to the removed (writing letters, voting, newspapers, films, theater, books), the magically removed (fantasies, dreams, daydreams), or the virtually removed (Internet) contexts. In short, the scope of human coordination is greatly expanded by the introduction of novel

tools and technologies. Thus the range of topics that can be covered under the social psychology of communication is vast. The focus of this chapter was on the biological, neural, psychological, and social foundations of the communication process with particular emphasis on synchronization processes as a core knowledge system, namely, the embodied grounding of communication that to our knowledge has not been addressed.

I tried to draw attention to the fact that human communication is not merely speech (e.g., Bavelas & Chovil, 2000). Language-based communication is about meanings and implications conveyed by articulatory gestures that are very important for coordination processes, but this is the tip of a *multimodal* communication iceberg that would not be possible without the cooperation of a range of automatic *neural and motor* processes that not only fulfill a continuous monitoring function by means of which parity is established but also provide scaffolds (among other things) for the production and comprehension of auditory messages. Face-to-face communication is composed of a simultaneously integrated set of audible and visible acts and any act of communication is a unified process resulting from the integration of information from multiple sensory modalities.

Thus, "language use"—located merely in terms of the product of speech—constitutes a crucial albeit very specific medium and modality for communication. While utterances in a dialogue may be the most salient features to the communication process, there are nevertheless a whole range of other modalities to communication aside from the production and perception of speech. These involve bodily movements (hand, arm, head, hip, leg, etc.) that accompany speech, the affective intonation of speech, and facial muscle activities involved in the expression of affect (*zygomatic major*, *corrugator supercilii* muscles, etc.), among other things. All these are finely integrated motor expressions along with communicative gestures and are not loosely or randomly occurring features of communication (e.g., Condon, 1982; Davies, 1982). This list is merely to draw attention to the fact that not only does human communication consist of *what is said* but that what is said (speech) is just a part of a multimodal expressive "whole." The communication modalities that are at the service of synchronization processes are prelinguistic and ancient in evolutionary terms. They display some continuity with primates. Other modalities such as modern language are—in evolutionary terms—recent modes of communication and are specifically human. In communication all modalities are integrated into a coordinated and complex whole.

There are a number of implications of the current formulation of synchronization processes as jointly activated. For one it treats communication as a "phenomenon" of a *coupled unit* and a phenomenon that comes about in terms of simultaneously recruited processes by the interlocutors to it, thus raising the analytic unit from that of an individual to a coupled unit. Conceptualizing communication as a process that has to be recruited simultaneously by two members involves thinking in terms of an analytic unit that is different from the traditional one. For historical and cultural reasons, the individual constitutes a prominent analytic focus and is the prevalent analytic unit in psychology and related disciplines. Such an analytic incision offers itself readily given the biological finitude of the individual. However, it does not necessarily lend itself as readily for conceptualizing what it means to be a "*social* species" in general and the analysis of communication in particular. Communication is a phenomenon that exists only when there is a *coupled unit* consisting of a producer and perceiver[9] and comes about in terms of processes recruited simultaneously by both producer and perceiver. Consequently, the proper analytic unit to communication is a coupled unit and jointly recruited processes.

Moreover, the current formulation suggests that the process of synchronization is a multimodal one serving multiple monitoring functions simultaneously. The present perspective introduces a different balance to the prevalent one that places a heavy emphasis on linguistically or symbolically driven communication. Language-driven communication—defined here as coordination processes—is the tip of the iceberg. For such purposeful communication to be successful hard work has to be done at the prereflexive level. Moreover, the argument derived from the particulate principle and its implication that higher levels of organization of communication (at the dialogical level) "hide" the constituent levels, thus obscuring conscious access to them, sheds a novel light on understanding a topic that is current, namely, the specific features of implicit cognitive processes. The distinctive flavor of approaching of what is implicit is that it comes from a perspective that is grounded on communication.

ACKNOWLEDGMENTS

NOTES

1. Others have suggested that chemosignals also have a significant role to play and that dance alone is insufficient to give other bees guidance (e.g., Sanoz, Laloi, Odoux, & Pham-Delegue, 2000; Wenner, Wells, & Johnson, 1969), although recent evidence suggests that dances provide the guidance to the vicinity of the source and chemosignals facilitate locating the precise source (Riley, Greggers, Smith, Reynolds, & Menzel, 2005).
2. The exception is learning by imitation.
3. There is, to my knowledge, only a brief reference to communication (Rizzollatti & Craighero, 2004, p. 183), while the theoretically provocative and stimulating accounts have been about the evolution of language (e.g., Arbib, 2005; Arbib & Rizzolatti, 1997; Rizzolatti & Arbib, 1998). Further theoretical speculations have been about what has been referred to as *shared intersubjective space* (e.g., Gallese, 2003).

4. There is a potential methodological issue with these two reports. This has to do with whether the observing subject's behavior is causally related to the movements of the executioner or the observed, or even to the anticipated object movements (as in the case of stacking blocks). The question is, Are the observers synchronizing with arm–finger movements or fixating on the endpoint of the object trajectory (Mataric & Pomplun, 1998)?

5. Thus, imitation is seen as different from *emulation* (cf. Mataric & Pomplun, 1998) where the observer's behavior may be casually related but not to the movement of the model but rather the observed object movement as in the case of stacking blocks. Emulation involves learning properties of the physical situation (Tomasello, 1990, 1998) rather than the movements themselves.

6. The plasticity of synchronization is demonstrated in the diverse studies with, for instance, monkeys whereby hearing specific sounds (e.g., peanut cracking and paper ripping) are sufficient to recruit the entire motor neuron sample involved in the action leading to the sound. See also the next section on the evolution of language. The flexibility of learning (as all the evidence with adults and monkeys suggests) synchronization is a very highly adaptive assimilative process, and acquisition of new forms is rapid. It is also noteworthy that synchronization as a process is also adaptively utilized in highly complex social interaction (e.g., syntactic priming), which indicates that the process has been adapted for the purposes of learning and executing more recent forms of communication.

7. I shall not discuss the issue of the evolution of language here for which there are diverse accounts (e.g., Aiello & Dunbar, 1993; Arbib, 2002, 2005; Corballis, 2004; Dunbar, 1996; Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004, *inter alia*). There is a neurophysiological account deriving from the research on mirror neurons, arguing for the gestural origins of language evolution (e.g., Arbib, 2002; Rizzolatti & Arbib, 1998) as well as models for the embodied grounding of language, which converge with the present argument on synchronization (e.g., Glenberg & Kaschak, 2002, 2003).

8. Most priming effects are successful well below the 1,000-millisecond margin. For instance, affective priming occurs within a short-lived window—preceding the target by 300 msec and succeeding it by maximally 150 msec (Fockenberg, Koole, & Semin, 2005). Synchronization of facial muscles with affective expressions takes place within 500 msecs (Dimberg et al., 2000, inter alia).

9. The nature of coupling has obviously undergone radical changes with the invention of revolutionary technologies such as writing, postal service, phone, video linking, Internet, and so on. These retain the embodied features of communication, but this subject is beyond the scope of this chapter.

REFERENCES

Abler W. L. (1989). On the particulate principle of self-diversifying systems. *Journal of Social and Biological Structures*, *12*, 1–13.

Adolphs, R., & Spezio, M. (in press). The neural basis of affective and social behavior. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed.). New York: Cambridge University Press.

Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, *52*, 61–69.

Aiello, L. C., & Dunbar, R. (1993). Neocortex size, group-size, and the evolution of language. *Current Anthropology*, *34*, 184–193.

Anisfield, M. (1991). Neonatal imitation: Fact or artifact? *Developmental Psychology*, *17*, 655–660.

Anisfield, M. (1996). Only tongue protrusion modeling is matched by neonates. *Developmental Review*, *16*, 149–161.

Arbib, M. A. (2002). Beyond the mirror system: Imitation and the evolution of language. In C. Nerhaniv & K. Dautenham (Eds.), *Imitation in animals and artifacts* (pp. 229–280). Cambridge, MA: MIT Press.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Brain and Behavioral Sciences, 28*, 165–167.

Arbib, M. A., & Rizzolatti, G. (1997) Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, *29*, 393–424.

Austin, J. L. (1962). *How to do things with words*. Oxford, UK: Oxford University Press.

Baldissera, F., Cavallari, P., Craighero, L., & Fadiga, L (2001). Modulation of spinal excitability during observation of hand actions in humans, *European Journal of Neuroscience*, *13*, 190–194.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong—Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R., & Mullett, L. (1988). Form and function in motor mimicry: Topographic evidence that the primary function is communicative. *Human Communication Research*, *14*, 275–299.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, L. (1986a). Experimental methods for studying elementary motor mimicry. *Journal of Nonverbal Behavior*, *10*, 102–119

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, L (1986b). I show how you feel motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, *50*, 322–329.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning - An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.

Beckers, R., Holland, O., & Deneubourg, J. (1994). From local actions to global tasks: Stigmergy and collective robotics. In R. Brooks & P. Meas (Eds.), *Artifical life* (Vol. 4, pp. 48–82). Cambridge, Ma: MIT Press.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher student interactions. *Journal of Nonverbal Behavior*, *12*, 120–138.

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin*, *20*, 303–311.

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother–infant interactions. *Journal of Personality and Social Psychology*, *54*, 243–253.

Bock, L. J. (1986). Syntactic priming in language production. *Cognitive Psychology*, *18*, 355–387.

Bock, L. J. (1989). Close class immanence in sentence production, *Cognition*, *31*, 163–189.

Bock, L. J., & Loebell, H. (1990). Framing sentences, *Cognition*, *35*, 1–39.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 119–135). London: Academic Press.

Bourhis, R. Y., Giles, H., Leyens, J. P., & Tajfel, H. (1979). Psycholinguistic distinctiveness: Language divergence in Belgium. In H. Giles & R. St Clair (Eds.), *Language and social psychology* (pp. 158–185). Oxford, UK: Blackwell.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Buccino, G., Lui, F., Canessa, L., Patteri, I., Lagravinese, G., Benuzzi, C., et al. (2004). Neural Circuits Involved in the Recognition of Actions Performed by Nonconspecifics: An fMRI Study. *Journal of Cognitive Neuroscience*, *16*, 114–127.

Capella, J. N. (1981). Mutual influence in expressive behavior: Adult-adult and infant–adult dyadic interaction. *Psychological Bulletin*, *89*, 101–132.

Capella, J. N., & Planalp, S. (1981). Talk and silence sequences in informal conversations III: Interspeaker influence. *Human Communication Research*, *7*, 117–132.

Caporael, L. R., Dawes, R. M., Orbell, J. M., & van der Kragt, A. J. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, *12*, 683–739.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, USA*, *100*, 5497–5502.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Cheng, C. M., & Chartrand, T. L. (2003). Self-monitoring without awareness: Using mimicry as a nonconscious affiliation strategy. *Journal of Personality and Social Psychology 85*, 1170–1179.

Clark, H. H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.

Clark, H. H. (1996a). Communities, commonalities, and communication. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 324–355). Cambridge, UK: Cambridge University Press.

Clark, H. H. (1996b). *Using language*. New York: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. Resnick, J. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 73–86). Washington DC: American Psychological Association.

Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, *22*, 245–258.

Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tamur (Ed.), *Questions about questions*. (pp. 15–48). New York: Russell Sage.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Cochin, S., Barthélémy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults. Electroencephalography. *Clinical Neurophysiology, 107*, 287–295.

Cochin, S., Barthélémy, B., Roux, S., & Martineau, J. (1999). Observation and execution of movement: Similarities demonstrated by quantified electroencephalography. *European Journal of Neuroscience*, *11*, 1839–1842.

Condon, W. S. (1982). Cultural microrhythms. In M. Davies (Ed.), *Interaction rhythms* (pp. 53–76). New York: Human Sciences Press.

Condon W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease*, *143*, 338.

Corballis, M. C. (2004). The origins of modernity: Was autonomous speech the critical factor? *Psychological Review*, *111*, 543–552.

Coupland, N. (1984). Accommodation at work–Some phonological data and their implications. *International Journal of the Sociology of Language*, *46*, 49–70.

Custance, D. M., Whiten, A., & Bard, K. A. (1995). Can young chimpanzees imitate arbitrary actions? Heyes and Heyes revisited, *Behavior*, *132*, 837–859.

Davies, M. (1982). *Interaction rhythms: Periodicity in communicative behavior*. New York: Human Sciences Press.

de Saussure, F. (1959). *Course in general linguistics*. New York: Philosophical Library. (Original work published 1915)

Dijksterhuis, A., & Bargh J. A. (2001). The perception–behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, *33*, 1–40.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*, 86–89.

DiPellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *NeuroImage*, *15*, 265–272.

Doise, W., Sinclair, A., & Bourhis, R. Y. (1976). Evaluation of accent convergence and divergence in cooperative and competitive intergroup situations. *British Journal of Social and Clinical Psychology*, *15*, 247–252.

Dunbar, R. (1996). *Grooming, gossip and the evolution of language*. Cambridge, MA: Harvard University Press.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402

Fadiga, L., & Craighero, L., & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Current Opinion in Neurobiology*, *15*, 213–218.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation—A magnetic stimulation study. *Journal of Neurophysiology 73*, 2608–2611.

Fitch, W. T. (2000) The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*, 258–267

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation, *Nature*, *424*, 769–771.

Fockenberg, D. A., Koole, S. L., & Semin, G. R. (2005). *Backward affective priming: Even when the prime is late, people still evaluate*. Manuscript under review.

Fussell, S. R., & Krauss, R. M. (1989a). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, *19*, 509–525.

Fussell, S. R., & Krauss, R. M. (1989b). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, *19*, 509–526.

Gallese, V. (2003). The manifold nature of interpersonal relations: the quest for a common mechanism. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences*, *358*, 517–528.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396–403.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation. *NeuroReport*, *12*, 1489–1492.

Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 57–83). Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Giles, H., Bourhis, R. Y., & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307–348). London: Academic Press.

Giles, H., & Coupland, J. (1991). *Language: Contexts and consequences*. Pacific Grove, CA: Brooks/Cole.

Giles, H., & Smith, P. M. (1979). Accommodation theory: Optimal levels of convergence. In H. Giles & R. St Clair (Eds.), *Language and social psychology* (pp. 45–65). Oxford, UK: Blackwell.

Giles, H., Taylor, D. M., & Bourhis, R. Y. (1973). Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in Society*, *2*, 177–192.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action, *Psychonomic Bulletin and Review*, *9*, 558–565.

Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 93–126). New York: Academic Press.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurons in the human brain: An fMRI study, *NeuroImage*, *18*, 928–937.

Grice, H. P. (1975). The logic of conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). New York: Academic Press.

Grice, H. P. (1978). Further notes on logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 9. Pragmatics* (pp. 113–128). New York: Academic Press.

Gump, B. B., & Kulik, A. (1997). Stress, affiliation, and emotional contagion. *Journal of Personality and Social Psychology*, *72*, 305–319.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences, USA*, *95*, 15061–15065.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotion contagion*. Cambridge, UK: Cambridge University Press.

Hauser, M. D., & Spelke, E. S. (in press). Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In M. Gazzaniga (Ed.), *The cognitive neurosciences III*. Cambridge, MA: MIT Press.

Heimann, M., Nelson, K. E., & Schaller, J. (1989). Imitative reactions among 14–21 day old infants. *Infant Mental Health Journal*, *6*, 31–39.

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, *5*, 253–261).

Higgins, E. T. (1981). The "communication game": Implications for social cognition and persuasion. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (pp. 343–392). Hillsdale, NJ: Erlbaum.

Higgins, E. T. (1992). Achieving "shared reality" in the communication game: A social action that creates meaning. *Journal of Language and Social Psychology*, *11*, 107–131.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchison, W. D., Davis, K. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (1999). Pain-related neurons in the human cingulate cortex. *Nature Neuroscience*, *2*, 403–405.

Iacobini, M. (2005). Understanding others: Imitation, language and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From mirror neurons to memes: Vol. 1. Mechanisms of imitation and imitation in animals* (pp. 77–100). Cambridge, MA: MIT Press.

Jaffe, J., & Feldstein, S. (1970). *Rhytmns of dialogue*. New York Academic Press.

Jarvelainen, J., Schürmann, M., & Hari, R. (2004). Activation of the human primary motor cortex during observation of tool use. *NeuroImage*, *23*, 187–192.

Kendon, A. (1970). Movement coordination in social interaction: Some examples described. *Acta Psychologica*, *32*, 101–125.

Keysar, B. (1997). Unconfounding common ground. *Discourse Processes*, *24*, 253–270.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, *153*, 628–636.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, *297*, 846–848.

Koomen, W. (1976). Of synchrony in conversation. *Gedrag: Tijdschrift voor Psychologie*, *4*, 219–229.

Krauss, R. M., & Chiu, C. (1998). Language and social behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 41–88). New York: McGraw-Hill.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 655–701). New York: Guilford Press.

LaFrance, M. (1985). Postural mirroring and intergroup relations. *Personality and Social Psychology Bulletin*, *11*, 207–217.

LaFrance, M., & Broadbent, M. (1976). Group rapport: Posture sharing as a nonverbal indicator. *Group and Organization Studies*, *1*, 328–333.

LaFrance, M., & Ickes, W. (1981). Posture mirroring and interactional involvement. Sex and sex typing effects. *Journal of Nonverbal Behavior*, *5*, 139–154.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*, 334–339.

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, *27*, 145–162.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology*, *32*, 1–62.

Legerstee, M. (1991). The role of person and object in eliciting early imitation. *Journal of Experimental Child Psychology*, *51*, 423–433.

Levin, H., & Lin, T. (1988). An accommodating witness. *Language and Communications*, *8*, 195–198.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*, 187–196.

Martineau, J., & Cochin, S. (2004). Visual perception in children: Human, animal and virtual movement activates different cortical areas. *International Journal of Psychophysiology*, *51*, 37–44.

Matarazzo, J. D., & Wiens, A. N. (1967). Interviewer influences on the durations of interviewee silence. *Journal of Experimental research in Personality*, *2*, 56–69.

Matarazzo, J. D., & Wiens, A. N. (1972). *The interview: Research on its anatomy and structure*. Chicago: Aldine-Alderton.

Mataric, M. J., & Pomplun, M. (1998). Foxation behavior in observation and imitation of human movement. *Cognitive Brain Research*, *7*, 191–202.

McGarva, A. R., & Warner, R. M. (2003). Attraction and social coordination: Mutual entrainment of vocal activity rhythms. *Journal of Psycholinguistic Research*, *32*, 335–354.

Mehan, H., & Wood, R. (1975). *The reality of ethnomethodology*. New York: Wiley.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London, B*, *358*, 491–500.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, *54*, 702–709.

Meltzoff, A. N., & Moore, M. K. (1992). Early imitation within a functional framework: The importance of person identity, movement, and development. *Infant Behavior and Development*, *15*, 479–505.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, *6*, 179–192.

Myowa, M. (1996). Imitation of facial gestures by an infant chimpanzee. *Primates*, *37*, 207–213.

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka M., & Matsuzawz, T. (2004). Imitation in neonatal chimpanzees (Pan troglodytes). *Developmental Science*, *7*, 437–442.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, *32*, 790–804.

Neumann, R., & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, *79*, 211–223.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*, 913–918.

Pasteels, M., & Bordereau, C. (1998). Releaser pheromones in termites. In R. K. vander Meer, M. D. Breed, M. L. Winston, & K. E. Espelie (Eds.), *Pheromone communication in social insects* (p. 193). Boulder, CO: Westview Press.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*, 495–497.

Pickering, M. J., & & Brannigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, *3*, 136–141.

Preston, S. D., & de Waal, F. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 1–72.

Ramnani, N., & Miali, R. C. (2003). A system in the human brain for predicting the action of others. *Nature Neuroscience*, *7*, 85–90.

Riley, U., Greggers, A., Smith, D., Reynolds, D. R., & Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature*, *435*, 205–207.

Rizzolatti, G., & Arbib, M. A. (1998) Language within our grasp. *Trends in Neurosciences*, *21*, 188–194.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rizzolatti, G., Fadiga L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action, *Nature Reviews Neuroscience*, *2*, 661–670.

Rommetweit, R. (1974). *On message structure: A framework for the study of language and communication*. New York: Wiley

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the analysis of turn taking in conversation. *Language*, *50*, 696–735.

Sandoz, J. C., Laloi, D., Odoux, J. F., & Pham-Delegue, M. H. (2000). Olfactory information transfer in the honeybee: Compared efficiency of classical conditioning and early exposure. *Animal Behavior*, *59*, 1025–1034.

Scheflen, A. E. (1982). Comments on the significance of interaction rhythms. In M. Davies (Ed.), *Interaction rhythms: Periodicity in communicative behavior* (pp. 13–23). New York: Human Sciences Press.

Schütz, A. (1962). *Collected papers*. The Hague, Netherlands: Nijhof.

Schober, M. F. (1998). Different kinds of conversational perspective-taking. In R. J. Kreuz & S. R. Fussell (Eds.), *Social and cognitive approaches to interpersonal communication* (pp. 145–174). Mahwah, NJ: Erlbaum.

Seanger, P. (1997). *Space between words: The origins of silent reading*. Palo Alto, CA: Stanford University Press.

Searle, J. R. (1969). *Speech acts*. Cambridge, UK: Cambridge University Press

Semin, G. R. (2000). Agenda 2000: Communication: Language as an implementational device for cognition. *European Journal of Social Psychology*, *30*, 595–612.

Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173.

Seyfarth, R., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic Communication. *Science*, *210*, 801–803.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation, *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 326–332.

Singer, T., & Frith, C. (2005). The painful side of empathy. *Nature Neuroscience*, *8*, 845–846.

Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, *30*, 526–537.

Snyder, M. (1987). *Public appearances, private realities: The psychology of self-monitoring*. New York: Freeman.

Sonnby-Borgstrom, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology*, *43*, 433–443.

Sonnby-Borgstrom, M., Jonsson, P., & Svensson, O. (2003). Emotional empathy as related to mimicry reactions at different levels of information processing. *Journal of Nonverbal Behavior*, *27*, 3–23.

Spelke, E. (2000). Core knowledge. *American Psychologist*, *55*, 1233–1243.

Spelke, E. (2003). What makes us smart?: Core knowledge and natural language. In D. Gertner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 277–312). Cambridge, MA: MIT Press.

Street, R. L. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research*, *11*, 139–169.

Taylor, D. M., & Royer, E. (1980). Group processes affecting anticipated language choice in intergroup relations. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 185–192). Oxford, UK: Pergamon Press.

Tettamanti, M., Buccino, B., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*, 273–281.

Tickle-Degnen, L., & Rosenthal, R. (1987). Group rapport and nonverbal behavior. *Review of Personality and Social Psychology*, *9*, 113–136.

Tinbergen, N. (1952). The curious behavior of the stickleback. *Scientific American*, *187*(6), 22–26.

Tomasello, M. (1990). Cultural transmission in the tool use and communicatory signaling of chimpanzees? In S. T. Parker & K. R. Gibson (Eds.), *Language and intelligence in monkeys and apes* (pp. 274–311). Cambridge, UK: Cambridge University Press.

Tomasello, M. (1998). Emulation learning and cultural learning. *Behavior and Brain Sciences*, *21*, 703–704.

Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., Kosugi, D., Mizuno, Y., et al. (2004). Development of social cognition in infant chimpanzees (Pan trogiodytes): Face recognition, smiling, gaze, and the lack of triadic interactions. *Japanese Psychological Research*, *46*, 227–235.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing. A neurophysiological study. *Neuron*, *31*(1), 155–165.

van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*, *15*, 71–74.

van Baaren, R. B., Holland, R.W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, *39*, 393–398.

van Baaren, R. B., Horgan, T. G., Chartrand, T. L., & Dijkmans, M. (2004). The forest, the trees, and the chameleon: Context dependence and mimicry. *Journal of Personality and Social Psychology*, *86*, 453–459.

van Baaren, R. B., Maddux, W. W., Chartrand, T. L., de Bouter, C., & van Knippenberg, A. (2003). It takes two to mimic: Behavioral consequences of self-construals. *Journal of Personality and Social Psychology*, *84*, 1093–1102.

Wallbott, H. G. (1991). The robustness of communication of emotion via facial expression: Emotion recognition from photographs with deteriorated pictorial quality. *European Journal of Social Psychology*, *21*, 89–98.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*, 989–994.

Webb, J. T. (1969). Subjective speech rates as a function of interviewer behavior. *Language and Speech*, *12*, 54–67.

Webb, J. T. (1972). Interview synchrony: An investigation of two speech rate measures. In A. W. Seligman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 115–133). New York: Pergamon Press.

Wenner, A. M., Wells, P. H., & Johnson, D. L. (1969). Honey bee recruitment to food sources: Olfaction or language? *Science*, *164*, 84–86.

Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *112*, 270–281.

Whiten, A., Custance, D. M., Gomez, J. C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children and chimpanzees (Pan troglodytes) *Journal of Comparative Psychology*, *110*, 3–14.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti G. (2003). Both of us disgusted in my Insula: The common neural basis of seeing and feeling disgust, *Neuron*, *40*, 655–664.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*, 701–702.

Attachment Theory and Research

Core Concepts, Basic Principles, Conceptual Bridges

PHILLIP R. SHAVER
MARIO MIKULINCER

A person coming to the literature of social psychology from ordinary life, in any society, or from the humanities—an expert on modern fiction, for example—would expect to hear a great deal about life in families, beginning with the total dependency of young children on their adult caregivers; the construction and maintenance across the lifespan of a self or identity in the context of relationships with family members and friends; sexual attraction, sexual behavior, mating, marriage, and parenting; and people's reliance on each other for protection, emotional support, and comfort in the face of life's inevitable disappointments, stresses, illnesses, conflicts, and losses. Although these topics do receive attention in contemporary social-psychological journals and textbooks, until fairly recently they were considered tangential to a field focused on social perception, attitudes, attitude change, and social behavior in task-oriented groups of various kinds.

There are multiple ways one could think about reasons for this intellectual history, but certainly one of them is that modern social psychology came of age during and after World War II, when propaganda and persuasion, group dynamics, and interethnic prejudices and conflicts were all very salient. Moreover, during those years the field was primarily a masculine enterprise, many of whose leaders had served in the armed forces and/or been captivated intellectually and emotionally by its ghastly examples of hatred, prejudice, and human cruelty. Less salient at that time were the many aspects of life that seemed more feminine and, at least to social psychologists, to fall within the disciplinary provinces of developmental, clinical, or personality psychology.

What then was hidden in shadows has steadily moved into the limelight as an increasing number of women have entered the field, family relationships (including abusive ones) and divorce have emerged as societal concerns, health and social contributions to health have become increasingly important, and research methods pioneered in more traditional social-psychological topic areas have proven useful in tackling subject matter that once seemed to defy empirical analysis (e.g., intimacy, trust, love, and grief). Also relevant to this chapter is the fact that social psychology has increasingly taken notice of evolutionary biology (e.g., Simpson & Kenrick, 1997). Earlier in its history (when social relations departments and interdisciplinary programs in social psychology were common), the field seemed more naturally linked with sociology and cultural anthropology, fields in which people's motives and attitudes were viewed as emerging from social and cultural roots without much basis in pan-human biological substrates.

Within the contemporary field of social psychology, attachment theory (Bowlby, 1969/1982) has proven to be a congenial and productive meeting place for researchers interested in close relationships, the development of self and personality in the context of such relationships, and the evolutionary background of core human needs and motives that find expression in relationships. From the beginning, attachment theory was interdisciplinary, crossing boundaries and borrowing concepts from ethology; evolutionary biology; cognitive, developmental, and clinical psychology; and psychoanalysis. The theory's creator, John Bowlby, was a British psychiatrist and psychoanalyst interested in normal and abnormal personality

development and its implications for social problems such as crime and delinquency, in addition to such clinical phenomena as anxiety disorders, disordered grieving following divorce or the death of a loved one, and depression. Bowlby's theory is now one of the leading intellectual frameworks in developmental, personality, and social psychology, partly because of methodological contributions made by Mary Ainsworth. Based on her empirical tests of the theory in observational and laboratory studies of infant–parent relationships, we and other social psychologists have been able to extend the theory to topics of interest to contemporary social psychologists: social schemas, affect regulation, romantic love, marital functioning, and even (going back to our World War II predecessors) group dynamics, prejudice, and intergroup relations.

In this chapter we explain attachment theory in its classic and contemporary forms, placing special emphasis on core concepts, basic principles, and conceptual comparisons and bridges with other theoretical frameworks. Although our primary goal is to explain the theory's core concepts and principles, we of course provide empirical evidence throughout. It is the theory's already proven ability to generate creative and revealing research that makes it worth considering.

CORE CONCEPTS IN ATTACHMENT THEORY AND RESEARCH

In his classic trilogy, *Attachment and Loss*—one of the most cited book series in contemporary psychology—Bowlby (1969/1982, 1973, 1980) asked and answered the following question: Why does "maternal deprivation" (loss of a primary caregiver, such as one's mother, during childhood) have such a potent effect on subsequent personality development (as documented in scores of studies, summarized in Cassidy & Shaver, 1999). By considering a vast array of sources ranging from clinical psychoanalysis to primate ethology and cognitive and cognitive-developmental psychology, Bowlby came to the conclusion that a person's fundamental sense of safety, social acceptance, and well-being rests on the quality of his or her social relationships with "attachment figures." Moreover, when a child has no reliable, trustworthy, secure relationship with one or more such figures, social and affective development is distorted in ways that can eventuate in emotional disorders ranging from anxiety and depression to antisocial personality and other personality disorders. In conceptualizing the importance of attachment relationships to personality development, Bowlby had already moved beyond related but subsequent notions in the literature of social psychology, such as Baumeister and Leary's (1995) claim that human beings have a fundamental "need to belong."

In explaining the motivational bases of personality development, Bowlby (1969/1982) borrowed from ethology the concept of *behavioral system*, a species-universal, biologically evolved neural program that organizes behavior in ways that increase the chances of survival and reproduction despite environmental dangers and demands. Theoretically, the attachment behaviors observed when a person encounters threats and stressors—for example, distress signals, proximity seeking, clinging to a caregiver, and relaxing once proximity and support are provided—are due to a hardwired "attachment behavioral system," just as a caregiver's reactions to a relationship partner's (especially a dependent child's) distress signals and attachment behaviors are due to an innate "caregiving behavioral system." By dividing motivational systems into functional categories such as attachment, caregiving, exploration, affiliation, and sex, Bowlby was able to conceptualize links among, and functional and dysfunctional properties of, these systems in a wide variety of life situations across the lifespan. In the following sections we outline these and other concepts of attachment theory, which have guided research for more than 30 years and contributed to a deep understanding of cognitive, emotional, and self-regulatory processes in close relationships.

The Attachment Behavioral System

According to Bowlby (1969/1982), the *attachment behavioral system* is part of a network of phylogenetically evolved behavioral systems, such as exploration, affiliation, caregiving, and sex, which govern the choice, activation, and termination of behavioral sequences designed to attain particular set goals—states of the person–environment relationship that have adaptive advantages for individual survival and reproduction of genes. These behavioral sequences are "activated" by certain stimuli or kinds of situations that make a particular set goal salient (e.g., loud noises signaling danger, sudden darkness, the presence of a predator) and "deactivated" or "terminated" by other stimuli or situations that signal attainment of the desired goal state.

In Bowlby's (1969/1982) view, behavioral systems also include "ontogenetically learned" components reflecting the particular history of behavioral system activation by a particular person in specific kinds of contexts. Although behavioral systems presumably operate mainly at a subcortical level and in a reflexive, mechanistic manner, their ability to achieve set goals depends on a person's actual transactions with the external world. Therefore, to make goal attainment more likely, behavioral systems include cognitive-behavioral mechanisms, such as monitoring and appraising the effectiveness of behaviors emitted in a particular context, which allow flexible, goal-corrected adjustment of the system's "programming" when necessary to put an individual back on the track of goal attainment. Over time, after operating repeatedly in certain environments, a person's behavioral systems become molded by social encounters, "programming" the neural/behavioral capacities so that they fit with important relationship partners (e.g., parents) and yield effective action in that relational environment. Through this process, a person learns to adjust his or her behavioral systems to fit contextual demands and form reliable expectations about possible access routes and barriers to goal attainment. These expectations, which operate partly at a more conscious and intentional level, become

part of a behavioral system's programming and are sources of both individual differences and within-person continuity in the system's operation.

The presumed biological function of the attachment behavioral system is to protect a person (especially during infancy and early childhood) from danger by assuring that he or she maintains proximity to caring and supportive others (*attachment figures*). In Bowlby's (1969/1982) view, the need to seek out and maintain proximity to attachment figures (what he called "stronger and wiser" caregivers) evolved in relation to the prolonged helplessness and complete dependence of human infants who cannot defend themselves from predators and other dangers. According to Bowlby's evolutionary reasoning, infants who maintained proximity to a supportive caregiver were more likely to survive and eventually to reproduce, causing genes that fostered proximity seeking and other attachment behavior in times of danger to be selected and passed on to subsequent generations. Bowlby (1969/1982, 1988) assumed that although the effects of attachment-system activation are most easily observed during infancy, and the attachment system may operate somewhat differently at different age periods, it continues to function throughout life, as indicated by adults' needs for proximity, support, comfort, and security (Hazan & Zeifman, 1999).

During infancy, primary caregivers (usually one or both parents, but also grandparents, older siblings, daycare workers) are likely to serve attachment functions. Research has shown that, when tired or ill, infants tend to seek proximity to a primary caregiver (e.g., Ainsworth, 1973) and be notably reassured and soothed in that person's presence (e.g., Heinicke & Westheimer, 1966). In adulthood, a wider variety of relationship partners can serve as attachment figures, including familiar coworkers, friends, and romantic partners. They form what Bowlby (1969/1982) called a person's *hierarchy of attachment figures*. There may also be context-specific attachment figures, who are real or potential sources of comfort and support in specific milieus, such as teachers and supervisors in academic settings or therapists in therapeutic settings. Moreover, groups, institutions, and symbolic personages (e.g., God) can become targets of proximity seeking. There is evidence that many young children have imaginary friends (e.g., Gleason, 2002); that some married adults who suffer the death of a spouse continue to experience the spouse's presence and seek his or her assistance and support in times of need (e.g., Klass, Silverman, & Nickman, 1996); and that many adults believe they can and do obtain protection and comfort from gods, angels, and saints (e.g., Fraley & Shaver, 1999; Kirkpatrick, 1999). In addition, there are components of the self that result from internalization of and identification with attachment figures' traits (we call these self-caregiving representations; Mikulincer & Shaver, 2004), and they can serve successfully as symbolic sources of support and comfort.

It is important to understand that the concept of "attachment figure" has a specific meaning in attachment theory. Attachment-related interactions with these people are not viewed as being simply the same as other forms of social interaction. According to attachment theory (e.g., Ainsworth, 1991; Hazan & Shaver, 1994; Hazan & Zeifman, 1994), an attachment figure should accomplish three functions. First, he or she should be a target for proximity seeking. People tend to seek and enjoy proximity to their attachment figures in times of need and to actively resist separation from them. Second, an attachment figure should be, or provide, a *safe haven* in times of need (i.e., reliably provide protection, comfort, support, and relief). Third, an attachment figure should be, or should function as, a *secure base*, allowing a child or adult relationship partner to pursue nonattachment goals in a safe environment. Based on this narrow definition of attachment figures, we view an interaction as attachment relevant when it occurs with a familiar other, or the mental representation of a familiar other, in a stressful context, with the expectation of receiving protection, comfort, or support. This protection and support in the realm of attachment allows a person to function better in nonattachment domains such as exploration, creative thinking, empathic and prosocial behavior toward others, and sexual mating.

In studies of adults, researchers are able to identify a research participant's attachment figures by using a standard measure, the WHOTO questionnaire, developed by Hazan and Zeifman (1994) and Fraley and Davis (1997). The measure asks a respondent to name the particular people on whom he or she relies for various forms of protection, guidance, and support and then to describe the role of each such person in the respondent's life (e.g., mother, father, sibling, romantic partner, and friend). We (Mikulincer, Gillath, & Shaver, 2002) conducted several experiments in which we subliminally primed participants with threat words (e.g., failure and separation) and then determined indirectly (using reaction times in a lexical decision or Stroop task) which names became more available for mental processing when a person felt threatened. It turned out that the names of attachment figures (identified with the WHOTO questionnaire) became more available in response to a threatening word, something that did not happen with the names of other close relationship partners not mentioned in the WHOTO. This and other evidence indicates that attachment figures are not just any relationship partners; rather, they are special individuals to whom a person turns when he or she needs protection and support.

Bowlby (1969/1982) also specified the set goal of the attachment system and described the typical cycle of attachment-system activation and deactivation. The goal of the system is a sense of protection or security (called by Sroufe & Waters, 1977, felt security), which normally terminates the system's activation. This goal is made particularly salient by encounters with actual or symbolic threats and by appraising an attachment figure as not sufficiently near, interested, or responsive. In such cases, the attachment system is activated and the individual is driven to seek and reestablish actual or symbolic proximity to an attachment figure (a process Bowlby called the "primary strategy" of the attachment system). These bids for proximity persist until protection and security are attained. When the set goal of security is attained, the at-

tachment system is deactivated and the individual calmly and coherently returns to nonattachment activities.

In infants, attachment-system activation includes non-verbal expressions of neediness and desire for proximity, such as crying, calling, and pleading, as well as locomotor behaviors aimed at reestablishing and maintaining proximity, such as moving toward the caregiver and clinging (Ainsworth, Blehar, Waters, & Wall, 1978). In adulthood, the primary attachment strategy is not necessarily to engage in actual proximity-seeking behavior. Instead, it may be sufficient to activate soothing, comforting mental representations of relationship partners who regularly provide care and protection or even self-representations associated with these partners (Mikulincer & Shaver, 2004). These cognitive representations can create a sense of safety and security, help a person deal successfully with threats, and allow the person to continue pursuing nonattachment goals without having to interrupt these activities to engage in actual proximity bids.

In support of these ideas, recent studies (Mikulincer, Gillath, et al., 2001; Mikulincer, Gillath, et al., 2003; Mikulincer, Hirschberger, Nachmias, & Gillath, 2001; Mikulincer & Shaver, 2001) show that a variety of experimental techniques designed to activate mental representations of internalized attachment figures (e.g., subliminal presentation of the names of people nominated as attachment figures in the WHOTO; guided imagery concerning the availability of these attachment figures; and visualization of the faces of these figures) improve participants' self-reported mood during an experimental session and unconsciously endow formerly neutral stimuli with positive affect. Specifically, activation of mental representations of attachment figures led to higher liking for unfamiliar Chinese ideographs even under threatening conditions and eliminated the detrimental effects that threats otherwise had on liking. Similar experimental interventions eliminated outgroup negativity, even when participants thought an outgroup member had insulted or challenged their ingroup. Thus, activation of mental representations of security-providing attachment figures seems to have a calming, soothing effect, which reduces threats and has positive effects on assessments of human and inanimate stimuli.

Bowlby (1969/1982), along with Harlow (1959), rejected classical psychoanalytic and Pavlovian behavioral frameworks that portrayed social attachment as a secondary effect of feeding (viewed in terms of drive reduction). In line with "object relations" approaches to psychoanalysis, Bowlby viewed human beings as naturally relationship seeking, naturally oriented to what Harlow called "contact comfort" (in his well-known studies of infant monkeys' attachments to and reliance on real and cloth-surrogate mothers), and as naturally inclined to seek proximity to familiar, comforting figures in times of need. That is, Bowlby viewed proximity–contact and maintenance over time of affectionate, trusting, and supportive interpersonal relationships as innately sought-after goal states and rejection, separation, and the loss of such relationships as aversive antigoal states. Moreover, he viewed successful bids for proximity and the attainment of felt security as necessary for the formation of sat-

isfying interpersonal relationships. Every attachment interaction that alleviates distress and enhances felt security reaffirms the adaptive advantage of closeness and strengthens affectional bonds with a particular relationship partner.

From an emotion-regulation perspective, smooth operation of the attachment system can be viewed as a dynamic, homeostatic process aimed at restoring emotional equanimity. In fact, emotional arousal (e.g., fear, anxiety, and anger) is associated with attachment-system activation; distress alleviation results in attachment-system deactivation; and basic emotions (e.g., love, joy, fear, anger, and sadness) frequently occur during attachment-related interactions. For this reason, the attachment system plays an important role in arousing, regulating, and deescalating emotional states, shaping—over time—a person's affective tone and maintaining emotional equanimity.

Bowlby (1973) and Ainsworth (1991) were especially interested in the reciprocal associations between the attachment system and the exploration and affiliation systems, because they viewed attachment insecurity as a hindrance to the full development of skills associated with the other systems. A child or adult who feels threatened and inadequately protected or supported has a difficult time directing attention and other psychological resources to free play, curious investigation of objects and environments, and the social skills necessary for satisfying affiliative relationships with peers. Considered more generally and extended over a longer period of development, this same interference process is thought to disrupt the normal development of self-efficacy, self-esteem, coping (affect-regulation) skills, and positive, trusting social orientation. Just as being harassed or distracted at school interferes with normal cognitive development, being forced by one's social environment to focus only on threats and insecurity distorts and interferes with social development and results in a person with measurably diminished capacities.

Individual Differences in Attachment-System Functioning

Attachment-Figure Availability, the Sense of Security, and Secondary Strategies

Although nearly all children are born with a normal attachment system, which motivates them to pursue proximity and security in times of need, proximity maintenance and security attainment also depend on the responses of particular relationship partners to one's bids for proximity and safety. According to Bowlby (1973, 1988), the quality of attachment-system functioning depends on the availability of a relationship partner in times of need; the partner's sensitivity and responsiveness to one's bids for closeness, comfort, and support; and the attachment figure's ability and willingness to alleviate distress and provide a secure base from which to activate other behavioral systems. These variations in the nature of the caregiver's, or attachment figure's, responses are thought to be the major sources of individual

differences in attachment-system functioning, because of their impact on operating parameters of the system. (There may also be genetically based temperamental causes of individual differences in infant attachment behavior, but if so they have yet to be convincingly demonstrated empirically; see, for example, O'Connor & Croft, 2001.)

When a relationship partner is available, sensitive, and responsive to an individual's proximity-seeking efforts in times of need, the individual is likely to feel an inner sense of attachment security—a sense that the world is a generally safe place, that attachment figures are helpful when called on, and that it is possible to explore the environment curiously and confidently and to engage rewardingly with other people. This sense is an inner signal that the attachment system is functioning well and that proximity seeking is an effective emotion-regulatory strategy. Moreover, the individual acquires important procedural knowledge about distress management, which becomes organized around a relational script (Waters, Rodrigues, & Ridgeway, 1998). This *secure-base script* includes something like the following if–then propositions: "If I encounter an obstacle and/or become distressed, I can approach a significant other for help; he or she is likely to be available and supportive; I will experience relief and comfort as a result of proximity to this person; I can then return to other activities." This script is a cognitive reflection of the phylogenetically "hardwired" program at the heart of the attachment system; as such, it requires little in the way of changes in the system's operating parameters.

However, when a primary attachment figure proves not to be physically or emotionally available in times of need, not responsive to a person's proximity bids, or poor at alleviating distress or providing a secure base, attachment-system functioning is disrupted and the set goal is not attained. In such cases, the individual does not experience comfort, relief, or felt security. Rather, the distress that initially activated the system is compounded by serious doubts about the feasibility of attaining a sense of security: "Is the world a safe place or not? Can I trust others in times of need? Do I have the resources necessary to manage my own negative emotions?" These worries about self and others, and the resulting sense of vulnerability, can maintain the attachment system in a continually activated state, keep a person's mind preoccupied with threats and the need for protection, and interfere drastically with the functioning of other behavioral systems.

Negative interactions with an inadequately available and responsive attachment figure also signal that the primary attachment strategy is failing to accomplish its set goal. As a result, the operating parameters of the attachment system have to be adjusted and certain *secondary attachment strategies* are likely to be adopted. Attachment theorists (e.g., Cassidy & Kobak, 1988; Main, 1990) have emphasized two such secondary strategies: *hyperactivation* and *deactivation* of the attachment system. Viewed in terms of the famous fight–flight distinction (Cannon, 1939), hyperactivating strategies are "fight" responses to the frustration of attachment needs (Bowlby called it

"protest"). This response comes about in relationships in which the attachment figure is sometimes responsive but only unreliably so, placing the attached person on a partial reinforcement schedule that seems to reward persistence of energetic, strident, noisy proximity-seeking attempts, because they sometimes appear to succeed. In such cases, the individual does not easily give up on proximity seeking and in fact intensifies it to coerce the attachment figure's love and support. The main goal of these strategies is to get an attachment figure, viewed as unreliable or insufficiently available and responsive, to pay attention and provide protection or support. The way to try to attain this goal is to maintain the attachment system in a chronically activated state until an attachment figure is perceived to be adequately available and responsive. This involves exaggerating appraisals of danger and signs of attachment-figure unavailability and intensifying one's demands for attention, affection, and assistance.

Deactivating strategies are a "flight" reaction to an attachment figure's unavailability, which seem to develop in relationships with figures who disapprove of and punish closeness and expressions of need or vulnerability. In such relationships, an individual learns to expect better outcomes if signs of need and vulnerability are hidden or suppressed, proximity-seeking efforts are weakened or blocked, the attachment system is deactivated despite a sense of security not being achieved, and the person attempts to deal with threats and dangers alone (what Bowlby, 1969/1982, called *compulsive self-reliance*.) The primary goal of deactivating strategies is to keep the attachment system turned off or downregulated to avoid frustration and distress caused by attachment-figure unavailability. This deactivation requires denying attachment needs, steering clear of closeness and interdependence in relationships, and distancing oneself from threats that can cause unwanted activation of the attachment system.

Attachment Working Models

According to Bowlby (1969/1982), variations in caregiver responses to an attached individual's bids for proximity and protection not only alter the operation of the attachment system in a particular interaction or short-term series of interactions but also gradually produce more enduring and pervasive changes in attachment-system functioning. According to Bowlby (1973), these long-term effects are explicable in terms of the storage of significant interactions with an attachment figure in an associative memory network. This stored knowledge allows a person to predict future interactions with the relationship partner and adjust proximity-seeking attempts without having to rethink each one. The repeated recording in memory of attachment-related interactions results in increasingly stable mental representations of self, partner, and the relationship. Bowlby called these mental representations *working models* and viewed them as the basis of stable individual differences in attachment-system functioning. The concept is interesting from a social-psychological standpoint, because it is similar to such concepts as "script" and "social schema." As with those

concepts, which originally seemed coolly cognitive because they were inspired by digital computer programs and cybernetic devices, Bowlby viewed them as cognitive-affective structures that include affective memories and contribute importantly to expectations and appraisals that evoke emotion (Shaver, Collins, & Clark, 1996).

Bowlby (1969/1982) thought that interactions with attachment figures were stored in at least two kinds of working models: representations of attachment figures' responses (*working models of others*) and representations of the self's lovability and competence (*working models of self*). Bowlby argued that "if an individual is to draw up a plan to achieve a set-goal not only does he have to have some sort of working model of his environment, but he must have also some working knowledge of his own behavioral skills and potentialities" (p. 112). Thus the attachment system, once it has been used repeatedly in a given relational setting, includes representations of the availability, responsiveness, and sensitivity of a relationship partner as well as representations of the self's own capabilities for mobilizing the partner's support and one's feelings of being loved and valued by the partner. These representations organize a person's memories of attachment interactions and guide future proximity-seeking efforts.

Because working models, at least initially, are based on the internalization of specific interactions with a particular attachment figure, a person can hold multiple working models that differ in the outcome of the interaction (success or failure to attain security) and the strategy used to deal with insecurity in that interaction (hyperactivating, deactivating). Like other mental representations, these working models form excitatory and inhibitory associations with one other (e.g., experiencing or thinking about an episode of security attainment activates memories of congruent episodes of successful proximity maintenance and renders memories of hyperactivation and deactivation less accessible), and these associations favor the formation of more abstract and generalized representations of attachment-system functioning with a specific partner. Thus, models with a specific attachment figure (relationship-specific models) are created, and through excitatory and inhibitory links with models representing interactions with other attachment figures, even more generic working models are formed to summarize different relationships. The end result of this process can be conceptualized as a hierarchical associative memory network that includes episodic memories, relationship-specific models, and generic models of security attainment, hyperactivation, and deactivation. As a result, with respect to a particular relationship and across different relationships, most people can sometimes think about interpersonal interactions in secure terms and at other times think about them in hyperactivating or deactivating terms.

In a recent study, Overall, Fletcher, and Friesen (2003) obtained preliminary evidence concerning the hierarchical nature of the cognitive network of attachment working models. They asked participants to complete attachment measures for three specific relationships within each of three domains—family, friendship, and romantic—and then examined the structural organization of these relationship descriptions. Confirmatory factor analyses revealed that a hierarchical arrangement of specific and global working models best fit the data, indicating that models for specific relationships (e.g., with particular family members) are nested within relationship-domain representations (e.g., family members), which in turn are nested within more global models.

The neural network of attachment-related models has all the usual properties of any cognitive network (e.g., differentiation, integration, and coherence between the various models) (Collins & Read, 1994). In addition, each working model within the network differs in cognitive accessibility (the ease with which it is activated and used to guide the functioning of the attachment system in a given attachment interaction). As with other mental representations, the strength or accessibility of each model is determined by the amount of experience on which it is based, the number of times it has been applied in the past, and the density of its connections with other working models (e.g., Baldwin, 1992; Collins & Read, 1994; Shaver et al., 1996). At a relationship-specific level, the model representing the typical interaction with an attachment figure has the highest accessibility in subsequent interactions with that person. At a generic level, the model that represents interactions with major attachment figures (e.g., parents and romantic partners) typically becomes the most chronically accessible attachment-related representation and has the strongest effect on attachment-system functioning across relationships and over time.

Consolidation of a chronically accessible working model is the most important psychological process accounting for the enduring, long-term effects on personality functioning of attachment interactions during infancy, childhood, and adolescence (Bowlby, 1973; Waters, Merrick, Treboux, Crowell, & Albersheim, 2000). Given a fairly consistent pattern of interactions with primary caregivers during infancy and childhood, the most representative or prototypical working models of these interactions become part of a person's implicit procedural knowledge, tend to operate automatically and unconsciously, and are resistant to change. Thus, what began as representations of specific interactions with a primary caregiver during childhood become core personality characteristics, tend to be applied in new situations and relationships, and shape attachment-system functioning in adulthood.

Although activation of a particular working model depends on the history of attachment interactions, attachment theory also emphasizes the importance of contextual factors that contribute to this activation (e.g., Collins & Read, 1994; Shaver et al., 1996). Recent studies have shown that contextual cues concerning a partner's availability as well as imagined encounters with supportive or nonsupportive others can activate congruent working models, even if they are incongruent with a person's chronically accessible working model (e.g., Mikulincer, Gillath, et al., 2001; Mikulincer & Shaver, 2001). In fact, this chronically accessible model coexists with less typical

working models in the memory network, and these models can be activated by contextual factors in a given situation or social interaction.

An Integrative Model of Attachment-System Functioning in Adulthood

In an attempt to integrate previous control-system representations of the attachment behavioral system (e.g., Bowlby, 1969/1982; Fraley & Shaver, 2000; Shaver & Hazan, 1993) and the immense accumulating empirical evidence on the functioning of this system in adulthood, we (Mikulincer & Shaver, 2003; Shaver & Mikulincer,

2002) proposed a three-phase model of attachment-system activation and dynamics in adulthood. The model (see Figure 28.1) includes three major components. The first component concerns the monitoring and appraisal of threatening events and is responsible for activation of the attachment system. The second component involves the monitoring and appraisal of the availability and responsiveness of attachment figures and is responsible for variations in the sense of attachment security. According to our model, once the attachment system is activated, an affirmative answer to the question, "Is an attachment figure available and likely to be responsive to my needs?" results in a sense of security, fosters the application of the
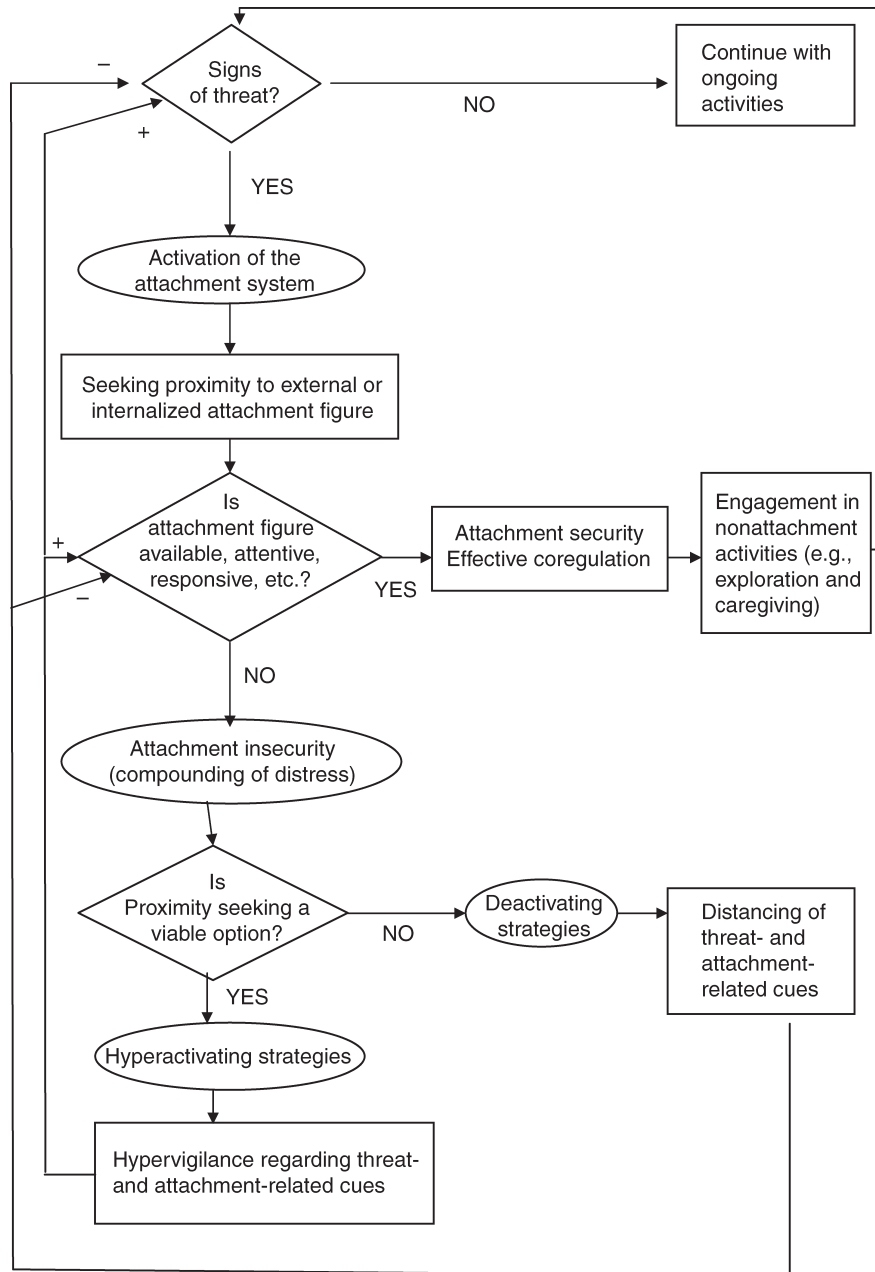
FIGURE 28.1. Shaver and Mikulincer's (2002) integrative model of the activation and dynamics of the attachment system in adulthood.

secure base script, and facilitates engagement in nonattachment activities. The third component concerns monitoring and appraisal of the viability of proximity seeking as a means of coping with attachment insecurity and is responsible for variations in the use of hyperactivating or deactivating strategies. Whereas the appraisal of proximity seeking as likely to be successful, assuming sufficient effort is expended, favors the reliance on hyperactivating strategies, appraising proximity seeking as unlikely to alleviate distress and perhaps even likely to exacerbate it favors the adoption of deactivating strategies. The model also includes hypothetical excitatory and inhibitory "neural circuits" (shown as arrows on the lefthand side of the diagram), resulting from the recurrent use of hyperactivating or deactivating strategies, which affect the monitoring of threats and attachment figures' availability.

The model is sensitive to both context and personality. On the one hand, each component of the model can be affected by specific contextual factors (e.g., actual threats and information about attachment-figure availability or proximity-seeking viability), which initiate a bottom-up process in a person's working models, activating congruent attachment representations, and producing immediate changes in attachment-system functioning. On the other hand, each component of the model is affected by chronically accessible working models, which bias the appraisals of threats, attachment-figure availability, and proximity-seeking viability. These biases are part of a top-down process by which the attachment system functions in accordance with a person's chronic attachment working models. Overall, the model acknowledges the importance of both the context in which the attachment system is activated on a particular occasion and person-specific variations resulting from attachment history and chronically accessible working models.

Conceptualization and Measurement of Attachment Style

According to attachment theory (Bowlby, 1988; Fraley & Shaver, 2000; Shaver & Hazan, 1993), a particular history of attachment experiences and the resulting consolidation of chronically accessible working models lead to the formation of relatively stable individual differences in attachment-system functioning. These stable and generalized individual differences can be empirically examined by measuring a construct called *attachment style*—a person's habitual pattern of expectations, needs, emotions, and behavior in interpersonal interactions and close relationships (Hazan & Shaver, 1987). Depending on how it is measured, attachment style characterizes the functioning of a person's attachment system in a particular relationship (relationship-specific style) or across relationships (global attachment style).

The concept of attachment style, although not given that name, was first proposed by Ainsworth (1967) to describe infants' patterns of responses to separations from and reunions with their mother in the laboratory "Strange Situation" assessment procedure. Using this procedure, infants were originally classified into one of three style categories: secure, avoidant, or anxious. Main and Solomon (1990) later added a fourth category, "disorganized/disoriented," characterized by odd, awkward behavior and unusual fluctuations between anxiety and avoidance.

Infants classified as secure seem to hold chronically accessible working models of security attainment, and their pattern of responses to separation and reunion reflects a stable sense of attachment security. Specifically, they react to separation from their mother with overt expressions of distress but then recover quickly and continue to explore the environment with interest. When reunited with mother, they greet her with joy and affection, respond positively to being held, and initiate contact with her (Ainsworth et al., 1978). Avoidant infants seem to hold chronically accessible working models of unsuccessful proximity-seeking attempts organized around attachment-system deactivation. This organization is manifested in their responses to separation and reunion episodes, where they show little distress when separated from mother and avoid her upon reunion (Ainsworth et al., 1978). Anxious infants also seem to hold chronically accessible working models of frustrated proximity-seeking attempts, but these models seem to be organized around attachment-system hyperactivation. This organization is manifested in the expression of protest and distress during separation episodes and conflictual, angry responses toward mother at reunion (Ainsworth et al., 1978).

In the 1980s, researchers from different psychological fields (developmental, clinical, personality, and social) constructed new measures of attachment style in order to extend attachment research into adolescence and adulthood. Based on a developmental and clinical approach, Main and her colleagues (George, Kaplan, & Main, 1985; Main, Kaplan, & Cassidy, 1985; see Hesse, 1999, for a review) devised the Adult Attachment Interview (AAI) to study adolescents and adults' mental representations of attachment to their parents during childhood. In the AAI, interviewees answer open-ended questions about their childhood relationships with parents and are classified into three categories paralleling Ainsworth's infant typology: "secure" (or free and autonomous with respect to attachment), "dismissing" (of attachment), or "preoccupied" (with attachment). Using the AAI coding system (George et al., 1985), a person is classified as secure if he or she describes parents as available and responsive and his or her memories of relationships with parents are presented in a clear, convincing, and coherent manner. Dismissing persons play down the importance of attachment relationships and tend to recall few concrete episodes of emotional interactions with their parents. Preoccupied individuals are entangled in worries and angry feelings about parents, are hypersensitive to attachment experiences, and can easily retrieve negative memories, but they have trouble discussing them coherently without anger or anxiety. Despite the richness of AAI narratives, which are particularly useful in clinical settings, the interview is costly to administer and score, and it deals exclusively with memories of child–parent relationships.

Working from a personality and social psychological perspective and attempting to apply Bowlby's ideas to the study of romantic relationships, Hazan and Shaver (1987) developed a self-report measure of adult attachment style suitable for use in experiments and surveys. In its original form, the measure consisted of three brief descriptions of feelings and behaviors in close relationships that were intended to embody adult romantic analogues of the three infant attachment styles identified by Ainsworth and colleagues (1978). Participants were asked to read the descriptions and then place themselves into one of the three attachment categories according to their predominant feelings and behavior in romantic relationships. The three descriptions were:

- *Secure*: I find it relatively easy to get close to others and am comfortable depending on them and having them depend on me. I don't worry about being abandoned or about someone getting too close to me.
- *Avoidant*: I am somewhat uncomfortable being close to others; I find it difficult to trust them completely, difficult to allow myself to depend on them. I am nervous when anyone gets too close and often, others want me to be more intimate than I feel comfortable being.
- *Anxious*: I find that others are reluctant to get as close as I would like. I often worry that my partner doesn't really love me or won't want to stay with me. I want to get very close to my partner and this sometimes scares people away.

Hazan and Shaver's seminal study was followed by hundreds of others that used the simple forced-choice self-report measure to examine the interpersonal and intrapersonal correlates of adult attachment style (see reviews by Shaver & Hazan, 1993; Shaver & Mikulincer, 2002). Over time, attachment researchers made methodological and conceptual improvements to the original self-report measure. These improvements included the use of Likert scales for rating the extent to which each of the three descriptions captured one's experiences in romantic relationships (e.g., Levy & Davis, 1988); decomposition of the descriptions into separate items that could be included in multi-item scales (e.g., Collins & Read, 1990; Feeney & Noller, 1990; Simpson, 1990); splitting the avoidant category into "dismissing" and "fearful" subtypes, thus moving from a 3- to a 4-category classification scheme (Bartholomew & Horowitz, 1991); and rewording the instructions and items to examine global attachment style in close relationships generally (not only in romantic relationships) and relationship-specific styles (e.g., Baldwin, Keelan, Fehr, Enns, & Koh Rangarajoo, 1996; LaGuardia, Ryan, Couchman, & Deci, 2000; Mikulincer, Florian, & Tolmacz, 1990).

Today, adult attachment researchers in the fields of personality and social psychology (e.g., Brennan, Clark, & Shaver, 1998) largely agree that attachment styles are best conceptualized as regions in a two-dimensional (anxiety-by-avoidance) space, partly because two dimensions are consistently obtained in factor analyses of attachment measures and partly because Fraley and Waller (1998) demonstrated that dimensional representations of attachment style are more accurate than categorical representations. Interestingly, the two-dimensional space is very similar to the one defined by two discriminant functions in Ainsworth and colleagues' (1978) early summary of research on infant–mother attachment (Fig. 10, p. 102).

The first dimension, attachment-related *avoidance*, is concerned with discomfort with closeness and dependence on relationship partners, preference for emotional distance and self-reliance, and the use of deactivating strategies to deal with insecurity and distress. The second dimension, attachment-related *anxiety*, is concerned with a strong desire for closeness and protection, intense worries about partner availability and one's own value to the partner, and the use of hyperactivating strategies for dealing with insecurity and distress. People who score low on both dimensions are said to be secure or to have a secure attachment style. This region of low anxiety and low avoidance is defined by a chronic sense of attachment security, trust in partners and expectations of partner availability and responsiveness, comfort with closeness and interdependence, and coping with threats and stressors in constructive ways. Throughout the remainder of this chapter we refer to people with secure, anxious, or avoidant attachment styles, or people who are relatively secure, anxious, or avoidant. Although our categorical shorthand can misleadingly foster typological thinking, we will always be referring to fuzzy regions in a two-dimensional space in which people are continuously distributed.

The two attachment-style dimensions can be measured with the 36-item Experiences in Close Relationships (ECR) scale (Brennan et al., 1998), which is reliable in both the internal-consistency and test–retest senses and has high construct, predictive, and discriminant validity (Crowell, Fraley, & Shaver, 1999). Eighteen items tap the avoidance dimension (e.g., "I try to avoid getting too close to my partner" and "I prefer not to show a partner how I feel deep down"), and the remaining 18 items tap the anxiety dimension (e.g., "I need a lot of reassurance that I am loved by my partner" and "I resent it when my partner spends time away from me"). The two scales were conceptualized as independent and have been found to be empirically uncorrelated in most studies.

Hundreds of studies using self-report measures of adult attachment style, some based on three categories, some on four categories, and some on two dimensions, have found theoretically coherent attachment-style variations in relationship quality, mental health, social adjustment, ways of coping, emotion regulation, self-esteem, interpersonal behavior, and social cognitions (see Mikulincer & Shaver, 2003; Shaver & Clark, 1994; Shaver & Hazan, 1993, for reviews). Moreover, recent studies have shown that, despite substantial differences in focus (parent–child vs. adult–adult relationships) and method (brief self-reports vs. extensive interview transcripts), self-report measures of adult attachment style are related to AAI coding scales (Bartholomew & Shaver, 1998; Shaver, Belsky, & Brennan, 2000). These findings

imply that scores on the two kinds of measures are related to each other in sensible ways, and that both are reflections of underlying attachment working models and strategies.

BASIC PRINCIPLES OF ATTACHMENT THEORY

In the previous section, we outlined the operating characteristics of the attachment system, explained some of the context-sensitive and partner-specific variations in the system's functioning, and showed how these variations are studied with interview and self-report measures of adult attachment style. We now consider in more detail some of the personal, dyadic, and broader social consequences of variations in attachment-system functioning. Specifically, we focus on three basic issues related to the core principles of attachment theory: (1) the involvement of the attachment system in emotion regulation; (2) the positive implications of attachment-figure availability and the resulting sense of attachment security for social judgments, self-image, personality development, mental health, and relationship quality; and (3) the defensive biasing of cognition, motivation, and behavior by secondary attachment strategies, either deactivating or hyperactivating, and potential emotional and adjustment problems resulting from these biases.

The Emotion-Regulatory Function of Attachment-System Activation: Seeking Proximity and Support

One basic principle of attachment theory is that encounters with threats and dangers automatically activate the attachment system and increase the salience and urgency of the goal of attaining felt security by gaining proximity to and comfort from an attachment figure. This goal can be viewed either in terms of actually attaining safety or security (i.e., being protected and eliminating or reducing environmental threats) or attaining a more desirable emotional state—felt security rather than fear, anxiety, or anger. According to Bowlby (1969/1982), proximity and support seeking are fundamental elements of a person's repertoire of self-regulation skills. Without the relief and reassurance provided by attachment figures, it is difficult for a person, especially a young child, to acquire and develop other potential social and coping capacities.

In our model of attachment-system functioning in adulthood (Mikulincer & Shaver, 2003; Shaver & Mikulincer, 2002), we describe a two-stage process by which the attachment system can perform its regulatory function when a person is threatened or distressed. In the first stage, threat appraisal results in preconscious activation of the attachment system and automatic heightening of the accessibility of security-providing cognitions in a person's associative memory network (e.g., representations of supportive attachment figures; episodic memories of comforting interactions with these figures; self-representations associated with these figures; and

thoughts related to love and support). These preconsciously activated nodes become ready for use in information processing and, based on recent findings from social cognition research (e.g., Wegner & Smart, 1997), affect a person's state of mind and behavioral intentions even before the person experiences any sign of them in his or her stream of consciousness. In the second stage, this preconscious activation can give rise to conscious thoughts about seeking proximity to security-providing figures as well as conscious behavioral intentions and actual proximity- and support-seeking behavior.

In adulthood, as mentioned earlier, preconscious activation of the attachment system does not necessarily lead to actual proximity-seeking behavior, because activation of mental representations of caring and protecting partners can create a sense of safety and security that soothes the person, allowing him or her to continue to deploy attention in desired directions, carry on with chosen activities, and so on. In such cases, the adult attachment system can accomplish its regulatory function intrapsychically without awareness. There are situations, however, such as physical and psychological traumas, illnesses, or losses, in which symbolic proximity to internalized figures is not sufficient to provide adequate comfort and relief, and in such situations attachment-system activation leads to proximity-seeking behavior. There are also periods of development, such as old age, in which people's resources may be taxed to the point where it becomes necessary to seek actual support from others (Shaver & Mikulincer, 2004).

The self-regulatory processes attendant upon attachment-system activation do not depend only on demands imposed by a specific threat encountered; they can also be biased by excitatory and inhibitory processes related to chronic reliance on hyperactivating or deactivating strategies. For example, these chronic secondary strategies can bias subjective appraisal of threats, with hyperactivating strategies leading to exaggerated threat appraisal and deactivating strategies leading to dismissal of threats and suppression of threat-related thoughts. Second, insecure strategies and working models can affect which attachment-related nodes are automatically activated by threat appraisals, increasing the accessibility of negative attachment-related /thoughts (e.g., worries about attachment-figure unavailability and thoughts about separation and rejection). Third, hyperactivating and deactivating strategies can affect actual engagement in proximity and support seeking, with avoidant people preferring self-reliance and perceiving support seeking as a risky or potentially humiliating strategy and anxious people exaggerating overt expressions of helplessness as a means of eliciting others' compassion and care (Shaver & Mikulincer, 2002).

Preconscious Activation

In recent studies, we examined preconscious activation of the attachment system (Mikulincer, Birnbaum, Woddis, & Nachmias, 2000; Mikulincer, Gillath, &

Shaver, 2002) and found that subliminal priming with a threat-related word (e.g., illness, failure, and separation), as compared with a neutral word (e.g., hat), heightened the cognitive accessibility of attachment-related mental representations. This heightened activation was indicated by faster lexical decision times for proximity-related words (e.g., love and closeness) and names of people designated in the WHOTO as security-providing attachment figures (e.g., the name of a parent, spouse, or close friend). Interestingly, these effects were circumscribed to attachment-related representations and were not found for attachment-unrelated words or the names of people, including very familiar ones, other than attachment figures.

We also documented attachment-style variations in preconscious activation of the attachment system. For example, we found that anxious, hyperactivating strategies lead to attachment-system activation even in neutral contexts and color this activation with worries about separation and rejection. Specifically, people who score relatively high on attachment anxiety (measured by the ECR) exhibit heightened accessibility of attachment-related themes and attachment figures' names following subliminal priming with either threatening or nonthreatening words, and they also exhibit heightened access to words associated with separation and rejection. This pattern of attachment-system activation serves the function of hyperactivating strategies, which is to hold threat-related thoughts in working memory, thereby maintaining chronic activation of the system.

Second, avoidant (deactivating) strategies involve suppressing attachment-related worries and inhibiting attachment-system activation during encounters with attachment-related threats. For people who score relatively high on the avoidance scale of the ECR, worries about rejection and separation are relatively inaccessible. However, these worries *do* become accessible to avoidant individuals in response to threat primes when a "cognitive load" is added to a lexical decision task. Social cognition research has demonstrated that the addition of a "cognitive load" results in increased accessibility of to-be-suppressed material (e.g., Wegner, Erber, & Zanakos, 1993). Thus, our results support the theoretical notion that avoidant people actively suppress attachment-related worries and concerns but have trouble continuing to do so when a cognitive load is added. In addition, when we used the word "separation" as a threat prime, avoidant individuals exhibited decreased access to the names of their attachment figures. It thus seems that avoidant people's attachment system is preconsciously inhibited following thoughts of separation, which may have something to do with prior experiences in which expressions of emotion led to attachment figures' threats to leave.

Our studies have also revealed the functional and adaptive nature of attachment-system activation among more securely attached individuals (those with low scores on the attachment avoidance and anxiety dimensions). For them, heightened accessibility of attachment-related representations has occurred only in response to priming with threat words, and the activation is circumscribed to attachment themes with positive affective connotations. That is, secure people's encounters with threats heighten access to positive thoughts about love, support, and comfort, which, in turn, lead to anticipated relief and comfort (Mikulincer & Shaver, 2003).

The functional and adaptive nature of attachment-system activation in secure individuals was also observed in a recent study focusing on the accessibility of what we call security-based self-representations (Mikulincer & Shaver, 2004)—components or subroutines of the self that originate in interactions with supportive attachment figures. Based on social cognition research dealing with the relational basis of self-representations (e.g., Andersen & Chen, 2002; Baldwin, 1992), we argued that securely attached people, through repeated interactions with responsive attachment figures, form two kinds of self-representations: (1) those derived from how a person sees and evaluates him- or herself during interactions with these figures (*self-in-relation-with-a-security-enhancing-attachment-figure*) and (2) those derived from identification with features and traits of these attachment figure (*self-caregiving representations*). In addition, we proposed that secure people, who react to threats with heightened accessibility of representations of security-enhancing attachment figures (Mikulincer, Gillath, & Shaver, 2002), tend to experience a parallel heightened access to security-based self-representations. These self-representations, which have been formed in connection with threats that were alleviated by attachment figures, are mentally associated with attachment-figure representations and the positive feelings that arise from interactions with these figures. As a result, security-based self-representations can be automatically activated in new situations appraised as threatening and have a self-soothing effect.

To test these ideas, we conducted two two-session studies. In the first session of each study, we asked participants to generate traits that described either a security-enhancing attachment figure (study 2) or their self-in-relation-with-this-figure (study 1). In the second session, we exposed participants to either a threatening or a neutral condition, noted the accessibility of various categories of traits within their self-descriptions, and assessed their emotional state. As predicted, securely attached people reacted to the threat condition with heightened accessibility of security-based self-representations. They rated traits that they had previously used to describe a security-enhancing attachment figure or themselves in relation to this figure as more descriptive of their current self in threatening as compared with emotionally neutral conditions. This heightened accessibility of security-based self-representations was not observed among insecurely attached persons. Moreover, security-based self-representations had a soothing effect: The higher the accessibility of these self-representations, the more positive was a participant's emotional state following a threat. Thus, it appears that securely attached individuals can mobilize caring qualities within themselves as well as representations of being loved and valued in times of need that provide real comfort and allow a person to feel unperturbed.

Seeking Proximity and Support

There is extensive evidence, emanating from different theoretical traditions, that encounters with threats activate proximity and support seeking (e.g., Lazarus & Folkman, 1984; Schachter, 1959). For example, a recent series of "terror management" studies (i.e., studies of ways in which people regulate their fear of death) have shown that the threat generated by death awareness leads people to seek proximity as a means of buffering death concerns (see Mikulincer, Florian, & Hirschberger, 2003). These studies also show that proximity seeking can override other defensive maneuvers, such as endorsing one's own cultural beliefs while derogating deviants and outgroup members or engaging in other efforts to bolster one's self-esteem. Interestingly, and surprisingly in terms of terror management theory, people exposed to death reminders are willing to have their cultural worldview challenged or their self-esteem threatened in order to maintain proximity to relationship partners (e.g., Hirschberger, Florian, & Mikulincer, 2003; Wisman & Koole, 2003).

Research has also documented effects of dispositional attachment style on the use of proximity and support seeking as self-regulatory devices. For example, death concerns have been found to heighten proximity seeking mainly among securely but not among insecurely attached people (Mikulincer & Florian, 2000; Taubman Ben-Ari, Findler, & Mikulincer, 2002). These results fit with Fraley and Shaver's (1998) findings about proximity seeking in response to an impending separation from a romantic partner. They unobtrusively coded behaviors of separating couples in an airport and found that higher scores on attachment avoidance were associated with less frequent contact seeking and more frequent avoidance behavior (turning away, looking elsewhere, watching TV) as separation loomed.

In adult attachment research, there is extensive evidence that the sense of attachment security favors the seeking of support from both informal sources, such as parents, friends, and romantic partners, and formal sources, such as physicians, teachers, and counselors (e.g., Larose, Bernier, Soucy, & Duchesne, 1999; Larose, Boivin, & Doyle, 2001). Moreover, many studies have shown that higher levels of attachment insecurity, mainly along the avoidance dimension, are associated with less reliance on support seeking as a way of coping with stressful events (e.g., Alexander, Feeney, Hohaus, & Noller, 2001; Birnbaum, Orr, Mikulincer, & Florian, 1997; Feeney, 1998; Mikulincer & Florian, 1995, 1998; Mikulincer, Florian, & Weller, 1993).

These attachment-style differences have also been noted in observational studies of actual support-seeking behaviors. For example, Simpson, Rholes, and Nelligan (1992) invited heterosexual dating couples to the laboratory, told the woman in each couple that she was about to experience an anxiety-provoking laboratory procedure, and then asked them to wait with their partner for 5 minutes while the experimenter prepared the apparatus. During this "waiting period," participants' behavior was unobtrusively videotaped, and raters later coded the ex-

tent to which each participant sought the partner's support. As expected, higher scores on the avoidance dimension were associated with less support seeking from the partner mainly when women reported relatively high levels of distress (i.e., where pretheoretical intuitions would lead an observer to expect greater support seeking). Avoidant women often attempted to distract themselves by reading magazines instead of asking for support. In a more recent study, Collins and Feeney (2000) asked people to talk with their dating partner about a personal problem. This interaction was videotaped and raters coded the extent to which participants used direct and indirect ways of seeking support (e.g., directly asking for help and conveying a need for help through expressions of distress). Whereas attachment avoidance was associated with less frequent support-seeking behavior, attachment anxiety was associated with indirect methods of seeking support.

Overall, research indicates that threats activate attachment-related representations and heighten the tendency to seek proximity and support, but the studies also reveal attachment-style differences in the use of proximity seeking as a self-regulatory device: Whereas attachment security heightens reliance on proximity seeking, attachment insecurities interfere with fruitful proximity and support seeking. Anxious individuals evince chronic preconscious activation of attachment-related representations, but the associated concomitant preconscious activation of worries about rejection and abandonment seem to disorganize or unbalance the overt expression of support seeking. Avoidant individuals exhibit a dissociated self-regulatory stance: They react to threats with preconscious activation of the attachment system, but this activation does not reach awareness and is not translated into proximity-seeking behavior. This dissociated stance reflects deactivating strategies, which suppress the ever-present need for love and block access to attachment-related representations (Fraley, Davis, & Shaver, 1998).

Secure Attachment as an Inner Resource: The "Broaden and Build" Cycle of Attachment Security

Another principle of attachment theory is that attachment security has positive effects on self-image, coping, adjustment, interpersonal functioning, and personal growth. According to Bowlby (1988), interactions with available and loving attachment figures are natural building blocks of a solid psychological foundation, and the sense of attachment security that results from these interactions makes people more resilient in the face of adversities and hardships. In our model of attachment-system functioning (Mikulincer & Shaver, 2003), repeated episodes of attachment-figure availability create what we, following Fredrickson (2001), call a *broaden-and-build* cycle of attachment security, which provides inner resources for maintaining emotional equanimity in times of stress, fosters formation of satisfying close relationships, broadens one's repertoire of skills and perspectives, and contributes to natural processes of growth and self-actualization. In the fol-

lowing sections we present theoretical ideas and review empirical evidence concerning these salutogenic effects of attachment security.

Self-Representations

One "build" component of the broaden-and-build cycle of attachment security is the formation of authentic, solidly grounded feelings of self-worth, competence, and mastery that allow people to find comfort and reassurance in their own attributes and qualities while confronting threats and stressors. As mentioned earlier, Bowlby (1973) argued that children construct a model of themselves while interacting with attachment figures in times of need. During episodes of attachment-figure availability, children can easily perceive themselves as valuable, lovable, and special, thanks to being valued, loved, and regarded as special by a caring attachment figure. Moreover, they learn to view themselves as active, strong, and competent because they can effectively mobilize a partner's support and restore emotional equanimity. In this way, interactions with available and responsive others and the resulting sense of attachment security become primary sources of feelings of self-worth and mastery and natural building blocks of what Rogers (1961) called the "real self"—positive self-perceptions derived from others' positive regard during the course of a person's development.

Research consistently shows that attachment security is strongly associated with positive self-representations. As compared with anxiously attached persons, secure persons report higher self-esteem (e.g., Bartholomew & Horowitz, 1991; Mickelson, Kessler, & Shaver, 1997), view themselves as more competent and efficacious (e.g., Cooper, Shaver, & Collins, 1998), and possess more optimistic expectations about their ability to cope with stressful events (e.g., Berant, Mikulincer, & Florian, 2001; Cozarelli, Sumer, & Major; 1998). For example, Cozarelli et al. (1998) found that securely attached women undergoing an abortion reported higher levels of self-efficacy for coping with the abortion beforehand and higher self-esteem several months afterward.

Attachment security is also associated with having a coherent, balanced, and well-organized model of self. In a series of four studies, Mikulincer (1995) found that, although participants who endorsed a secure attachment style tended to recall more positive than negative self-relevant traits, they had ready cognitive access to both positive and negative self-attributes in a Stroop task. In addition, they revealed a highly differentiated and integrated self-organization in trait-sorting tasks and had relatively small discrepancies between actual-self representations and self-standards (ideal-self and ought-self representations). That is, attachment security not only encourages positive self-appraisals but also seems to allow people to tolerate weak points of the self and integrate them within a coherent and overall positive self-structure. Hence, securely attached people are able to feel good about themselves and maintain a stable sense of self-esteem even when they become aware of personal faults or imperfections.

Emotion Regulation and Mental Health

Another "build" component of the broaden-and-build cycle of attachment security is a set of constructive ways of coping, by which people can effectively handle problematic situations and manage distress without creating negative side effects. According to attachment theory, interactions with available attachment figures and the resulting sense of attachment security provide actual and symbolic contexts in which to learn constructive coping strategies (Mikulincer & Shaver, 2003). Beyond strengthening a person's confidence in the effectiveness of support seeking, episodes of attachment-figure availability and support facilitate the adoption of other constructive regulatory strategies embodied in the "secure base script" (Waters et al., 1998): acknowledgment and display of distress, positive reappraisal of the distress-eliciting situation, and engagement in instrumental problem solving.

Interactions with emotionally accessible and responsive others provide the context in which a child comes to openly and flexibly organize and express emotions and understand their functional assets (Cassidy, 1994; Mikulincer & Shaver, 2003). During these interactions, expression of negative affect is responded to sensitively by the attachment figure and reliably leads to distress-alleviating interventions by this caring person. The child thus learns that emotional states can be tolerated and transformed, that acknowledgment and display of emotions are functional steps toward restoring emotional equanimity, and that one can feel comfortable exploring, acknowledging, and expressing one's own emotions (Cassidy, 1994).

In adult attachment research, there is extensive evidence that self-reports of attachment security are associated with higher scores on self-report and behavioral measures of emotional expressiveness (e.g., Feeney, 1995, 1999; Searle & Meara, 1999) and self-disclosure (e.g., Bradford, Feeney, & Campbell, 2002; Keelan, Dion, & Dion, 1998; Mikulincer & Nachshon, 1991). For example, Mikulincer and Nachson (1991) content-analyzed participants' face-to-face verbal disclosure of personal information to another person in a laboratory situation and found that secure participants disclosed more intimate and emotion-laden information than did avoidant participants. Moreover, using a biographical memory task in which participants were asked to recall specific, early memories of positive and negative emotions, Mikulincer and Orbach (1995) found that participants who classified themselves as securely attached had ready mental access to painful memories of anger, sadness, and anxiety and were able to reexperience some of the accompanying negative affect. However, they still had better access to positive memories of happiness and did not experience an automatic spread of associations to memories of other negative emotional experiences. This allows secure people to maintain a positive cognitive context and a well-differentiated emotion-memory architecture, which in turn allows them to process negative memories without becoming overwhelmed by negativity, as often happens in the laboratory and real life for anxious individuals.

According to attachment theory (Bowlby, 1988; Mikulincer & Shaver, 2003), interactions with available and supportive attachment figures promote and reaffirm optimistic and hopeful appraisals of person–environment transactions. During positive interactions with good attachment figures, children gradually become convinced that distress is manageable, external obstacles can be overcome, and restoration of emotional equanimity is only a matter of time. As a result, secure people can make self-soothing reappraisals of aversive events that help them resolve distressing episodes with less strain than experienced by less secure people. That is, they can use what Lazarus and Folkman (1984) called "reappraisal strategies" and Rothbaum, Weisz, and Snyder (1982) called "secondary control"—construal of aversive events as controllable, temporary, and context-specific and construal of the self as capable of managing problematic situations (see Ochsner & Gross, 2004, for documented benefits of reappraisal strategies).

The association between self-reports of attachment security and positive, optimistic appraisals of stressful events has been well documented in social psychological studies (e.g., Berant et al., 2001; Birnbaum et al., 1997; Mikulincer & Florian, 1995, 1998). For example, Berant and colleagues (2001) found that securely attached mothers of infants who were diagnosed with congenital heart defects reported more positive appraisals of motherhood-related tasks, both immediately after the diagnosis and 1 year later, than anxious or avoidant mothers. Moreover, self-reports of attachment security were associated with less use of threat/loss frames in thinking about relationships (e.g., Boon & Griffin, 1996) and the appraisal of romantic interactions, daily social interactions, and small-group interactions in more positive terms (e.g., Pietromonaco & Feldman Barrett, 1997; Rom & Mikulincer, 2003; Tidwell, Reis, & Shaver, 1996).

Experiences of attachment-figure availability also offer opportunities to learn that one's own instrumental actions are often able to reduce distress. For example, a child learns that his or her bids for proximity alter a partner's behavior and result in the restoration of emotional equanimity. As a result, security-providing interactions strengthen a person's reliance on active, instrumental approaches to problem solving. This heightened reliance is further facilitated by another core feature of episodes of attachment-figure availability: Experiencing attachment figures as loving and approving allows secure people to revise erroneous beliefs without excessive fear of criticism or rejection, thus facilitating cognitive changes that are often necessary when designing an effective plan for solving a problem. That is, secure people's confidence that support is available in case of confusion, uncertainty, or disorganization allows them to open their cognitive structures to new information and flexibly adjust their plans for dealing realistically with problematic situations.

In support of this view, secure people have been found to rely on problem-focused strategies while coping with stressful events (e.g., Lussier, Sabourin, & Turgeon, 1997; Mikulincer & Florian, 1998) and to deal with interpersonal conflicts by compromising and integrating their own and their partner's positions (e.g., Carnelley,

Pietromonaco, & Jaffe, 1994; Levy & Davis, 1988) as well as openly discussing the problem and resolving the conflict (e.g., Scharfe & Bartholomew, 1995; Simpson, Rholes, & Phillips, 1996). This constructive approach to emotion regulation was illustrated by Mikulincer (1998b), who found that secure participants' recollections of personal experiences of anger were characterized by adaptive problem-solving actions aimed at repairing the relationship with the instigator of anger.

Attachment security promotes what Lazarus (1991) called a "short circuit of threat," sidestepping the interfering and dysfunctional aspects of emotions while retaining their functional, adaptive qualities. Efficient management of distress results in more and longer periods of positive mood, thereby rendering mood disorders, maladjustment, and psychopathology less likely. Indeed, several studies have documented positive associations between secure attachment and measures of well-being (e.g., Berant et al., 2001; Birnbaum et al., 1997) and negative associations between security and symptoms of depression, anxiety, and hostility (e.g., Cooper et al., 1998; Mickelson et al., 1997; Roberts, Gotlib, & Kassel, 1996), and between security and eating disorders, substance abuse, and conduct disorders (Brennan & Shaver, 1995; Cooper et al., 1998; Mickelson et al., 1997). Recent studies indicate that both dispositional measures of attachment security and contextual manipulations of the sense of attachment security are associated with lower levels of posttraumatic symptoms (e.g., intrusion of traumatic thoughts) among people who were exposed to the traumas of war or terrorism (Mikulincer, Shaver, & Horesh, 2006).

Relationship Quality

The "build" components of the broaden-and-build cycle of attachment security are also manifested in interpersonal behaviors and close relationships. Episodes of attachment-figure availability promote and reaffirm positive beliefs about others' sensitivity, responsiveness, and goodwill. The secure child learns that he or she can count on others' good intentions and depend on others as providers of comfort and relief. These experiences ensure a person that proximity maintenance is rewarding and that interdependent relationships are important for regulating emotions and satisfying needs. As a result, secure people find it relatively easy to trust others, experience and express gratitude, and feel affection toward relationship partners; they also find it easier to tolerate and accept ambiguous or even negative partner behaviors. Accordingly, they feel comfortable with intimacy and interdependence, emphasize the benefits of being together, and organize their interactions around the perceived benefits of intimate, mutually supportive relationships. Thus, attachment security enhances the motivation to be involved in stable couple relationships and contributes to the quality of those relationships.

There is now good evidence that secure people maintain more stable romantic relationships and report higher levels of relationship satisfaction and adjustment (see Mikulincer, Florian, Cowan, & Cowan, 2002, for an

extensive review). This pattern has been repeatedly documented in studies of both dating and married couples and cannot be explained by other personality factors, such as the "big five" personality traits or self-esteem (Mikulincer, Florian, et al., 2002). For example, Davila, Karney, and Bradbury (1999) collected data every 6 months for 3 years from newlywed couples and found that changes in husbands' and wives' reports of secure attachment predicted concurrent changes in both partners' reports of marital satisfaction. Studies have also linked attachment security with greater intimacy (e.g., Feeney & Noller, 1990; Hazan & Shaver, 1987) and stronger commitment (e.g., Shaver & Brennan, 1992; Simpson, 1990).

There are many studies examining associations between attachment security and positive perceptions of romantic partners. As compared to insecure individuals, securely attached people have more positive views of their romantic partners (e.g., Collins & Read, 1990; Feeney & Noller, 1991), perceive their partners as more supportive (e.g., Collins & Read, 1990; Ognibene & Collins, 1998), and feel more trusting and affectionate toward their partners (e.g., Collins & Read, 1990; Mikulincer, 1998c; Simpson, 1990). Attachment security is also associated with positive expectations concerning partner behaviors (e.g., Baldwin, Fehr, Keedian, & Seidel, 1993; Baldwin et al., 1996; Mikulincer & Arad, 1999). For example, Baldwin and colleagues (1993) examined the cognitive accessibility of expectations concerning partner's behaviors in a lexical decision task and found that secure people had poorer access to negative partner behaviors (e.g., partner being hurtful) than did anxious and avoidant people. Attachment security is also associated with more positive explanations of a relationship partner's behavior (e.g., Collins, 1996; Mikulincer, 1998b, 1998c). Collins (1996) asked participants to explain hypothetical negative behaviors of a romantic partner and found that more secure individuals were more likely to attribute partner's negative behaviors to unintentional, unstable, and highly specific causes and less likely to provide explanations that had negative implications for relationship stability.

Broadening of Skills and Perspectives

As mentioned earlier, Bowlby (1969/1982) conceptualized a dynamic interplay between the attachment system and other behavioral systems (e.g., exploration, caregiving, affiliation, and sex). We view this dynamic interplay as the basis for the "broaden" aspect of the broaden-and-build cycle of attachment security, which contributes to the expansion of nonattachment skills, the opening of cognitive structures to novel perspectives, and the actualization of a person's natural talents. We endorse and have pursued Bowlby's idea that insecurity interferes with the activation and unfettered operation of other behavioral systems. Only when an attachment figure is available and a sense of attachment security is restored can a temporarily insecure person devote full attention to nonattachment activities. Moreover, being confident that support is available when needed, securely attached people can take necessary risks and accept important challenges in an effort to expand their skills and perspectives and actualize their potentials.

This implies that attachment security should enhance curiosity and encourage relaxed exploration of new, unusual information and phenomena and favor the formation of open and flexible cognitive structures despite the uncertainty and confusion that a broadening of knowledge might entail. Indeed, several studies have shown that attachment security is associated with greater trait curiosity (Mikulincer, 1997), more willingness to explore new environments (Green & Campbell, 2000), stronger endorsement of mastery-approach goals in achievement settings (Elliot & Reis, 2003), heightening of creativity following induction of positive affect (Mikulincer & Sheffi, 2000), greater cognitive openness and tolerance for ambiguity (Mikulincer, 1997), and less dogmatic thinking (Mikulincer, 1997).

Attachment security is also associated with the incorporation of novel and even inconsistent information into existing cognitive structures (e.g., Green-Hennessy & Reis, 1999; Mikulincer, 1997; Mikulincer & Arad, 1999). For example, Mikulincer (1997) assessed the tendency to make judgments on the basis of early information and to ignore later data. He found that securely attached individuals were less likely than anxious or avoidant individuals to rate a target person based on the first information received. Interestingly, Green and Campbell (2000) found that contextual priming of attachment security heightened people's willingness to explore novel stimuli, and Mikulincer and Arad (1999) reported that asking participants to visualize a supportive other increased cognitive openness and led even chronically anxious or avoidant people to revise their beliefs based on new information.

These effects of attachment security on cognitive openness have also been documented in a recent series of studies on attitudes toward outgroup members (Mikulincer & Shaver, 2001). In these studies we showed that the greater a person's chronic sense of attachment security, the weaker his or her hostile responses to a variety of outgroup members. In addition, priming techniques that momentarily heightened the sense of attachment security eliminated hostile responses to outgroup members. That is, the sense of attachment security promotes tolerant and accepting attitudes toward people who do not belong to one's own group.

Theoretically, the "broadening" effect of attachment security should promote optimal functioning of the caregiving system, which should show itself in a person's proneness and willingness to provide support and care to others who are chronically dependent or temporarily in need. In line with this prediction, studies have shown that attachment security is associated with higher scores on self-report scales tapping responsiveness to a relationship partner's needs (e.g., Feeney, 1996; Kunce & Shaver, 1994) and more supportive actual behaviors toward a distressed partner (e.g., Fraley & Shaver, 1998; Simpson et al., 1992). In a recent study, Westmaas and Silver (2001) found that attachment-related avoidance was associated with negative attitudes toward a person who had been diagnosed with cancer, and attachment anxiety was associ-

ated with high levels of distress during an interaction with the ill person. In addition, Mikulincer, Gillath, and colleagues (2001) found that both dispositional and situationally augmented attachment security were associated with heightened empathy and compassion for a suffering individual.

There is also evidence that attachment security promotes prosocial values. Mikulincer, Gillath, and colleagues (2003) reported that chronic and contextually augmented attachment security was associated with stronger endorsement of personal values reflecting concern for other people's welfare. Recently, Gillath and colleagues (2004) found that avoidant attachment was negatively associated with engagement in various altruistic activities such as caring for the elderly and donating blood. Although attachment anxiety was not related to overall involvement in such volunteer activities, it was associated with more self-enhancing or self-soothing motives for volunteering (e.g., to feel better about oneself and to enjoy a sense of belonging). Overall, these studies indicate that attachment security provides a solid foundation for compassion and altruistic caregiving, whereas attachment insecurities interfere with prosocial feelings and behaviors.

The Defensive Nature of Secondary Attachment Strategies

A third principle of attachment theory is that defensive biases associated with secondary (insecure) attachment strategies distort and damage emotion regulation, negatively color mental representations of self and others, and contribute to psychological and social problems. According to attachment theory (Main, 1990; Mikulincer & Shaver, 2003; Shaver & Mikulincer, 2002), secondary attachment strategies (hyperactivation and deactivation) include psychological defenses against the frustration and pain caused by attachment-figure unavailability. Although they are attempts at adaptation carried out under adverse environmental circumstances, they end up being maladaptive when used in later relationship situations where security would be more productive. Each of the secondary strategies is aimed originally at achieving a workable relationship with an inconsistently available or consistently distant or unavailable attachment figure. To sustain these strategies, a person has to build otherwise distorted or constraining working models and affect-regulation mechanisms that are likely to interfere with subsequent development and attempts to create rewarding close relationships. In the following pages, we review theoretical proposals and empirical evidence regarding the various defensive biases imposed by secondary attachment strategies and their potentially pathogenic effects on adjustment and mental health.

Emotion Regulation

Secondary attachment strategies defensively bias emotion regulation and alter, obstruct, or suppress the experience and expression of emotions (Cassidy & Kobak, 1988; Mikulincer & Shaver, 2003). The deactivating strategies used by avoidant individuals are intended to dodge or suppress every emotional state associated with threat-related thoughts (e.g., fear, sadness, and shame), because these thoughts can activate unwanted attachment-related needs, memories, and behaviors. Moreover, avoidant people often view negative emotions and expressions of weakness or vulnerability as incompatible with their desire for and maintenance of self-reliance. This causes them to inhibit natural emotional reactions to relationship threats, such as rejection, separation, and loss; and to try to keep these feelings out of consciousness.

Unlike relatively secure people, those who are avoidant cannot engage readily in optimal problem solving because this often requires opening knowledge structures to new information, admitting frustration and possible defeat, dealing with uncertainty and confusion, and running freely through one's memories without attempting to block attachment-system activation (Mikulincer, 1997). Avoidant people have difficulty reappraising emotion-eliciting events because, during frustrating interactions with unavailable, unresponsive, or disapproving attachment figures, they have been forced to doubt the general goodness of the world and good intentions of other people. They have trouble looking on the bright side of troubling events, transforming threats into challenges, and anticipating other's support if they allow themselves to become demoralized.

Deactivating strategies may also block direct confrontation with undesirable emotional states. Avoidant people often prefer to dissociate their emotions from their thoughts and actions, using what Lazarus and Folkman (1984) called *distancing coping*. This requires suppression of emotion-eliciting thoughts, repression of painful memories, diversion of attention from emotion-related material, and inhibition of verbal and nonverbal expressions of emotion.

Bowlby (1980) characterized avoidant individuals' deactivation of emotions in terms of defensive exclusion and segregated mental systems. Bowlby (1988) suggested that the excluded information is stored in mental representations that are blocked from consciousness, not integrated into the stream of consciousness and the conscious determination of behavior, and inaccessible to new information or constructive reappraisal. Defensive exclusion lowers the accessibility of threat- and attachment-related cognitions and creates difficulties in encoding material that is congruent with them. When encoded, this information tends to be processed in a shallow way, because it has no strong excitatory associations with other accessible cognitions.

Unlike secure and avoidant people, who perceive threat-related emotions as goal-incongruent states that should either be managed effectively or suppressed, anxiously attached people perceive these emotions as congruent with their goal of attachment-system hyperactivation. In the process of emotion regulation, anxious hyperactivating strategies are manifested in effortful attempts to generate and intensify emotional states. These states include every emotion that plays a role in activating the attachment system—threats, dangers, and negative interactions with attachment figures. They also include

emotions that emphasize a person's wounds and incompetence, such as sadness, anxiety, shame, and guilt, because these make it natural to insist on attachment figures' attention and care (Cassidy, 1994).

How do anxious people accomplish their goal of intensifying their emotions? One method is to overemphasize the potentially threatening aspects of even benign events, another is to transform challenges into threats, and another is to ruminate on pessimistic beliefs about one's inability to manage distress (Mikulincer & Florian, 1998). Another method is what Lazarus and Folkman (1984) called *emotion-focused coping*—shifting attention to internal indications of distress, thereby making them seem more urgent and destabilizing. This maneuver includes hypervigilance to the physiological aspects of emotional states, heightened recall of threat-related experiences, rumination on real and potential threats, exacerbation of negative feelings, and exaggerated displays of distress (Shaver & Mikulincer, 2002). Another hyperactivating strategy is to engage in wild, counterphobic behavior that makes danger more real or to adopt ineffective courses of action that are likely to be self-defeating and result in failure. All these strategies create a self-amplifying cycle of distress, which is maintained cognitively by ruminative thoughts and feelings even after a threat objectively ends.

Interestingly, although hyperactivating and deactivating strategies lead to opposite patterns of emotional expression (intensification versus suppression), both result in dysfunctional emotional experiences. Avoidant people lose out on the adaptive aspects of emotional experiences because they have poor access to their emotions; anxious people lose out because their attention is devoted to the threatening and interfering aspects of emotions more than their functional aspects. These tendencies, once the province of psychoanalytical clinicians, have now been extensively documented in empirical studies of attachment style and ways of coping with stressful events (see Fuendeling, 1998; Mikulincer & Florian, 1998; Shaver & Mikulincer, 2002, for reviews). In these studies, higher avoidance scores are associated with higher scores on measures of coping by distancing, and attachment anxiety is associated with higher scores on measures of emotion-focused coping. For example, Mikulincer and Orbach (1995) reported that avoidant attachment was associated with a repressive coping style, Feeney (1995) reported that avoidance was related to behavioral blunting (seeking distractions when dealing with stress), and Mikulincer and Florian (1998) found that people who classified themselves as anxiously attached tended to report more frequent task-related, ruminative worries after failing cognitive tasks than were reported by their secure and avoidant counterparts.

Attachment strategies are also manifested in the ways people cope with attachment-related threats (see Mikulincer & Shaver, 2003; Shaver & Mikulincer, 2002, for reviews). For example, Mayseless, Danieli, and Sharabany (1996) and Scharf (2001) found that whereas anxiously attached people reacted to imagined separations in the projective Separation Anxiety Test with strong emotional responses (distress intensification, self-blame), avoidant people refrained from dealing with the threat of separation. In a related pair of studies, Fraley and Shaver (1997) examined the role of secondary attachment strategies in the suppression of separation-related thoughts. Participants wrote continuously about whatever thoughts and feelings they were experiencing while being asked to suppress thoughts about their romantic partner leaving them for someone else. Attachment anxiety was associated with poorer ability to suppress separation-related thoughts—more frequent thoughts of breakup following the suppression task and higher skin conductance during the task. In contrast, more avoidant individuals were better able than less avoidant individuals not only to stop thinking about separation but also to reduce the intensity of their autonomic responses to these painful thoughts. Fraley, Garner, and Shaver (2000) showed that these avoidant defenses act in a *preemptive* manner by holding attachment-related material out of awareness right from the initial encoding of the information.

In a series of studies examining the experience and management of death anxiety (e.g., Mikulincer & Florian, 2000; Mikulincer et al., 1990), anxious individuals were found to intensify death concerns and keep death-related thoughts active in memory. Specifically, attachment anxiety was associated with heightened fear of death at both conscious and unconscious levels, as well as heightened accessibility of death-related thoughts even when no death reminder was present. In contrast, avoidant individuals tended to suppress death concerns and dissociate their conscious claims from their unconscious (but measurable) anxiety. Although avoidance was related to low levels of self-reported fear of death, it was also related to heightened death anxiety in projective Thematic Apperception Test (TAT) stories.

Avoidant people's dissociative tendencies were also documented by Mikulincer (1998a), who found that avoidant individuals, as compared with secure ones, reacted to anger-eliciting episodes with lower levels of self-reported anger and higher levels of physiological arousal (heart rate). Two other studies examined access to emotions during the AAI, finding that avoidant people expressed fewer negative feelings during the interview but displayed higher levels of physiological arousal (heightened electrodermal activity; Dozier & Kobak, 1992) and a high rate of facial expressions of anger, sadness, and negative surprise (Zimmerman, 1999) while speaking about their relationships with parents.

The biases associated with anxious and avoidant approaches to emotion regulation were also documented in Mikulincer and Orbach's (1995) study of emotional memories. Anxious individuals quickly accessed negative emotional memories and then had difficulty controlling the spread of activation from one negative emotional memory to another (a process associated with being classified as anxious, or preoccupied, in the AAI; Hesse, 1999). These findings fit with the theoretical portrayal of anxious people as having an undifferentiated, chaotic emotional architecture, which makes emergence from negative emotional spirals difficult. In contrast, Mikulincer and Orbach found that avoidant individuals

had poor access to negative emotional memories, and those that were recalled were rather shallow (a pattern also characteristic of dismissingly avoidant individuals in the AAI).

Defensive Distortions of Mental Representations of Self and Others

According to attachment theory, secondary attachment strategies defensively bias insecure persons' working models of self (e.g., Bowlby, 1988; Main, 1990; Mikulincer & Shaver, 2003). Whereas hyperactivating strategies negatively bias anxious people's sense of self-esteem, deactivating strategies favor defensive processes of self-enhancement and self-inflation. On the one hand, the excitatory pathways (in Figure 28.1) running from hyperactivating strategies to the monitoring of threat-related cues causes attention to be directed to self-relevant sources of distress (e.g., thoughts about personal weaknesses), thereby fostering chronic doubts about self-worth. This low self-esteem can be exacerbated by self-defeating self-presentational tendencies, which involve emphasizing helplessness and vulnerability as a way of eliciting other people's compassion and support. On the other hand, the inhibitory circuits associated with deactivating strategies (in Figure 28.1) divert attention from self-relevant sources of distress and therefore inhibit consideration of negative self-aspects and contribute to the maintenance of high self-esteem. This defensive inflation of self-esteem is further reinforced by adopting a self-reliant attitude, which requires exaggeration of strengths and self-worth, and by strategic attempts to convince others that one does not need their support.

In a direct examination of these defensive biases, Mikulincer (1998a) exposed people to various kinds of threatening and neutral situations and assessed self-appraisals following the manipulations. Participants with an avoidant attachment style made more explicit and implicit positive self-appraisals following threatening, as compared with neutral, situations. In contrast, anxiously attached participants reacted to threat with self-devaluation, making more explicit and implicit negative self-appraisals following threatening than neutral conditions. Mikulincer also noted that introducing contextual factors that inhibited defensive tendencies (a "bogus pipeline" device that measures "true feelings about things" or the presence of a friend who knew the participants) inhibited avoidant participants' self-inflation response to threats as well as anxious participants' self-devaluation response. That is, insecure people's self-appraisals seemed truly to be strategic defensive maneuvers aimed at convincing other people of the strength of the avoidant self or the neediness of the anxious self.

Secondary attachment strategies are also likely to bias person perception. In the case of avoidant individuals, who want to maintain distance from others and view themselves as strong and perfect, their deactivating strategies are likely to be directed toward increasing distinctiveness, uniqueness, and devaluation of others. In contrast, in the case of anxiously attached people, who want

to be loved and accepted, their hyperactivating strategies are likely to be directed toward increasing the sense of connectedness and belongingness and creating a false sense of consensus. Indeed, Mikulincer, Orbach, and Iavnieli (1998) found that whereas anxious individuals were more likely than their secure counterparts to perceive others as similar to themselves, and to exhibit a false consensus bias in both trait and opinion descriptions, avoidant individuals were more likely than secure individuals to perceive others as dissimilar to them and to exhibit a false distinctiveness bias. Importantly, Mikulincer and colleagues also found that anxious individuals reacted to threats by generating a self-description that was more similar to their partner's self-description, thereby increasing the justification for solidarity. Avoidant individuals, in contrast, reacted to the same threats by generating self-descriptions that were less similar to their partner's self-description and by forgetting more traits that they and their partner shared.

In a subsequent study, Mikulincer and Horesh (1999) found that avoidant people defensively projected their own unwanted traits onto others, which increased self–other differentiation and, by comparison, enhanced their own sense of self-worth. In contrast, anxiously attached participants projected their own traits onto others, which increased their sense of self–other similarity, compatibility, and closeness. Importantly, these two seemingly different mechanisms resulted in a negative appraisal of others. In the case of avoidant persons, the negative appraisal was derived from the projection onto others of negative self-relevant traits. In the case of anxious individuals, it was derived from a tendency to perceive others the way they perceive themselves—as relatively weak, helpless, unworthy, and unlovable.

Problems in Mental Health and Adjustment

Attachment theorists view secondary attachment strategies as risk factors that reduce resilience in times of stress and contribute to emotional problems and poor adjustment (Bowlby, 1988; Mikulincer & Shaver, 2003). Hyperactivating strategies lead to distress intensification and a chaotic emotional architecture that impairs anxious people's ability to regulate negative emotions. As a result, the anxious person experiences an endless and uncontrollable flow of negative thoughts and emotions, which in turn can lead to cognitive disorganization and, in certain cases, culminate in psychopathology. Although avoidant, deactivating strategies contribute to defensive maintenance of a façade of security and calmness, they block access to emotions and hence can impair a person's ability to confront and cope with life's adversities. This impairment is particularly likely to be manifested during prolonged, highly demanding stressful experiences that require active confrontation of a problem and mobilization of external sources of support. In addition, although deactivating strategies involve suppressing the conscious experience and display of distress, the distress can still be indirectly manifested in somatic symptoms, sleep problems, and other health problems. Moreover, negative attitudes toward close relationships and rela-

tionship partners can channel unresolved distress into feelings of hostility, loneliness, and estrangement from others.

With regard to hyperactivating strategies, a large number of studies have shown that attachment anxiety is inversely associated with well-being and positively associated with global distress, depression, anxiety, eating disorders, substance abuse, conduct disorders, and severe personality disorders (see Lopez & Brennan, 2000; Mikulincer & Florian, 2001; Mikulincer & Shaver, 2003, for extensive reviews). These associations have been found in different age groups, ranging from adolescents to elderly adults, community samples, psychiatric inpatients and outpatients, and individuals experiencing acute stressful events (e.g., abortion) or more chronic stressful conditions (e.g., chronic pain).

For avoidance, the findings are more complex. On the one hand, a host of studies yielded no significant associations between avoidant attachment and self-report measures of well-being and global distress (see Mikulincer & Florian, 2001; Mikulincer & Shaver, 2003, for reviews). On the other hand, several studies indicate that avoidant attachment is associated with particular patterns of emotional and behavioral problems that may result from the underlying action of deactivating strategies. Specifically, significant associations have been found between avoidance and a pattern of depression characterized by perfectionism, self-punishment, and self-criticism (e.g., Zuroff & Fitzpatrick, 1995), heightened reports of somatic complaints (e.g., Mikulincer et al., 1993), a hostile view of other people (e.g., Mikulincer, 1998b), substance abuse and conduct disorders (e.g., Brennan & Shaver, 1995; Cooper et al., 1998; Mickelson et al., 1997), and schizoid and avoidant personality disorders (e.g., Brennan & Shaver, 1998).

In addition, whereas no consistent association has been found in community samples between avoidant attachment and global distress, studies that focus on highly demanding and distressing events reveal that avoidant attachment is related to higher levels of reported distress. For example, Berant and colleagues (2001) assessed mothers' reactions to the birth of an infant with a congenital heart defect and found that avoidance, as assessed at the time of the initial diagnosis of the infant's disorder, was the most potent predictor of maternal distress a year later. Moreover, Mikulincer, Horesh, Eilati, and Kotler (1999) found that avoidance was positively associated with global distress among Israeli Jewish settlers whose lives were in danger because of residing in disputed territory controlled by the Palestinian Authority.

It seems that deactivating strategies may contribute to mental health under fairly normal circumstances characterized by only mild encounters with stressors. Under highly demanding conditions, however, these strategies seem to collapse, and in such cases avoidant individuals may exhibit high levels of distress and emotional problems. This conclusion is supported by two of our recent laboratory studies (Mikulincer, Dolev, & Shaver, 2004), which showed that the addition of a demanding cognitive task, which had previously been shown to interfere with mental suppression (e.g., Wegner et al., 1993), impaired avoidant individuals' ability to block the activation of attachment-related worries. Under low cognitive-load conditions, avoidant individuals were able to suppress thoughts related to the painful breakup of a romantic relationship and did not exhibit activation of negative self-traits after being asked to think about this painful episode. However, when a cognitive load was imposed (a secondary but demanding cognitive task), avoidant individuals exhibited a strong rebound of previously suppressed thoughts about the painful separation and heightened activation of negative self-traits. In other words, under high-load conditions, avoidant participants resembled their anxiously attached counterparts, exhibiting high accessibility of separation-related thoughts and an automatic spread of activation from these attachment-related thoughts to negative self-representations.

Adult attachment studies also provide insights into the psychological mechanisms that may account for the association between insecure attachment patterns and measures of emotional and adjustment problems. For example, Roberts and colleagues (1996) found that negative concepts of the self, others, and the future mediated both cross-sectional and prospective associations between attachment anxiety and depression. In addition, the core procedural components of hyperactivating and deactivating strategies play an important mediating role. Whereas the association between attachment anxiety and negative affectivity is explained by heightened reliance on emotion-focused coping and mental rumination on threat-related thoughts (e.g., Birnbaum et al., 1997; Cozzarelli et al., 1998), the association between avoidance and negative affectivity is mediated by heightened reliance on distancing coping, high levels of emotional control, and reluctance to engage in support seeking (e.g., Birnbaum et al., 1997; Cozzarelli et al., 1998).

BUILDING CONCEPTUAL BRIDGES BETWEEN ATTACHMENT THEORY AND OTHER THEORETICAL FRAMEWORKS

In the remaining sections, we point to similarities and differences between attachment theory and four other broad psychological approaches to understanding the human mind: psychoanalysis; relational interdependence theories; social cognition theories; and humanistic and "positive psychology" perspectives on personal development. In so doing, we hope to deepen the reader's understanding of the implications of attachment theory and build conceptual bridges to other theoretical approaches.

Psychodynamic Foundations of Attachment Theory

The links between attachment theory and psychoanalysis were evident in Bowlby's early writings (e.g., Bowlby, 1956). He was trained as a child psychiatrist and psychoanalyst, and like other psychoanalytic thinkers, he assumed that the explanation of adult behavior lay

somewhere in childhood, especially in early social relationships. Although he was dissatisfied with the conventional psychoanalysis of his time, especially the ideas of Anna Freud and Melanie Klein, he still believed that the quality of a child's emotional ties with mother had tremendous effects on normal and abnormal patterns of personal, interpersonal, and social functioning across the life span. Furthermore, Bowlby constructed attachment theory around themes that defined most of the psychoanalytic theories of his time: satisfaction and frustration of basic inner wishes (for security and protection), inner conflicts associated with barriers to wish fulfillment, psychological defenses aimed at avoidance or suppression of negative emotions associated with inner conflicts, and emotional problems related to the overuse of defenses.

These conceptual commonalities become more evident when analyzing the basic postulates that define contemporary psychodynamic approaches (Westen, 1998). In his impressive review of contemporary psychoanalysis, Westen (1998) asserted that all contemporary psychodynamic theorists agree with five core postulates. First, a large portion of mental life is unconscious. Second, cognitive and affective processes operate in parallel so that people can have conflicting motives, thoughts, and feelings toward the same situation or person, and psychological defenses are often used to deal with these conflicts. Third, childhood experiences play a crucial role in the formation of adult personality. Fourth, mental representations of the self and others are major components of personality; they often explain a person's behavior in interpersonal and social settings, and account for or contribute to psychological disorders. Fifth, healthy personality development is a journey from social dependence to mature autonomy.

Attachment theorists and researchers adhere to all five postulates. According to attachment theory, many components of the attachment behavioral system can operate unconsciously (Mikulincer & Shaver, 2003). As reviewed earlier, recent studies have shown that activation of the attachment system can occur at an unconscious level and can shape a person's processing of information and behavior before he or she reflects on any of it in the stream of consciousness (e.g., Mikulincer, Gillath, & Shaver, 2002). In addition, deactivating strategies seem to operate at an unconscious level. Avoidant people often seem not to be consciously aware of suppressing or denying attachment needs and attachment-related thoughts and memories (Cassidy & Kobak, 1988). Furthermore, according to Bowlby (1988) and some of our own research, these suppressed needs, memories, and thoughts continue to remain active in unconscious, segregated mental systems and at times resurface in experience and action when deactivating strategies prove insufficiently strong given other cognitive or emotional demands on mental resources.

In attachment theory, the concepts of inner conflict and psychological defense are central to the characterization of the goals and operation of secondary attachment strategies. Specifically, these strategies seem to reflect the underlying presence of distress-eliciting, conflicting tendencies toward the self and relationship partners and are organized around specific defensive maneuvers against these attachment-related sources of distress (Mikulincer & Shaver, 2003). Hyperactivating strategies reflect a compromise between conflicting, ambivalent tendencies toward attachment figures—anger and hostility toward unavailable attachment figures together with an intense need for proximity to and love from these frustrating figures (Cassidy & Kobak, 1988). Deactivating strategies are organized around conflicting tendencies at different levels of awareness, with lack of negative emotions and a detached attitude evident at the conscious level while high levels of unresolved attachment-related distress exist at an unconscious level (Shaver & Mikulincer, 2002).

Three additional features of attachment theory fit with the remaining postulates of contemporary psychodynamic theories. According to Bowlby (1973), childhood experiences with primary caregivers have important effects on attachment-system functioning in adulthood, and as stated earlier, mental representations of the self and others (attachment working models) explain how mental residues of these early experiences become building blocks of a person's cognitions and behaviors in adulthood and have a shaping influence on emotion regulation, interpersonal relations, and mental health. Furthermore, in attachment theory, the consolidation of dispositional attachment security, a sign of healthy personality development and functioning, provides a foundation for increased exploration, self-regulation, and a flexible balance between self-reliance and reliance on others, which facilitates a move toward maturity and relative autonomy combined with an ability to rely comfortably on others when necessary (Mikulincer & Shaver, 2004). This developmental progression stands in marked contrast with the overly dependent, infantile position of the anxious person and the rigidly self-reliant attitude of the avoidant person.

This does not mean, however, that attachment theory can be simply equated with psychoanalysis. In fact, attachment theory offers a unique perspective on the developmental trajectory of working models and the role played by contextual factors in shaping cognitions and behaviors in adulthood. While contemporary psychoanalysis still views mental representations of self and others in adulthood as mental residues of childhood experiences, Bowlby (1988) believed that the developmental trajectory of working models is not linear or simple and that these mental representations in adulthood are not exclusively based on early experiences. Rather, they can be updated throughout life and can be affected by a broad array of contextual factors, such as current interactions with a relationship partner, the partner's attachment style and dynamics, and a person's current life situation, which can moderate or even override the effects of mental residues of past experiences. Research supports this complex version of the developmental trajectory of working models, with longitudinal studies showing only a moderate level of stability in attachment orientations

from infancy to adolescence and indicating that life events (e.g., parental death) can substantially alter a person's working models (see Fraley, 2002, for a review and meta-analysis of these studies). In this respect, attachment theory, especially as fleshed out by social psychologists, owes a great deal to other conceptual and methodological paradigms in social psychology.

The changing nature of attachment styles and underlying mental representations and affect-regulation strategies is also evident in our model of attachment-system functioning (Mikulincer & Shaver, 2003), in which bottom-up processes initiated by the presence of contextual cues about attachment-figure availability or proximity-seeking viability can alter the functioning of the system. As discussed earlier, attachment theory and research suggest that a particular individual can possess multiple, even conflicting, working models of self and relationship partners beyond the working models that evolved from childhood experiences with parents, and these different models can be contextually activated in experimental settings and have varying effects on cognitions and behaviors (e.g., Baldwin et al., 1996; Collins & Read, 1994; Shaver et al., 1996). Indeed, recent studies have shown that positive effects of contextual priming of security-enhancing representations are also found even among chronically insecure people (e.g., Mikulincer, Gillath, et al., 2001; Mikulincer & Shaver, 2001). Thus, attachment theory does not assert that a person's current attachment orientation must mirror or match his or her attachment orientations with parents during childhood. Rather, the current orientation is a complex amalgam of historical and contemporary contextual factors, which enable the "reworking" of mental representations of self and attachment figures across the lifespan.

There is more to learn about how these changes occur and what psychological mechanisms are involved. Two recent studies provide initial information about the mechanisms, while highlighting the crucial role played by the subjective appraisal of person–environment transactions (Davila & Cobb, 2004; Simpson, Rholes, Campbell, & Wilson, 2003). Simpson and colleagues (2003) examined changes in attachment orientations during the transition to parenthood and found that prenatal appraisals of support and anger explained the way attachment orientations changed across this transition. Specifically, women who perceived less spousal support and more spousal anger during pregnancy became more anxiously attached across the transition, whereas husbands who perceived themselves as providing more spousal support during pregnancy became less avoidant across the transition to parenthood. Davila and Cobb (2004) conducted an 8-week daily diary study, during which participants reported on daily life events and daily levels of attachment security, and found that negative fluctuations in attachment security were explained by the extent to which people appraised the events as involving interpersonal loss. The findings from these two studies indicate that life events can change attachment orientations if people construe the events as disconfirming their chronically accessible working models.

The Relational Basis of Attachment Theory

The preceding discussion of updating working models highlights the importance of the relational context in which the attachment system is activated. Although attachment-system functioning is a reflection of intrapsychic processes related to a person's wishes, fears, and defenses, it can be expressed in behavior (proximity seeking to a relationship partner) and is sensitive to the relational context in general and to the relationship partner's particular responses (availability, responsiveness) on a specific occasion (Shaver & Hazan, 1993). In fact, attachment-system functioning involves real or imagined interpersonal interactions with actual or internalized attachment figures and can be altered by these figures' responses to one's proximity bids. In this respect, attachment theory has a lot in common with interdependence theories of close relationships (e.g., Thibault & Kelley, 1959; Van Lange, De Cremer, Van Dijk, & Van Vugt, Chapter 23, this volume), which focus on the interpersonal interaction as the unit of analysis and emphasize the powerful influence that one person's responses can exert on a partner's cognitions and behaviors. In attachment theory, this interdependence is evident when the attachment system becomes activated and the responses of an attachment figure can affect the operation of the attachment system.

Attachment theory acknowledges the important effects of the relational context on a person's attachment orientation and on relational cognitions and behaviors in a particular interaction (Shaver & Hazan, 1993). In our model (Mikulincer & Shaver, 2003), the three modules of attachment-system functioning can be affected by a partner's behaviors: The partner can be a source of threat and therefore trigger attachment-system activation (e.g., by threatening abandonment or violence) and can affect the appraisal of attachment-figure availability as well as the viability of proximity seeking as a means of achieving security. Moreover, a person's relational cognitions and behaviors depend not only on the functioning of his or her attachment system but also on the partner's attachment behaviors. Indeed, several studies have shown that both partners' attachment orientations contribute uniquely to the prediction of both partners' relationship satisfaction (e.g., Brennan & Shaver, 1995; Collins & Read, 1990). In addition, other studies, using observational techniques, diary keeping, and narrative accounts, have revealed that a person's attachment anxiety and avoidance have differential effects on relational emotions, cognitions, and behaviors depending on the partner's attachment scores (e.g., Collins & Feeney, 2000; Feeney, 2002; Simpson et al., 1992).

It is important to recall, however, that attachment theory is not exclusively relational. As discussed earlier, the theory includes the important idea that the internalization of interactions with attachment figures can be biased by defensive processes related to secondary attachment strategies (e.g., gaining proximity to an insufficiently available attachment figure by hyperactivating the attachment system or avoiding punishment or perpetual frustration by deactivating the system). Because of such bi-

ases, working models of the self and others do not exclusively reflect the ways the person and the partner actually behave in a given interaction. Rather, they are blended reflections of what actually happens in a social encounter as well as subjective biases resulting from attachment working models and strategies.

These defensive biases can also be noted in the subjective appraisal of a partner's responses to one's proximity bids. Whereas anxious people's hyperactivating strategies slant perception in the direction of noticing or imagining insufficient interest, availability, and responsiveness on the part of a partner, avoidant individuals' deactivating strategies increase the likelihood that genuine and clear-cut signals of attachment-figure availability are missed (Mikulincer & Shaver, 2003). These biases reflect a top-down process by which the most chronically activated working models moderate or override the potential influence of a relationship partner's actions, thereby constraining the nature of the interdependent interaction.

Furthermore, attachment-system activation in adulthood can occur intrapsychically without any overt expression in interpersonal behavior and without demanding the intervention of an actual relationship partner. In such cases, a person can search for comfort and security in his or her own mental representations without seeking proximity to or support from an actual relationship. In our recent analysis of security-based self-representations (Mikulincer & Shaver, 2004), we argued that these mental representations can be applied even in situations that are not explicitly social–relational. We showed empirically that secure people are likely to have internalized both self-soothing processes and some of their attachment figures' personal qualities, which they then use when encountering the frustration of failing repeatedly at a laboratory task, even though dealing with feelings of task failure is not particularly social or relational. This is just one example of ways in which attachment-related experiences may affect a person's cognitions and behaviors outside relational contexts. Another example is the case of the avoidant person, in which the dynamics associated with attachment needs, concerns, worries, and pain tend to occur at an intrapsychic, even unconscious, level without necessarily being expressed in interpersonal behavior.

Attachment theory has both intraindividual and interpersonal aspects; it is a prime example of "person by situation" approaches to human behavior. The person in this case is represented by the "hardwired" programming of the attachment behavioral system, the attachment working models of self and others, the procedural knowledge implicit in attachment strategies, and the associative neural networks connecting these strategies to the appraisal of person–environment transactions. The situation consists of the relationship partner's responses and other relevant contextual cues that can affect appraisal of social transactions and alter the functioning of the attachment system. The complexities in this equation stem from the fact that major parts of the "person" component were originally based on variations in the availability and responsiveness of primary caregivers in threatening situa-

tions, and major parts of the "situation" component are shaped by the person's attachment style, which may affect both the appraisal of the situation and the partner's own expressions of love, intimacy, or care. Dropping either the person or the situation component of the explanatory story results in the transformation of attachment theory into either an interdependence theory or a psychoanalytic theory.

Social Cognition Approaches and Attachment Theory

The role assigned by attachment theorists to working models of self and others in guiding attachment-system functioning is similar to the role played by schemas in the field of social cognition (e.g., Baldwin, 1992; Fiske & Taylor, 1991). Both attachment theory and social-cognitions theories emphasize the extent to which people subjectively construe person–environment transactions, store representations of typical transactions (working models in attachment theory terms; schemas, prototypes, or scripts in social-cognition language), and use these representations for understanding new transactions and organizing future action plans. In both theoretical approaches, these mental representations guide and coordinate emotion regulation, self-image, person perception, and cognitions, goals, feelings, and behavior in interpersonal settings. Furthermore, attachment theory conceptualizes working models in the same way social-cognition theorists conceptualize mental representations: They are stored in an associative memory network, maintain excitatory and inhibitory connections with other representations, and have a particular level of accessibility determined by past experiences and other factors, and this accessibility can be heightened in a given situation by relevant contextual cues (e.g., Baldwin, 1992; Collins & Read, 1994; Shaver et al., 1996).

The commonalities between attachment theory and social-cognition theories become even more evident in the analysis of the topics and methods appearing in recent adult attachment studies. As explained earlier, attachment researchers have invested a great deal of energy in assessing attachment-style variations in cognitive structures that had previously been conceptualized and examined in social-cognition research, such as person perception (e.g., Zhang & Hazan, 2002), the accessibility and organization of self-representations (e.g., Mikulincer, 1995), the accessibility of expectations about others' behavior (e.g., Baldwin et al., 1993), the accessibility of memories of social interactions (e.g., Miller & Noirot, 1999), and the way people interpret relationship partners' behavior (e.g., Collins, 1996). Moreover, adult attachment research tends to rely more and more on techniques and methods borrowed from social-cognition research, such as implicit memory tasks, semantic priming techniques, and measuring reaction times in lexical decision and Stroop color-naming tasks (e.g., Baldwin et al., 1993; Mikulincer, Gillath, et al., 2001; Mikulincer, Gillath, & Shaver, 2002).

Despite these commonalities, however, it would be a mistake to equate attachment working models with the

cognitive structures usually studied in social-cognition re- search. In their thoughtful review of the nature, content, structure, and functions of attachment working models, Shaver and colleagues (1996) enumerated four differ- ences between these constructs: As compared to other mental representations, (1) working models also tend to deal with a person's wishes, fears, conflicts, and psycho- logical defenses, and they can be affected by these psy- chodynamic processes; (2) working models seem to have a larger and more powerful affective component than most social schemas and tend to be shaped more by emotion-regulation processes; (3) working models tend to be construed in more relational terms and to organize representations of the self, others, and social interactions in a highly interdependent fashion; and (4) working mod- els are broader, richer, and more complex structures, and can include tandem or opposite representations of the same person–environment transaction at episodic, semantic, and procedural levels of encoding. Overall, at- tachment working models cannot be equated with most other social cognitions, because they evolve not only from simple memories of actual experiences but from dy- namic processes of goal pursuit, emotion regulation, and psychological defenses involved with wishes for proxim- ity and security and fears of separation and helplessness. As a result, they can distort a person's perceptions of so- cial reality, even though many are based on actual social interactions.

These differences call attention to the dialectical ten- sion between the goal-oriented and emotion-regulation functions that working models accomplish. On the one hand, due to the goal-oriented and goal-corrected nature of the attachment system, working models have to be what Bowlby (1973, p. 235) called "tolerably accurate re- flections" of what actually happened in attachment rela- tionships; otherwise, people would not be able to plan effective goal-oriented behavior and attain important re- lational goals. In this respect, working models resemble other cognitive representations that store factual knowl- edge and semantic and procedural information about re- ality constraints and demands. On the another hand, due to the emotion-regulation function of working models, they sometimes have to distort declarative and proce- dural knowledge in order to manage attachment-related fears, worries, and insecurities and protect a person from the distress and pain of attachment-figure unavailability. This dialectical tension between the goal-oriented and emotion-regulation functions of working models seems to be unique to attachment theory and differentiates it from most theories of social cognition.

Positive Psychology and Attachment Theory

The broaden-and-build cycle of attachment security calls attention to the optimistic, hopeful, constructive, and actualization-oriented tone of attachment theory, which makes it different from most other psychodynamic, rela- tional, and social-cognition theories. As already noted, people who possess a stable sense of attachment security generally feel safe and worthy, hold an optimistic and

hopeful outlook of life, rely on constructive ways of cop- ing and regulating distress, and interact with others in a confident and open manner. Moreover, they can devote mental resources that otherwise would be employed in defensive maneuvers to growth-oriented activities that contribute to the broadening of their perspectives and capacities and the actualization of their natural talents. This health- and growth-oriented theme in attachment theory has much in common with the "humanistic psychology" movement of the 1950s and 1960s (e.g., Maslow, 1968; Rogers, 1961) and today's "positive psy- chology" movement (Aspinwall & Staudinger, 2003; Seligman, 2002). Both movements are attempts to coun- terbalance psychology's traditional focus on con- flicts, fears, egoistic defenses, destructive tendencies, and psychopathology by directing attention to human strengths, potentials, and virtues that contribute to self- actualization and the development of what Rogers (1961) called a *fully functioning person*.

According to attachment theory, the sense of security is a basic human strength (Mikulincer & Shaver, 2005). It facilitates the development of other personality charac- teristics that fall under the rubric of "positive" psycholog- ical traits, such as resilience, optimism, hope, positive affectivity, curiosity and exploration, healthy autonomy, a capacity for love and forgiveness, feelings of intercon- nectedness and belongingness, tolerance, and kindness (see Lopez & Brennan, 2000; Mikulincer & Shaver, 2003, 2005; Shaver & Mikulincer, 2002). Moreover, one can easily recognize major similarities between the way the broaden-and-build cycle of attachment security evolves from repeated episodes of attachment-figure availabil- ity and ideas discussed by humanistic psychologists about the parenting style that facilitates a child's self- actualization. For example, the notion of having an available, caring, and loving attachment figure resonates with Maslow's (1968) concept of B-perception— nonjudgmental, forgiving, loving acceptance of another human being, and with Rogers's (1961) view of optimal parenting in terms of "unconditional positive regard." The common idea that recurs across the various theoreti- cal frameworks is that experiences of being loved, ac- cepted, and supported by others constitute the most im- portant form of personal protection and provide a solid psychological foundation for confronting adversity and maintaining equanimity and effective functioning in times of stress without interrupting natural processes of growth and self-actualization.

Recently, we (Mikulincer & Shaver, 2005) reviewed ex- tensive evidence showing that chronic or contextual acti- vation of the sense of attachment security attenuates the defensive motives that social psychologists tend to view as universal, such as the need for self-enhancement, needs for consensus and uniqueness, intergroup biases, defense of knowledge structures, and defense of cultural worldviews. Adult attachment studies have consistently shown that a sense of attachment security acts as a de- fault inner resource superseding defensive needs and rendering defensive maneuvers less necessary. In fact, as explained earlier, these defensive maneuvers and the re-

sulting biases in the appraisals of self, others, and social reality tend to be more characteristic of insecurely attached people. Mikulincer and Shaver (2005) noted that that these defensive needs and maneuvers seem to indicate that a person has been forced by social experiences to transact with the environment without adequate mental representations of attachment security and has had to struggle for a sense of self-worth, despite experiencing serious doubts about being lovable and possessing good inner qualities.

Despite these commonalities, there is an important difference between attachment theory and humanistic or positive psychology. Whereas the positive, humanistic approaches focus mainly on growth-oriented, promotion-focused aspects of development and personality, attachment theory emphasizes both the prevention and the promotion aspects of the attachment system. This dual focus is well illustrated in the two basic functions of "safe haven" and "secure base" served by available, responsive, caring, and loving attachment figures. These figures need to protect a person from threats and dangers; prevent the experience of negative, painful outcomes; and calm the person's fears and conflicts. At the same time, they need to provide a "secure base" from which the individual can take risks, explore the environment, and engage in promotion-oriented activities. Failure to provide either a "safe haven" or a "secure base" results in attachment-related worries and doubts as well as the development of psychological defenses that sometimes compensate for the lack of a sense of security, but at the cost of cognitive distortion, rigidity, narrowness, alienation, and an increase in interpersonal and intergroup conflict. Unlike positive psychology, attachment theory emphasizes both the "dark" and the "bright" sides of human nature and experience and explains how the attachment system deals with fears, anger, conflicts, and defenses, as well as the equally natural capacities for happiness, love, growth, and self-actualization. We believe that it makes sense to explore positive psychology within a framework that also illuminates negative psychology, because they are two natural sides of the same human coin.

CONCLUSIONS

In this chapter, we have emphasized three fundamental, interrelated principles of attachment theory that are crucial for studying and understanding the social mind. *Principle 1* concerns the adaptive, self-regulatory functions of proximity and support seeking: When a person encounters threats and dangers, whether stemming from environmental demands or internal rumination, the attachment system is activated and urgent goals become salient—to gain proximity to, and protection and comfort from, attachment figures. Beyond childhood, a person is also likely to rely on internal images of being supported by attachment figures and feelings of being comforted and supported. *Principle 2* concerns the temporally extended beneficial effects of interactions with available and responsive relationship partners and the resulting

sense of attachment security. The sense of security (or, in cases in which insufficient support is provided, the corresponding sense of insecurity) affects a person's resilience in the face of adversities and hardships, coping strategies and effectiveness, self-image, personal and social adjustment, behavior in social relationships, and personal growth. Many mental and social processes studied by personality and social psychologists working outside the attachment paradigm are affected and moderated by attachment style. *Principle 3* concerns the predictable defensive biases that arise and become established in the mind when failure of the primary attachment strategy, which is to maintain proximity to a security-providing attachment figure, results in hyperactivation or deactivation of the attachment system. According to attachment theory, these two strategies are attempts at adaptation to an inconsistently available or consistently distant or unavailable attachment figure, but once established as salient coping strategies they distort and interfere with emotion regulation, destructively color mental representations of self and others, and contribute to psychological and social problems. Because of the centrality of these principles to any understanding of the human mind, especially its social or relational aspects, the ideas and findings generated by attachment researchers tie together many of the basic concepts and findings of personality, social, developmental, and clinical psychology—whether "positive" or "negative."

Attachment theory acknowledges and integrates different, even seemingly contradictory views of human nature and maintains dialectical tension between opposites of four kinds: (1) the constraining influence of past experiences versus the forces for change in current experiences; (2) the intrapsychic nature of the attachment system, working models, and attachment strategies versus the relational, interdependent nature of attachment-related feelings, experiences, and behaviors; (3) the goal-oriented, self-regulatory function of the attachment system versus its distress-regulation, self-protective function; and (4) the importance of fears, conflicts, and prevention-focused mechanisms versus the importance of promotion-focused mechanisms and the capacity for growth and self-actualization. Given this complexity, when tied to an impressive array of research techniques and paradigms, including those borrowed from contemporary social-cognition research, attachment theory provides a unique and highly generative framework for conceptualizing and empirically exploring the full range of human constructive as well as destructive potentials. It has a remarkable 30-year record of suggesting creative, probing empirical studies that can be integrated into an expanding yet coherent scientific story. The theory shows every indication of being able to benefit from new methods, such as neuroimaging. Future theorizing and research within the attachment tradition promise to result in a comprehensive understanding of social and emotional processes, their development and incorporation into conscious and unconscious mental structures, and their amenability to education and clinical intervention.

REFERENCES

Ainsworth, M. D. S. (1967). *Infancy in Uganda: Infant care and the growth of love*. Baltimore: Johns Hopkins University Press.

Ainsworth, M. D. S. (1973). The development of infant–mother attachment. In B. M. Caldwell & H. N. Ricciuti (Eds.), *Review of child development research* (Vol. 3). Chicago: University of Chicago Press.

Ainsworth, M. D. S. (1991). Attachment and other affectional bonds across the life cycle. In C. M. Parkes, J. Stevenson-Hinde, & P. Marris (Eds.), *Attachment across the life cycle* (pp. 33–51). New York: Routledge.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: Assessed in the Strange Situation and at home*. Hillsdale, NJ: Erlbaum.

Alexander, R., Feeney, J., Hohaus, L., & Noller, P. (2001). Attachment style and coping resources as predictors of coping strategies in the transition to parenthood. *Personal Relationships*, *8*, 137–152.

Andersen, S. M., & Chen, S. (2002). The relational self: An interpersonal social-cognitive theory. *Psychological Review*, *109*, 619–645.

Aspinwall, L. G., & Staudinger, U. M. (Eds.). (2003). *A psychology of human strengths: Fundamental questions and future directions for a positive psychology*. Washington, DC: American Psychological Association.

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin*, *112*, 461–484.

Baldwin, M. W., Fehr, B., Keedian, E., & Seidel, M. (1993). An exploration of the relational schemata underlying attachment styles: Self-report and lexical decision approaches. *Personality and Social Psychology Bulletin*, *19*, 746–754.

Baldwin, M. W., Keelan, J. P. R., Fehr, B., Enns, V., & Koh Rangarajoo, E. (1996). Social-cognitive conceptualization of attachment working models: Availability and accessibility effects. *Journal of Personality and Social Psychology*, *71*, 94–109.

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology*, *61*, 226–244.

Bartholomew, K., & Shaver, P. R. (1998). Methods of assessing adult attachment: Do they converge? In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 25–45). New York: Guilford Press.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Berant, E., Mikulincer, M., & Florian, V. (2001). Attachment style and mental health: A one-year follow-up study of mothers of infants with congenital heart disease. *Personality and Social Psychology Bulletin*, *8*, 956–968.

Birnbaum, G. E., Orr, I., Mikulincer, M., & Florian, V. (1997). When marriage breaks up: Does attachment style contribute to coping and mental health? *Journal of Social and Personal Relationships*, *14*, 643–654.

Boon, S. D., & Griffin, D. W. (1996). The construction of risk in relationships: The role of framing in decisions about relationships. *Personal Relationships*, *3*, 293–306.

Bowlby, J. (1956). The growth of independence in the young child. *Royal Society of Health Journal*, *76*, 587–591.

Bowlby, J. (1969/1982). *Attachment and loss: Vol. 1. Attachment* (2nd ed.). New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation: Anxiety and anger*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and loss: Vol. 3. Sadness and depression*. New York: Basic Books.

Bowlby, J. (1988). *A secure base: Clinical applications of attachment theory*. London: Routledge.

Bradford, S. A., Feeney, J. A., & Campbell, L. (2002). Links between attachment orientations and dispositional and diary-based measures of disclosure in dating couples: A study of actor and partner effects. *Personal Relationships*, *9*, 491–506.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46–76). New York: Guilford Press.

Brennan, K. A., & Shaver, P. R. (1995). Dimensions of adult attachment, affect regulation, and romantic relationship functioning. *Personality and Social Psychology Bulletin*, *21*, 267–283.

Brennan, K. A., & Shaver, P. R. (1998). Attachment styles and personality disorders: Their connections to each other and to parental divorce, parental death, and perceptions of parental caregiving. *Journal of Personality*, *66*, 835–878.

Cannon, W. B. (1939). *The wisdom of the body* (2nd ed.). New York: Norton.

Carnelley, K. B., Pietromonaco, P. R., & Jaffe, K. (1994). Depression, working models of others, and relationship functioning. *Journal of Personality and Social Psychology*, *66*, 127–140.

Cassidy, J. (1994). Emotion regulation: Influence of attachment relationships. *Monographs of the Society for Research in Child Development*, *59*, 228–249.

Cassidy, J., & Kobak, R. R. (1988). Avoidance and its relationship with other defensive processes. In J. Belsky & T. Nezworski (Eds.), *Clinical implications of attachment* (pp. 300–323). Hillsdale, NJ: Erlbaum.

Cassidy, J., & Shaver, P. R. (Eds.). (1999). *Handbook of attachment: Theory, research, and clinical applications*. New York: Guilford Press.

Collins, N. L. (1996). Working models of attachment: Implications for explanation, emotion, and behavior. *Journal of Personality and Social Psychology*, *71*, 810–832.

Collins, N. L., & Feeney, B. C. (2000). A safe haven: An attachment theory perspective on support seeking and caregiving in intimate relationships. *Journal of Personality and Social Psychology*, *78*, 1053–1073.

Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *Journal of Personality and Social Psychology*, *58*, 644–663.

Collins, N. L., & Read, S. J. (1994). Cognitive representations of attachment: The structure and function of working models. In K. Bartholomew & D. Perlman (Eds.), *Attachment processes in adulthood* (pp. 53–92). London: Jessica Kingsley.

Cooper, M. L., Shaver, P. R., & Collins, N. L. (1998). Attachment styles, emotion regulation, and adjustment in adolescence. *Journal of Personality and Social Psychology*, *74*, 1380–1397.

Cozzarelli, C., Sumer, N., & Major, B. (1998). Mental models of attachment and coping with abortion. *Journal of Personality and Social Psychology*, *74*, 453–467.

Crowell, J. A., Fraley, R. C., & Shaver, P. R. (1999). Measurement of individual differences in adolescent and adult attachment. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 434–465). New York: Guilford Press.

Davila, J., & Cobb, R. J. (2004). Predictors of change in attachment security during adulthood. In W. S. Rholes & J. A. Simpson (Eds.), *Adult attachment: Theory, research, and clinical implications* (pp. 133–156). New York: Guilford Press.

Davila, J., Karney, B. R., & Bradbury, T. N. (1999). Attachment change processes in the early years of marriage. *Journal of Personality and Social Psychology*, *76*, 783–802.

Dozier, M., & Kobak, R. R. (1992). Psychophysiology in attachment interviews: Converging evidence for deactivating strategies. *Child Development*, *63*, 1473–1480.

Elliot, A. J., & Reis, H. T. (2003). Attachment and exploration in adulthood. *Journal of Personality and Social Psychology*, *85*, 317–331.

Feeney, J. A. (1995). Adult attachment and emotional control. *Personal Relationships*, *2*, 143–159.

Feeney, J. A. (1996). Attachment, caregiving, and marital satisfaction. *Personal Relationships*, *3*, 401–416.

Feeney, J. A. (1998). Adult attachment and relationship-centered anxiety: Responses to physical and emotional distancing. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 189–219). New York: Guilford Press.

Feeney, J. A. (1999). Adult romantic attachment and couple relationships. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 355–377). New York: Guilford Press.

Feeney, J. A. (2002). Attachment, marital interaction, and relationship satisfaction: A diary study. *Personal Relationships*, *9*, 39–55.

Feeney, J. A., & Noller, P. (1990). Attachment style as a predictor of adult romantic relationships. *Journal of Personality and Social Psychology*, *58*, 281–291.

Feeney, J. A., & Noller, P. (1991). Attachment style and verbal descriptions of romantic partners. *Journal of Social and Personal Relationships*, *8*, 187–215.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition.* New York: McGraw-Hill.

Fraley, R. C. (2002). Attachment stability from infancy to adulthood: Meta-analysis and dynamic modeling of developmental mechanisms. *Personality and Social Psychology Review*, *6*, 123–151.

Fraley, R. C., & Davis, K. E. (1997). Attachment formation and transfer in young adults' close friendships and romantic relationships. *Personal Relationships*, *4*, 131–144.

Fraley, R. C., Davis, K. E., & Shaver, P. R. (1998). Dismissing-avoidance and the defensive organization of emotion, cognition, and behavior. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 249–279). New York: Guilford Press.

Fraley, R. C., Garner, J. P., & Shaver, P. R. (2000). Adult attachment and the defensive regulation of attention and memory: Examining the role of preemptive and postemptive defensive processes. *Journal of Personality and Social Psychology*, *79*, 816–826.

Fraley, R. C., & Shaver, P. R. (1997). Adult attachment and the suppression of unwanted thoughts. *Journal of Personality and Social Psychology*, *73*, 1080–1091.

Fraley, R. C., & Shaver, P. R. (1998). Airport separations: A naturalistic study of adult attachment dynamics in separating couples. *Journal of Personality and Social Psychology*, *75*, 1198–1212.

Fraley, R. C., & Shaver, P. R. (1999). Loss and bereavement: Attachment theory and recent controversies concerning "grief work" and the nature of detachment. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 735–759). New York: Guilford Press.

Fraley, R. C., & Shaver, P. R. (2000). Adult romantic attachment: Theoretical developments, emerging controversies, and unanswered questions. *Review of General Psychology*, *4*, 132–154.

Fraley, R. C., & Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 77–114). New York: Guilford Press.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*, 218–226.

Fuendeling, J. M. (1998). Affect regulation as a stylistic process within adult attachment. *Journal of Social and Personal Relationships*, *15*, 291–322.

George, C., Kaplan, N., & Main, M. (1985). *The Adult Attachment Interview*. Unpublished manuscript, Department of Psychology, University of California, Berkeley.

Gillath, O., Shaver, P. R., Mikulincer, M., Nitzberg, R., Erez, A., & van Ijzendoorn, M. H. (2005). Attachment, caregiving, and volunteering: Placing volunteerism in an attachment-theoretical framework. *Personal Relationships*, *12*, 425–446.

Gleason, T. R. (2002). Social provisions of real and imaginary relationships in early childhood. *Developmental Psychology*, *38*, 979–992.

Green, J. D., & Campbell, W. K. (2000). Attachment and exploration in adults: Chronic and contextual accessibility. *Personality and Social Psychology Bulletin*, *26*, 452–461.

Green-Hennessy, S., & Reis, H. T. (1998). Openness in processing social information among attachment types. *Personal Relationships*, *5*, 449–466.

Harlow, H. F. (1959). Love in infant monkeys. *Scientific American*, *200*, 68–86.

Hazan, C., & Shaver, P. R. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511–524.

Hazan, C., & Shaver, P. R. (1994). Attachment theory as an organizational framework for research on close relationships. *Psychological Inquiry*, *5*, 1–22.

Hazan, C., & Zeifman, D. (1994). Sex and the psychological tether. In K. Bartholomew & D. Perlman (Eds.), *Advances in personal relationships: Vol. 5. Attachment processes in adulthood* (pp. 151–177). London: Jessica Kingsley.

Hazan, C., & Zeifman, D. (1999). Pair bonds as attachments: Evaluating the evidence. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 336–354). New York: Guilford Press.

Heinicke, C., & Westheimer, I. (1966). *Brief separations*. New York: International Universities Press.

Hesse, E. (1999). The Adult Attachment Interview: Historical and current perspectives. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 395–433). New York: Guilford Press.

Hirschberger, G., Florian, V., & Mikulincer, M. (2003). Strivings for romantic intimacy following partner complaint or partner criticism: A terror management perspective. *Journal of Social and Personal Relationships*, *20*, 675–687.

Keelan, J. P. R., Dion, K. K., & Dion, K. L. (1998). Attachment style and relationship satisfaction: Test of a self-disclosure explanation. *Canadian Journal of Behavioral Science*, *30*, 24–35.

Kirkpatrick, L. A. (1999). Attachment and religious representations and behavior. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 803–822). New York: Guilford Press.

Klass, D., Silverman, P. R., & Nickman, S. L. (Eds.). (1996). *Continuing bonds: New understandings of grief*. Washington, DC: Taylor & Francis.

Kunce, L. J., & Shaver, P. R. (1994). An attachment-theoretical approach to caregiving in romantic relationships. In K. Bartholomew & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 5, pp. 205–237). London: Jessica Kingsley.

La Guardia, J. G., Ryan, R. M., Couchman, C. E., & Deci, E. L. (2000). Within-person variation in security of attachment: A self-determination theory perspective on attachment, need fulfillment, and well-being. *Journal of Personality and Social Psychology*, *79*, 367–384.

Larose, S., Bernier, A., Soucy, N., & Duchesne, S. (1999). Attachment style dimensions, network orientation and the process of seeking help from college teachers. *Journal of Social and Personal Relationships*, *16*, 225–247.

Larose, S., Boivin, M., & Doyle, A. B. (2001). Parental representations and attachment style as predictors of support-seeking behaviors and perceptions of support in an academic counseling relationship. *Personal Relationships*, *8*, 93–113.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Levy, M. B., & Davis, K. E. (1988). Lovestyles and attachment styles compared: Their relations to each other and to various relationship characteristics. *Journal of Social and Personal Relationships*, *5*, 439–471.

Lopez, F. G., & Brennan, K. A. (2000). Dynamic processes underlying adult attachment organization: Toward an attachment theoretical perspective on the healthy and effective self. *Journal of Counseling Psychology*, *47*, 283–300.

Lussier, Y., Sabourin, S., & Turgeon, C. (1997). Coping strategies as moderators of the relationship between attachment and marital adjustment. *Journal of Social and Personal Relationships, 14*, 777–791.

Main, M. (1990). Cross-cultural studies of attachment organization: Recent studies, changing methodologies, and the concept of conditional strategies. *Human Development*, *33*, 48–61.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy,

childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development*, *50*, 66–104.

Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/ disoriented during the Ainsworth strange situation. In M. T. Greenberg, D. Cicchetti, & M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention* (pp. 121–160). Chicago: University of Chicago Press.

Maslow, A. H. (1968). *Toward a psychology of being*. New York: Van Nostrand.

Mayseless, O., Danieli, R., & Sharabany, R. (1996). Adults' attachment patterns: Coping with separations. *Journal of Youth and Adolescence*, *25*, 667–690.

Mickelson, K. D., Kessler, R. C., & Shaver, P. R. (1997). Adult attachment in a nationally representative sample. *Journal of Personality and Social Psychology*, *73*, 1092–1106.

Mikulincer, M. (1995). Attachment style and the mental representation of the self. *Journal of Personality and Social Psychology*, *69*, 1203–1215.

Mikulincer, M. (1997). Adult attachment style and information processing: Individual differences in curiosity and cognitive closure. *Journal of Personality and Social Psychology*, *72*, 1217–1230.

Mikulincer, M. (1998a). Adult attachment style and affect regulation: Strategic variations in self-appraisals. *Journal of Personality and Social Psychology*, *75*, 420–435.

Mikulincer, M. (1998b). Adult attachment style and individual differences in functional versus dysfunctional experiences of anger. *Journal of Personality and Social Psychology*, *74*, 513–524.

Mikulincer, M. (1998c). Attachment working models and the sense of trust: An exploration of interaction goals and affect regulation. *Journal of Personality and Social Psychology*, *74*, 1209–1224.

Mikulincer, M., & Arad, D. (1999). Attachment working models and cognitive openness in close relationships: A test of chronic and temporary accessibility effects. *Journal of Personality and Social Psychology*, *77*, 710–725.

Mikulincer, M., Birnbaum, G., Woddis, D., & Nachmias, O. (2000). Stress and accessibility of proximity-related thoughts: Exploring the normative and intraindividual components of attachment theory. *Journal of Personality and Social Psychology*, *78*, 509–523.

Mikulincer, M., Dolev, T., & Shaver, P. R. (2004). Attachment-related strategies during thought-suppression: Ironic rebounds and vulnerable self-representations. *Journal of Personality and Social Psychology*, *87*, 940–956.

Mikulincer, M., & Florian, V. (1995). Appraisal of and coping with a real-life stressful situation: The contribution of attachment styles. *Personality and Social Psychology Bulletin*, *21*, 406–414.

Mikulincer, M., & Florian, V. (1998). The relationship between adult attachment styles and emotional and cognitive reactions to stressful events. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 143–165). New York: Guilford Press.

Mikulincer, M., & Florian, V. (2000). Exploring individual differences in reactions to mortality salience: Does attachment style regulate terror management mechanisms? *Journal of Personality and Social Psychology*, *79*, 260–273.

Mikulincer, M., & Florian, V. (2001). Attachment style and affect regulation: Implications for coping with stress and mental health. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 537–557). Oxford, UK: Blackwell.

Mikulincer, M., Florian, V., Cowan, P. A., & Cowan, C. P. (2002). Attachment security in couple relationships: A systemic model and its implications for family dynamics. *Family Process*, *41*, 405–434.

Mikulincer, M., Florian, V., & Hirschberger, G. (2003). The existential function of close relationships: Introducing death into the science of love. *Personality and Social Psychology Review*, *7*, 20–40.

Mikulincer, M., Florian, V., & Tolmacz, R. (1990). Attachment styles and fear of personal death: A case study of affect regulation. *Journal of Personality and Social Psychology*, *58*, 273–280.

Mikulincer, M., Florian, V., & Weller, A. (1993). Attachment styles, coping strategies, and posttraumatic psychological distress: The impact of the Gulf War in Israel. *Journal of Personality and Social Psychology*, *64*, 817–826.

Mikulincer, M., Gillath, O., Halevy, V., Avihou, N., Avidan, S., & Eshkoli, N. (2001). Attachment theory and reactions to others' needs: Evidence that activation of the sense of attachment security promotes empathic responses. *Journal of Personality and Social Psychology*, *81*, 1205–1224.

Mikulincer, M., Gillath, O., Sapir-Lavid, Y., Yaakobi, E., Arias, K., Tal-Aloni, L., et al. (2003). Attachment theory and concern for others' welfare: Evidence that activation of the sense of secure base promotes endorsement of self-transcendence values. *Basic and Applied Social Psychology*, *25*, 299–312.

Mikulincer, M., Gillath, O., & Shaver, P. R. (2002). Activation of the attachment system in adulthood: Threat-related primes increase the accessibility of mental representations of attachment figures. *Journal of Personality and Social Psychology*, *83*, 881–895.

Mikulincer, M., Hirschberger, G., Nachmias, O., & Gillath, O. (2001). The affective component of the secure base schema: Affective priming with representations of attachment security. *Journal of Personality and Social Psychology*, *81*, 305–321.

Mikulincer, M., & Horesh, N. (1999). Adult attachment style and the perception of others: The role of projective mechanisms. *Journal of Personality and Social Psychology*, *76*, 1022–1034.

Mikulincer, M., Horesh, N., Eilati, I., & Kotler, M. (1999). The association between adult attachment style and mental health in extreme life-endangering conditions. *Personality and Individual Differences*, *27*, 831–842.

Mikulincer, M., & Nachshon, O. (1991). Attachment styles and patterns of self-disclosure. *Journal of Personality and Social Psychology*, *61*, 321–331.

Mikulincer, M., & Orbach, I. (1995). Attachment styles and repressive defensiveness: The accessibility and architecture of affective memories. *Journal of Personality and Social Psychology*, *68*, 917–925.

Mikulincer, M., Orbach, I., & Iavnieli, D. (1998). Adult attachment style and affect regulation: Strategic variations in subjective self-other similarity. *Journal of Personality and Social Psychology*, *75*, 436–448.

Mikulincer, M., & Shaver, P. R. (2001). Attachment theory and intergroup bias: Evidence that priming the secure base schema attenuates negative reactions to out-groups. *Journal of Personality and Social Psychology*, *81*, 97–115.

Mikulincer, M., & Shaver, P. R. (2003). The attachment behavioral system in adulthood: Activation, psychodynamics, and interpersonal processes. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 53–152). New York: Academic Press.

Mikulincer, M., & Shaver, P. R. (2004). Security-based self-representations in adulthood: Contents and processes. In W. S. Rholes & J. A. Simpson (Eds.), *Adult attachment: Theory, research, and clinical implications* (pp. 159–195). New York: Guilford Press.

Mikulincer, M., & Shaver, P. R. (2005). Mental representations of attachment security: Theoretical foundation for a positive social psychology. In M. W. Baldwin (Ed.), *Interpersonal cognition* (pp. 233–266). New York: Guilford Press.

Mikulincer, M., Shaver, P. R., & Horesh, N. (2006). Attachment bases of emotion regulation and posttraumatic adjustment. In D. K. Snyder, J. A. Simpson, & J. N. Hughes (Eds.), *Emotion regulation in families: Pathways to dysfunction and health* (pp. 77–90). Washington, DC: American Psychological Association.

Mikulincer, M., & Sheffi, E. (2000). Adult attachment style and cognitive reactions to positive affect: A test of mental categorization and creative problem solving. *Motivation and Emotion*, *24*, 149–174.

Miller, J. B., & Noirot, M. (1999). Attachment memories, models and information processing. *Journal of Social and Personal Relationships*, *16*, 147–173.

Ochsner, K. N., & Gross, J. J. (2004). Thinking makes it so: A social-

cognitive neuroscience approach to emotion regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 229–255). New York: Guilford Press.

O'Connor, T. G., & Croft, C. M. (2001). A twin study of attachment in preschool children. *Child Development*, *72*, 1501–1511.

Ognibene, T. C., & Collins, N. L. (1998). Adult attachment styles, perceived social support, and coping strategies. *Journal of Social and Personal Relationships*, *15*, 323–345.

Overall, N. C., Fletcher, G. J. O., & Friesen, M. (2003). Mapping the intimate relationship mind: Comparisons between three models of attachment representations. *Personality and Social Psychology Bulletin*, *29*, 1479–1493.

Pietromonaco, P. R., & Feldman Barrett, L. (1997). Working models of attachment and daily social interactions. *Journal of Personality and Social Psychology*, *73*, 1409–1423.

Roberts, J. E., Gotlib, I. H., & Kassel, J. D. (1996). Adult attachment security and symptoms of depression: The mediating roles of dysfunctional attitudes and low self-esteem. *Journal of Personality and Social Psychology*, *70*, 310–320.

Rogers, C. R. (1961). *On becoming a person*. Boston: Houghton Mifflin.

Rom, E., & Mikulincer, M. (2003). Attachment theory and group processes: The association between attachment style and group-related representations, goals, memory, and functioning. *Journal of Personality and Social Psychology*, *84*, 1220–1235.

Rothbaum, F., Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. *Journal of Personality and Social Psychology*, *42*, 5–37.

Schachter, S. (1959). *The psychology of affiliation*. Stanford, CA: Stanford University Press.

Scharf, M. (2001). A "natural experiment" in childrearing ecologies and adolescents' attachment and separation representations. *Child Development*, *72*, 236–251.

Scharfe, E., & Bartholomew, K. (1995). Accommodation and attachment representations in couples. *Journal of Social and Personal Relationships*, *12*, 389–401.

Searle, B., & Meara, N. M. (1999). Affective dimensions of attachment styles: Exploring self-reported attachment style, gender, and emotional experience among college students. *Journal of Counseling Psychology*, *46*, 147–158.

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York: Free Press.

Shaver, P. R., Belsky, J., & Brennan, K. A. (2000). The adult attachment interview and self-reports of romantic attachment: Associations across domains and methods. *Personal Relationships*, *7*, 25–43.

Shaver, P. R., & Brennan, K. A. (1992). Attachment styles and the "big five" personality traits: Their connections with each other and with romantic relationship outcomes. *Personality and Social Psychology Bulletin*, *18*, 536–545.

Shaver, P. R., & Clark, C. L. (1994). The psychodynamics of adult romantic attachment. In J. M. Masling & R. F. Bornstein (Eds.), *Empirical perspectives on object relations theory* (Vol. 5, pp. 105–156). Washington, DC: American Psychological Association.

Shaver, P. R., Collins, N., & Clark, C. L. (1996). Attachment styles and internal working models of self and relationship partners. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 25–61). Hillsdale, NJ: Erlbaum.

Shaver, P. R., & Hazan, C. (1993). Adult romantic attachment: Theory and evidence. In D. Perlman & W. Jones (Eds.), *Advances in personal relationships* (Vol. 4, pp. 29–70). London: Jessica Kingsley.

Shaver, P. R., & Mikulincer, M. (2002). Attachment-related psychodynamics. *Attachment and Human Development*, *4*, 133–161.

Shaver, P. R., & Mikulincer, M. (2004). Attachment in the later years: A commentary. *Attachment and Human Development*, *6*, 451–464.

Simpson, J. A. (1990). Influence of attachment styles on romantic relationships. *Journal of Personality and Social Psychology*, *59*, 871–980.

Simpson, J. A., & Kenrick, D. T. (Eds.). (1997). *Evolutionary social psychology*. Hillsdale, NJ: Erlbaum.

Simpson, J. A., Rholes, W. S., Campbell, L., & Wilson, C. L. (2003). Changes in attachment orientations across the transition to parenthood. *Journal of Experimental Social Psychology*, *39*, 317–331.

Simpson, J. A., Rholes, W. S., & Nelligan, J. S. (1992). Support seeking and support giving within couples in an anxiety-provoking situation: The role of attachment styles. *Journal of Personality and Social Psychology*, *62*, 434–446.

Simpson, J. A., Rholes, W. S., & Phillips, D. (1996). Conflict in close relationships: An attachment perspective. *Journal of Personality and Social Psychology*, *71*, 899–914.

Sroufe, L. A., & Waters, E. (1977). Attachment as an organizational construct. *Child Development*, *48*, 1184–1199.

Taubman Ben-Ari, O., Findler, L., & Mikulincer, M. (2002). The effects of mortality salience on relationship strivings and beliefs: The moderating role of attachment style. *British Journal of Social Psychology*, *41*, 419–441.

Thibault, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tidwell, M. C. O., Reis, H. T., & Shaver, P. R. (1996). Attachment, attractiveness, and social interaction: A diary study. *Journal of Personality and Social Psychology*, *71*, 729–745.

Waters, E., Merrick, S., Treboux, D., Crowell, J., & Albersheim, L. (2000). Attachment security in infancy and early adulthood: A twenty-year longitudinal study. *Child Development*, *71*, 684–689.

Waters, H. S., Rodrigues, L. M., & Ridgeway, D. (1998). Cognitive underpinnings of narrative attachment assessment. *Journal of Experimental Child Psychology*, *71*, 211–234.

Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology*, *63*, 903–912.

Wegner, D. M., & Smart, L. (1997). Deep cognitive activation: A new approach to the unconscious. *Journal of Consulting and Clinical Psychology*, *65*, 984–995.

Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin*, *124*, 252–283.

Westmaas, J. L., & Silver, R. C. (2001). The role of attachment in responses to victims of life crises. *Journal of Personality and Social Psychology*, *80*, 425–438.

Wisman, A., & Koole, S. L. (2003). Hiding in the crowd: Can mortality salience promote affiliation with others who oppose one's worldview. *Journal of Personality and Social Psychology*, *84*, 511–527.

Zhang, F., & Hazan, C. (2002). Working models of attachment and person perception processes. *Personal-Relationships*, *9*, 225–235.

Zimmerman, P. (1999). Structure and functions of internal working models of attachment and their role for emotion regulation. *Attachment and Human Development*, *1*, 291–306.

Zuroff, D. C., & Fitzpatrick, D. K. (1995). Depressive personality styles: Implications for adult attachment. *Personality and Individual Differences*, *18*, 253–365.

Social Power

SUSAN T. FISKE
JENNIFER BERDAHL

Basic principles of social power all follow from power as control over valued resources. People who have what they want—and what other people want—create social forces. Powerholders can be forces for ill or good, reflecting the tension between unconstrained greed and unconstrained generosity. With the unconstrained freedom to do what they want, people's behavior is tempered by their self-interest, ignorance, and exploitative intent, on the one hand, or by their felt responsibility for the welfare of others, on the other hand. We argue that the person, the position, and the situation combine to predict social power's effects for harmful or beneficial behavior.

DEFINING POWER

To analyze power we must define it, not an easy task. Arguably, the concept of power runs so deep that searching for a simple definition is a mistake (Lukes, 1986; Ng, 1980). Many consider power to be the most basic force behind human behavior (e.g., Cartwright, 1959; Russell, 1938), so no wonder power concerns disciplines spanning the social sciences and humanities (i.e., philosophy, psychology, sociology, economics, history, political science, and anthropology). Indeed, the construct of power is so broad that "some students of the subject think that the whole study of 'power' is a 'bottomless swamp' " (Dahl, 1957, p. 201). Despite this morass, the need to be explicit has led to several definitions,[1] which we can group into three categories: (1) power as influence, (2) power as potential influence, and (3) power as outcome control (see Figure 29.1).

Power as Influence

Definitions of *power as influence* define power by its effect (see right side of Figure 29.1). When one person causes, or influences, another to behave a certain way, the former has power over the latter. For example, Simon (1957) defined "A" as having power over "B" when "A's behavior causes B's behavior" (p. 5). According to Dahl (1957), "A has power over B to the extent that [A] can get B to do something that B would not otherwise do" (p. 202). Russell (1938) claimed that "A has more power than B, if A achieves many intended effects and B only a few" (p. 35). Most of the power literature in fact addresses social influence: Compliance, identification, and internalization were processes defined by French and Raven (1959), as well as by Kelman (1958). The social influence literature focuses on the strategies that change behavior as a result of interpersonal interaction (Chaiken, Wood, & Eagly, 1996; Cialdini & Trost, 1998; Turner, 1995). All these focus on the target's response to influence attempts.

We believe defining power in terms of influence is problematic because it defines power in terms of what it *does*, not what in terms of what it *is*. This is a we-know-it-when-we-see-it way of defining power that requires back-

Outcome Control (source of power)	→	Potential for Influence (capacity for effect)	→	Influence (effect of power)

Thibaut & Kelley (1959, p. 101): "The power of A over B increases with A's ability to affect the quality of outcomes attained by B."

Emerson (1962, p. 32): "The power to control or influence the other resides in control over the things he values."

Kipnis (1976, p. x): "access to resources needed by others."

Dépret & Fiske (1993): "asymmetrical control over another person's outcomes."

Max Weber (1914/1978, p. 152): "the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance."

French & Raven (1959, p. 152): "The strength of power of O/P in some system a is defined as the maximum potential ability of O to influence P in a."

Cartwright (1965, p. 4): "If O has the capability of influencing P, we say that O has power over P."

Russell (1938, p. 35): "A has more power than B, if A achieves many intended effects and B only a few."

Simon (1957, p. 5): "For the assertion, 'A has power over B,' we can substitute the assertion, 'A's behavior causes B's behavior.' "

Dahl (1957, p. 202): "A has power over B to the extent that he can get B to do something that B would not otherwise do."

FIGURE 29.1. Definitions of power placed within a causal continuum from source to effect of power.

ward inference to identify the presence of power by its consequence. This definition does not address what it is about power that might lead to influence in the first place.

Power as Potential Influence

Others have defined *power as potential influence*. For example, Weber (1914/1978) defined power as "the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance" (p. 152). Lewin (1944, 1951/1997) described power as the possibility of inducing forces. In one of the most cited papers on power to date, French and Raven (1959) based their "model of influence" on Lewin's conceptions of psychological force fields (Raven, 1993) and defined power this way: "The *strength of power* of O/P in some system *a* is defined as the maximum potential ability of O to influence P in *a*" (p. 152). Cartwright (1965) echoed with, "If O has the capability of influencing P, we say that O has power over P" (p. 4). According to these definitions, power can exist without *actual* influence.

The definitions themselves do not explain the origins of capacity to influence. Many who define power as potential influence view this capacity as stemming from control over valued resources. Cartwright (1965), for example, argued that influence derives from "the possession, or control, of valued resources" (p. 5). Dahl (1957) defined the base of power as "all the resources—opportunities, acts, objects, etc.—that [A] can exploit in order to affect the behavior of [B]" (p. 203). More recently, Manz and Gioia (1983) defined power as "the ability or potential to influence others" (p. 461) through resource control, and Vescio, Snyder, and Gervais (in press) define it as "the potential to influence others in meaningful ways . . . through the awarding and withholding of resources or through punishment." In short, these approaches all define A as having potential influence over B (i.e., power) *because* A controls resources of value to B.

Power as Outcome Control

Others have dropped influence from the definition of power altogether and have homed in on *power as resource or outcome control* (e.g., Dépret & Fiske, 1993; Emerson, 1962; Keltner, Gruenfeld, & Anderson, 2003; Kipnis, 1976; Thibaut & Kelley, 1959). Early advocates of this approach in social psychology include Thibaut and Kelley (1959), who defined power as the ability to affect another's reward–cost position, and Emerson (1962), who claimed "the power to control or influence the other resides in control over the things he values" (p. 32). More recently, Dépret and Fiske (1993) defined power as *asymmetrical control over another person's outcomes* (see also Fiske, 1993), and Keltner and colleagues (2003) defined it as "an individual's relative capacity to modify others' states by providing or withholding resources or administering punishments" (p. 265). Rather than identifying power by its consequences, these definitions identify it in the structural properties of social relations. This is the approach we adopt. We define power as *relative control over another's valued outcomes*.[2]

We believe influence effects need to be separated from the control of outcomes per se, whether or not the influence attempt is successful or even intended in the first place. People who control others' outcomes have power, like it or not. They may not even intend or want to use their power, though the control of valued outcomes gives them the potential. Making influence part of the definition of power presupposes target volition (Raven, 1993). It conflates a structural feature, the powerholder's control over valued outcomes, with a psychological factor, the target's choice to cooperate with an influence attempt (Dépret & Fiske, 1993; Fiske & Dépret, 1996). In that view, power is ceded by the powerless, which seems ironic. Control over valued outcomes still constitutes power, even if the target chooses to resist. That is, the target may care about the outcomes, even while refusing to enact the powerholder's conditions for obtaining the outcomes desired. *Power over* someone persists, even without the *power to* make the person comply.

**Implications of Defining Power
as Relative Outcome Control**

Defining power as relative control over another's valued outcomes has implications for how power is conceptualized and measured as well as the different types of power possible.

Power Is Relative

As the relative control over another's valued outcomes, power is always socially situated, hence our title, "Social Power." *Relative* means that power is an inherently comparative construct, defined within a particular social relationship. Power is not an "attribute" of an individual but a structural property of a social relation that derives from relative control over outcomes. That is, one cannot say that someone has power without specifying over *whom* (Arendt, 1960; Emerson, 1962; French & Raven, 1959; Wartenberg, 1990). A person can be powerful in one context and powerless in another. Some individuals may be more likely than others to obtain power in social situations when it is up for grabs; some may be used to having power and conditioned to experience and view the world from that perspective. In a given social relationship, however, these people may be interacting with another who is even more likely to obtain power or to experience it in daily life. Therefore, researchers must measure the relative power between individuals in a social situation, particularly their relative control over each other's outcomes, and not treat power as an individual-level construct or characteristic. Later, we distinguish social power from having a dominant personality and related constructs.

Control Varies

The relative *control* a powerholder has over another's outcomes stems from the powerholder's ability to administer or withhold rewards or punishments (Dépret & Fiske, 1993; Emerson, 1962; Keltner et al., 2003; Kipnis, 1976; Thibaut & Kelley, 1959). This control can vary in how formal, stable, and legitimate it is. Formal control is accompanied by an official title or explicit role (e.g., "boss"), whereas informal control is designated through social roles and status (e.g., typically, rich people are dominant, and poor people are subordinate). Control can range from being stable, not easily lost or removed (e.g., a physical advantage) to being unstable, easily usurped (e.g., an elected position). When control is unstable, power differences between two people (or parties) is likely to vary over time and is a state, rather than a trait, of the relationship. Finally, the control over another's valued outcomes may be considered legitimate (e.g., based on merit) or illegitimate (e.g., based on luck). The nature of a powerholder's control over another's outcomes affects the use and consequences of power. For example, power should be greater, and influence more likely, when control is stable and legitimate. A powerholder should be more likely to use "harsh" tactics of influence (Raven,

1992, 1993) when control is formal, stable, and illegitimate but "soft" tactics of influence when it is informal, unstable, and legitimate (Kipnis, 1972).

Outcomes Vary

As the relative control over another's *valued outcomes*, having power depends on the ability to control outcomes that others desire, in other words, outcomes that affect others' well-being. Food, safety, material goods, social acceptance, and the well-being of loved ones are but a few examples. Given the number of valued outcomes that someone else can control, classifying outcomes into a broad framework is useful. Primary categories certainly include physical, economic, and social outcomes.

   Physical outcomes define a person's health and safety, such as food, shelter, physical harm, and physical pleasure. *Economic outcomes* define a person's material well-being. Money and tangible goods are obvious examples, but evaluations, opportunities, and formal positions that determine economic rewards also fit into this category.

   Social outcomes define a person's societal and interpersonal well-being. These outcomes include at the core (1) social belonging (acceptance), arguably essential to people's survival in both ancestral and current times (Fiske, 2004). Belonging is intensely motivating (Baumeister & Leary, 1995), and it has even been defined as the core basis for power (Vescio, et al., in press). Other social outcomes that facilitate belonging all acquire value as means or measures of it. For example, another core outcome, (2) socially shared understanding, provides information about norms, which enable people to belong; providing or withholding normative information can be a source of power. Likewise, providing or withholding (3) a sense of social control is a crucial outcome; a powerholder can be predictable and responsive, encouraging people's own effectiveness, or not. Senses of understanding and control are social cognitive outcomes that powerholders can facilitate or impede. Less cognitive, more social–affective outcomes include facilitating or impeding another's (4) self-enhancement (via support, respect, and affection) and (5) social trust (creating benign, reliable ingroup expectations). These social outcomes—belonging, understanding, controlling, enhancing, and trusting—together constitute strong social incentives held by the powerful.

BASES OF POWER

All three types of outcome—physical, economic, and social—affect the well-being of individuals and are therefore likely to be valued results that, when controlled by someone else, give the controller power over the individual whose outcomes are controlled. In this way, physical, economic, and social outcomes form different bases of power. In the single most influential view of power to date, French and Raven (1959) outlined six different bases of power. Their bases of power resulted from analyzing the kinds of resources a person might use to exer-

cise influence: rewards (e.g., support, benefits, favors), coercion (annoyance, abuse, making trouble), legitimacy (socially accepted obedience, duty to conform, right to order), expertise (reputation for knowledge, intelligence, judgment), reference (ideal, model, aspiration), and information (persuasion, logic). According to our definition of power as outcome control, French and Raven's reward and coercive power *is* power, period. Their other bases of power describe the variety of resources a powerholder can wield to try to influence another, consistent with their definition of power as potential influence. In our view, legitimacy, expertise, reference, and information power are bases of power insofar as they yield reward or coercive power—in other words, control over another's outcomes.

French and Raven's bases of power have provided useful insight into different types of power. Global measures of perceived power relate to these bases of power and more (Nesler, Aguinis, Quigley, Lee, & Tedeschi, 1999). A later revision, the interpersonal power interaction model (IPIM; Raven, 1993, 2001) expanded these bases by subdividing sanctions (into impersonal, e.g., tangible, vs. personal, e.g., approval), legitimacy (into formal position, reciprocity, equity, and responsibility), expertise and reference (both into positive and negative), and information (into direct and indirect, e.g., overheard). This provides a detailed framework of the resources one might leverage to influence others.

The IPIM examines both target and agent. IPIM takes a cost–benefit approach to strategy choice (Koslowsky & Schwarzwald, 2001). Consistent with this approach, it notes that all the original French–Raven bases of power involve volition on the part of the target (Raven, 1993). The target theoretically can decide whether or not to cooperate, depending on the relative costs and benefits. The IPIM adds two new bases of power—force and manipulation—that do not fit their earlier volitional view. Brute force ignores the target's volition, by definition. And manipulation of the target's environment to force or prevent behavior likewise bypasses target volition, unless the target is able to exit the situation, knows it, and chooses to do so. Including target volition in the conception of power is problematic because targets may not have or perceive better alternatives. When a target refuses to go along, theorists have taken it as evidence of reduced power differential, when in fact someone controlling the target's outcomes has power over the target regardless of the target's choice.

Harsh and Soft Power

A series of scales has emerged to measure Raven's expanded list of perceived bases of power. The Interpersonal [Bases of] Power Inventory (IPI; Raven, Schwarzwald, & Koslowsky, 1998) validates 7 factors of the revised 11 bases from the subordinate's perspective. A further factor analysis of these 7 scales demonstrates a two-factor solution, identifying (1) *soft or weak bases*: use of credibility (expert and information power), referent and personal reward power, or legitimate power (de-

pendence) to influence a target through the trust, respect, and liking associated with these bases of power; and (2) *harsh or strong* bases: use of sanctions (impersonal rewards and punishments and personal punishments) and legitimate power (equity, reciprocity, and position) to influence a target with the mostly economic rewards associated with these bases of power (Raven et al., 1998).

Another way to view the distinction between soft and harsh forms of power is that soft power uses social outcomes (e.g., liking and respect) and harsh power uses economic and physical ones to attempt to influence a target. Outcomes vary in how concrete they are, with implications for ease of analysis by both lay and scientific observers. Physical, economic, and social outcomes together array on a continuum from tangible to intangible, or objective to subjective, with physical outcomes on one end (objective/tangible) and social outcomes on the other (subjective/intangible). Power differentials should be more obvious to the degree that they reflect objective, tangible outcomes, but differentials should be less obvious to the degree that they reflect subjective, intangible outcomes. Thus, the use of power is easy to recognize when it entails tangible outcomes (e.g., an economic bribe or a physical threat), and may more likely be considered "harsh," than when it entails intangible outcomes (e.g., liking and respect), more likely considered "soft."

The terms "harsh" and "soft," however, should not be confused with how threatening and effective these different forms of power are. The use of physical power is not necessarily more threatening or effective than the use of economic power and the use of economic power is not necessarily more threatening or effective than the use of social power. What matters is value. The strength (i.e., threat) of power depends on the value of the outcome controlled. Because social belonging is a primary value of critical importance to most people, using this outcome to wield power is likely to be quite threatening and effective, even if the exercise of this power appears "soft" or is difficult to recognize as a coercive tactic.

Raven and colleagues' (1998) "harsh" and "soft" power distinction fits other analyses (Bass, 1981; Elias & Loomis, 2004; Erchul, Raven, & Ray, 2001; Kipnis, 1984; Yukl & Falbe, 1991). Soft, relational power works well. Relational power predicts subordinate job satisfaction and compliance, consistent with the leadership literature (Bass, 1956; Erchul, Raven, & Whichard, 2001; Gerster & Day, 1997; Imai, 1991, 1994; Koslowsky, Schwarzwald, & Ashuri, 2001; Lowe, Kroeck, & Sivasubramaniam, 1996; Mullen, 1991; Raven et al., 1998; Wofford & Liska, 1993). In one setting, reference (relational) power mediated the effects of all other kinds of power on job performance (Rahim, Antonioni, & Psenicka, 2001). Women tend to show relational types of leadership (i.e., transformational and contingent transactional), rather than management by laissez-faire or by exception (Eagly, Johannesen-Schmidt, & van Engen, 2003). The relational style tends to work well for both genders when they use it. Note, however, that the contrast here falls between relational as respect, liking, and task-contingent rewards, on the one hand, and autocratic as sheer legitimacy with sporadic

punishment, on the other hand. Thus the distinctions overlap but are not identical.

In a praiseworthy variety of applied research settings, harsh and soft power demonstrably vary in effectiveness, but the pattern tends to favor soft power: Relating instructor power bases and undergraduate compliance, soft bases of power were especially effective for African American men (Elias & Loomis, 2004). In school consultants, teachers, and psychologists favored soft power (standard forms of information, expert, legitimate, and reference power; Erchul, Raven, & Ray, 2001; Erchul, Raven, & Wilson, 2004). In hypothetical doctor–patient scenarios, soft power again is rated most effective (information, referent, and expert power; Brown & Raven, 1994). Relating faculty power bases and graduate student outcomes, expert power (soft) and reward power (can be either) were especially effective (Aguinis, Nesler, Quigley, Lee, & Tedeschi, 1996). This taxonomy also applies to case studies analyzing historic political encounters (Gold & Raven, 1992; Raven, 1990, 2001) and mechanisms of religious influence (Raven, 1999, 2001). In numerous organizational settings, as just noted, soft power predicts subordinate job satisfaction, and expert power specifically (a component of soft power) seems to influence effectively everywhere.

Some studies do not use French and Raven's taxonomy but still examine influence tactics that could be mapped onto soft and harsh. Managerial preference across cultures (U.S. and Latin American) for conflict resolution tactics rank advising and mediation (soft) first in the United States, Argentina, and Mexico, and autocratic arbitration ("inquisitorial," harsh) second (although those ranks reversed in Dominican Republic); all rejected avoidance/ignoring and mostly rejected providing impetus/incentives; an adversarial/arbitration process was seen neutrally (Cropanzano, Aguinis, Schminke, & Denham, 1999).

Across vignettes varying jobs, roles, and goals, student-rated consensus ranked rationality, ingratiation, assertiveness, and exchange, in that order; note that the first two are softer than the last two (Aguinis, Nesler, Hosoda, & Tedeschi, 1994). Manager role-congruent strategy (direct influence) may also be favored (Aguinis & Adams, 1998); this could be harsh or soft. Note, however, that some studies merely ask people what they prefer, creating issues of demand, self-presentation, and lack of introspective access. Other studies ask about subordinate satisfaction or examine measures of effectiveness, which probably better index the different exercises of power.

One of the founders of work on the social psychology of power, David Kipnis (2001), just before his death, contrasted strong, controlling (harsh) tactics with rational and weak (soft) tactics. His metamorphic theory of power (Kipnis, 1976, 1984) notes that when the powerholder controls someone overtly, both of them attribute the target's actions to force; the powerholder then demeans the subordinate for being weak, and the subordinate resents the powerholder. Also, the desired behavior ceases as soon as surveillance ceases. When powerholders use rational (expert), relational tactics, both parties perceive

choice, each respects the other, the subordinate has self-respect, and the desired behavior is self-sustaining.

POWER AND STATUS

Much informal debate in the field contests whether power and status, often used interchangeably, are in fact distinct constructs. In our view, both are distinct structural positions. Status is a structural position (either formal or informal) involving relative respect and esteem, and it can be a form of power (outcome control). Status relates to power in that it yields control over social outcomes of value to others, namely, liking and respect. That is, people generally want to be liked and respected by those who are relatively high in status. As such, high-status individuals often have social power over others. Although both are structural positions, status and power are distinguishable, because only power necessitates outcome control. One may have high status and low outcome control, as in a lame-duck administrator. Or, one may control resources but have a position that lacks the respect, as in a low-status gatekeeper. Status entails formal position (legitimacy, in French & Raven's terms), but outcome control is separable.

An alternative taxonomy of power (Bass, 1960) provides a different way to think about status by focusing on personal and positional power, at a more abstract level than the French–Raven bases of influence (Yukl & Falbe, 1991). *Positional power* includes legitimacy and sanctions (reward, coercion), whereas *personal power* includes referent (relational) and expert power, which can be independent of one's position in an organization. *Political power* adds coalition building and decision control (Yukl, 1989). Because social status correlates with various kinds of outcome control (economic, social, and physical), observers therefore view it as a form of power.

ANTECEDENTS OF POWER

Demographic Antecedents

Whence comes power? Expectation states theory argues that certain characteristics confer status and afford power (Berger, Cohen, & Zelditch, 1972). For example, when they barely know each other, juries are more likely to pick as foreperson men over women, older people over younger ones, and whites over people of color (e.g., for gender, see Beckham & Aronson, 1978). In effect, certain fixed characteristics are treated as if they are resources, which then contribute to people in fact controlling resources.

Similarly, social role theory (Eagly, 1987) also argues that certain demographic groups are more likely than others to land in certain social roles, leading observers to attribute stereotypic personality characteristics (e.g., leadership qualities) to group members that may result more from their most frequent roles than from the individuals' own dispositions. Thus, women and men disproportionately and respectively inhabit caretaking and

breadwinning roles, so people infer that women are communal nurturers and men are agentic leaders.

The result is that demographic characteristics predict who is seen as powerful. Women are implicitly associated with egalitarian structures and men with hierarchical ones, especially by male perceivers (Schmid Mast, 2004). Hierarchical associates include *rank*, *status*, and *hierarchy*, whereas egalitarian associates include *equality*, *egalitarian*, and *similar*. Note that all terms reflect status (social position) not simply power (resource control). These gendered associations doubtless reflect a mix of expectations and self-selection. That is, women are devalued in traditionally male domains (for a meta-analysis, see Eagly, Makhijani, & Klonsky, 1992). But women are also less motivated to lead in hierarchically organized settings (for a meta-analysis, see Eagly, Karau, Miner, & Johnson, 1994) and use a less autocratic or directive style when they do lead (for a meta-analysis, see Eagly & Johnson, 1990). Consistent with these gender differences in hierarchy, all-male groups are more hierarchical than all-female groups, as indicated by speaking time (Schmid Mast, 2002a), yielding to interruptions (Schmid Mast, 2002b), and patterns of leadership centralization (Berdahl & Anderson, 2005). Also consistent are gender differences in social dominance orientation (SDO); men on average score higher on beliefs that group hierarchies are inevitable and good (Pratto, Stallworth, & Sidanius, 1997) and are less likely to prefer egalitarian norms in their ideal company or team (Berdahl & Anderson, 2005). Hierarchies differentiate on the basis of status, most often yielding power to those at the top.

Traditionally male domains also favor perceived male power. Perceived coercive, expert, and informational power all favor men, as rated by men (Elias, 2004). Men link bodily force to power, whereas women link bodily force to loss of power. Making a fist activates power concepts for both genders, but for men it also activates hope for power and favoring an assertive target, with the opposite results for women (Schubert, 2004). All these patterns correlate with the roles that men and women each disproportionately inhabit, but of course the causality can go from roles to associations, from associations to roles, and from third variables (e.g., testosterone) to both.

Although gender has been most studied, other fixed indicators of status include height (Wilson, 1968) and age (Berger et al., 1972; Mazur, 1985). Physical attractiveness predicts men's actual social status in groups (where status is defined as prominence, respect, and influence; Anderson, John, Keltner, & Kring, 2001). But baby-faced individuals come across as weak (Zebrowitz-McArthur & Montepare, 1989), and large eyes signal low dominance (Keating & Doyle, 2002). Of course, this type of appearance indirectly implicates gender as well, which appears throughout the relationships between status and power.

Personality Antecedents

Just as the Lewinian tradition has equated power with its frequent consequences (social influence), the personality approaches to power equate power with needs for general agency (causing changes in the environment), for personal control, for social dominance, and for social influence. The general need for agency dates back a century. Groos (1901) argued that people feel "joy in being a cause" (p. 385); White (1959) described a fundamental effectance motive that seeks competence; deCharmes (1968) posited the importance of personal causation. A need for mastery (Maslow, 1970; McClelland, 1976) theoretically builds self-esteem. A basic striving for power appears in work by Sullivan (1947) and Adler (1966). Ever since, many psychologists agree that people generally, as social beings, have motives both to predict and to control their outcomes, although people in the West tend to prefer personal control more than people in the East, who tend to prefer harmony and group control (Fiske, 2004; Morling & Fiske, 1999).

Individuals also differ in the extent to which they seek to control their own outcomes and to dominate other people. For example, people vary in need for personal control (Rotter, 1966), (own) mastery orientation (Dweck & Elliott, 1983), and self-efficacy beliefs (Bandura, 1977). And individuals differ in need for power over others (McClelland, 1976; Winter, 1975); high need for power predicts making oneself visible and taking risks to do so (McClelland & Teague, 1975; McClelland & Watson, 1973; Winter, 1973). Leaders high in need for power handle conflict and stress in their groups less well (Fodor, 1984, 1985), share less information and consider fewer ideas (Fodor & Smith, 1982), and produce more negative self-views among their subordinates (Fodor & Riordan, 1995). People high in need for power positively evaluate followers who ingratiate themselves (Fodor & Farrow, 1979). People high on need for dominance also like ingratiators (Gough, 1990; Murray, 1938; Operario & Fiske, 2001; Rotter & Hochreich, 1975), but they do not respect them (Operario & Fiske, 2001).

Similarly, personality matters to preferred influence strategies (Koslowsky & Schwarzwald, 2001). For example, high need for achievement correlates with tight control strategies (Raven, Freeman, & Haley, 1982), and high authoritarianism correlates with coercive and legitimacy-based influence strategies (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950; Altemeyer, 1988). An often-forgotten point: High authoritarians also submit to the influence of their own superiors. These results suggest that authoritarians are especially sensitive to power differences and are likely to exaggerate normal reactions to power inequality, such as using more coercive tactics when powerful and more ingratiating ones when powerless. An intriguing connection may link some forms of self-esteem to authoritarianism: Low self-esteem individuals use harsh influence strategies, without regard for the target's feelings (Raven & Kruglanski, 1970). Perhaps they feel threatened or see the world as more dog-eat-dog. People high on power motives (ability, need, and enjoyment of influence; Bennett, 1988) also see others as wielding power (Imai, 1993); from their perspective, power is everywhere, and they have to behave accordingly.

ACTUAL VERSUS PERCEIVED POWER

Power is not a consequence (influence) or an individual attribute (personality), it is outcome control. To be effective, people must know or believe the powerholder controls their outcomes. Actually controlling physical, economic, and social outcomes is power, with the potential and sometimes the consequence to influence, in our perspective. For example, listening to a discussion but withholding decisive information would constitute actual power but not perceived power. Surprisingly little work addresses how people discern who controls or might control outcomes.

Perceived power is partially captured by the IPI (Raven et al., 1998), described earlier, but also by a Perceived Social Power Scale (Imai, 1989), tapping perceptions that another person controls valued outcomes corresponding to five of the French–Raven bases: reward, coercion, expertise, legitimacy, and reference. People tend to try to influence those who control rewards (Imai, 1993). In vignette studies, coercive and reward power each inflate ratings of both referent and legitimate power (Aguinis, Nesler, Quigley, & Tedeschi, 1994). Although rewards matter, legitimate position as a power basis can be central: Being the supervisor, having the right to make requests, creating an obligation to obey (all features of legitimacy) matters most in both American and Israeli (Raven et al., 1998), as well as Japanese, data (Imai, 1989). While the former datasets pertain to supervisors, the latter ranges from parents to friends to supervisors and subordinates, so the importance of legitimacy generalizes.

Perceived status does not just come from actual roles and positions but also from correlated cues. Expectation states theory, described earlier, explains demographic effects on perceived status by nonverbal mediation. People can in fact assess each other's organizational status accurately, based on nonverbal cues in photographs (Schmid Mast & Hall, 2004). Note that this study defines status as we do power: ratings of control or influence and access to restricted resources, so it seems relevant to actual power. Accurate judgments of women's power used downward *head tilt*, whereas accurate judgments of men used more *formal dress* and *forward lean*. Perceivers also used lowered eyebrows and forward lean for women, and age for men, to judge power, but these cues did not increase their actual accuracy. To judge the related personality trait of assertiveness (Schmid Mast, Hall, Murphy, & Colvin, 2003), accurate judgments of women used *stumbling speech* and *erect posture*. Accurate judgments of men also used *erect posture*, as well as *fillers* and *looking while listening*. For accurately judging women, observers specifically had to avoid judgments based on speaking time and looking while speaking; to be accurate for men, they marginally had to avoid pleasant speech style and talking with hands.

In addition to the demographic and physical traits, personal traits such as perceived credibility impact perceived power (Nesler, Aguinis, Quigley, & Tedeschi, 1993). Credibility—defined as objective truthfulness, follow-through, and accuracy—resembled informational and expert bases of influence but may be viewed as a more personal version of perceived power. One could possess information or have a reputation for knowledge (informational or expert bases) without being an honest broker of it (personally credible). Nevertheless, credibility is closely related to the French–Raven control over information as a basis of potential influence (Eyuboglu & Atac, 1991; Pettigrew, 1972; Raven, 1974, 1988; Raven & Kruglanski, 1970; Yukl & Fabe, 1991).

But how are personal characteristics, such as credibility, communicated? Although many general principles of power's nonverbal cues currently elude us, power does influence eye gaze, which also influences perceived power. Gender enters here, again. In actual interaction, visual dominance ratio (gaze while talking vs. listening) links to manipulations of both expert and reward power for both genders (Dovidio, Ellyson, Keating, Heltman, & Brown, 1988). Both genders, when in tasks linked to their own gender (placing them in positions of expert power), show visual dominance, and men show it in mixed-sex dyads on neutral tasks (Dovidio et al., 1988), consistent with expectation states theory. These differences evaporate when both participants receive training on the topic, whether gender related or neutral (Brown, Dovidio, & Ellyson, 1990), again supporting the importance of expertise. Direct manipulations of status per se, as well as personality measures of expressed control, similarly affect visual dominance (Ellyson, Dovidio, Corson, & Vinicur, 1980).

People apparently know all this, because when observing videotaped speakers with varying gaze dominance ratios, attributions of power increased with look–speak and decreased with look–listen rates (Dovidio & Ellyson, 1982). In vignettes describing male dyads, direct eye contact indicated credibility to perceivers (Aguinis, Simonsen, & Pierce, 1998). For a woman interacting with a man, direct eye contact instead increased her perceived coercive power (Aguinis & Henle, 2001).

In the same pair of studies, going beyond gaze, the manipulated descriptions of nervous versus relaxed expressions reliably indicated both females' and males' perceived referent, reward, legitimate, expert, and credibility power, as well as males' coercive power. Other facial expressions have influenced perceived dominance: happy, angry, and surprised faces, but not sad and fearful faces (Montepare & Dobish, 2003), as well as not smiling (Halberstadt & Saitta, 1987; Keating et al., 1981). Expressing anger as opposed to sadness for a transgression can confer status (Tiedens, 2001). Apparently, conveying one's harmlessness or agreeableness can communicate submission. Indeed, subordinate women who prefer that position did smile more (Schmid Mast & Hall, 2004; see also Schmid Mast & Hall, 2003). Thus, gaze and facial expression interact with gender to communicate power.

Posture can also communicate power. We saw earlier that erect posture communicates status (probably power). Postural expansion (limbs spread out from body) also communicates power (Argyle, 1988; Aries, Gold, & Weigel, 1983; Eibl-Eibesfeldt, 1989; Gifford, 1991; Mehrabian, 1972; Spiegel & Machotka, 1997). When people adopt expanded or conversely constricted posture, other people get the message and tend to recipro-

cate the complementary posture (Tiedens & Fragale, 2003). Unlike other kinds of nonverbal communication, where mimicry creates comfort (Chartrand & Bargh, 1999), for dominance and submission, complementary posture instead creates comfort and liking.

Talking time reliably elicits perceived dominance, especially for men (for a meta-analysis, see Schmid Mast, 2002a). People are not necessarily wrong to judge so; people express dominance through speaking time, especially assigned a dominant role. But perceivers apparently exaggerate the relationship.

SOCIAL PSYCHOLOGY OF POWER: AMBIVALENT MOTIVES

Relative control over valued resources—our definition of power—creates for powerholders a tension between independence from others and responsibility for others.[3] On the one hand, powerholders are relatively free to pursue their own goals with little interference. On the other hand, powerholders are in a position to assist others in pursuing their own goals. These two themes, the evil and virtue of power, reflect two overall approaches in the literature, which contrasts the exercise of power for worse or better, for greed or for good. This dual nature of power taps the basic issue of self versus other, individual versus group, which may help explain why power is considered such a fundamental concept to so many social science disciplines.

Power as Independence

Most of the literature on power to date emphasizes the independence that relative control over valued outcomes affords the powerful and the relative dependence that a lack of control means for the less powerful, along with the positive implications of this arrangement for the powerful and the negative implications for the powerless. For example, power's lack of constraint enables rewards and freedom, encouraging positive affect and approach tendencies, and lacking power activates a threat orientation, negative affect, and inhibition (Keltner et al., 2003). Consistent with this perspective, people who are situationally powerful are more likely than those who are not to speak their mind, see others as liking them, and experience more positive relative to negative emotions (Anderson & Berdahl, 2002; Berdahl & Martorana, 2006), and when power is primed in individuals, they are more likely to take action, whether it is pro- or antisocial in nature (Galinsky, Gruenfeld, & Magee, 2003).

People need to feel control, some contingency between how they behave and what happens to them, between what they do and what they get (as noted earlier, Fiske, 2004, reviews the control literature). Because powerholders control valued resources, their own needs for control are met, whereas the relatively powerless are left out of control, uncertain about their fates. The *power-as-control (PAC) model* takes these ideas seriously (Fiske, 1993).

Attention and Accuracy

For the powerless, their dependence means that they must observe the powerful to understand the contingencies for receiving valued outcomes. Manipulating outcome dependency reliably makes the powerless attend to the powerful, especially their unexpected qualities (most informative); the powerless view the powerholders in more individuated and idiosyncratic ways. Beginning with symmetrical outcome dependency (Erber & Fiske, 1984; Neuberg & Fiske, 1987; Ruscher & Fiske, 1990; Ruscher, Fiske, Miki, & van Manen, 1991), people reliably attend to the person who controls their outcomes, especially to the most diagnostic (unexpected) information; they make dispositional attributions and reach impressions idiosyncratic to their own perspective. Subordinates better recall their superiors' nonverbal behavior than vice versa (e.g., Hall, Carter, & Horgan, 2001). They do not necessarily increase accuracy, but they are motivated by a need to feel accurate. The initial work on symmetrical outcome dependency confounded own dependency with responsibility for the other, as when teammates or competitors both depend on and control each other's fates. However, subsequent studies showed the same attentional and accuracy motivation effects for those who are asymmetrically outcome dependent, here defined as powerless (Dépret & Fiske, 1999; Goodwin, Gubin, Fiske, & Yzerbyt, 2000; Stevens & Fiske, 2000).

Motivation to be accurate mediates the effects of outcome dependency on impression formation (Neuberg & Fiske, 1987). Consistent with the idea of accuracy motivation, less powerful negotiators ask more diagnostic questions of their partners, as a result of greater accuracy and impression motivation (De Dreu & Van Kleef, 2004). Although people are motivated to believe they are accurate, of course they may or may not actually be more accurate. Powerful groups indeed can be less accurate (Ebenbach & Keltner, 1998; Keltner & Robinson, 1997), but that need not always be the case. At a minimum, people reliably attend upward more than downward, consistent with the effort to control (or at least predict) their own outcomes, the main mechanism of the PAC model. Powerful people are freed from constraints imposed by external control, so they can operate as they would by default.

Stereotyping and Discrimination

Besides being unconstrained, as predicted by PAC, powerholders may be motivated to neglect or oppress their subordinates in order to maintain the status quo. That is, they may operate not only by default but by design. Powerholders could use negative stereotypes to legitimate their privileged position. Some evidence fits this interpretation. Powerholders use stereotypic categories more than individuating traits (Goodwin et al., 2000), and this stereotypic focus holds especially for negative stereotypes and for illegitimate power positions, consistent with the idea that threatened powerholders use demeaning stereotypes to legitimate their positions (Rodriguez-Bailon, Moya, & Yzerbyt, 2000).

The first experiments to show self-serving effects of power used the minimal group paradigm, in which people are arbitrarily assigned to ingroup and outgroup and then have the opportunity to distribute rewards. When power was explicitly manipulated, powerful group members discriminated against subordinate group members; they felt comfortable and satisfied in doing so (Sachdev & Bourhis, 1985). As power levels increase, evaluations of others become increasingly negative and evaluations of the self become increasingly positive (Georgesen, & Harris, 1998, 2000). When status (prestige, skill) was manipulated orthogonally to outcome control and group size, both power and status predicted discrimination, especially for numerical minorities. High-status minority powerholders discriminated the most, and low-status minority subordinates discriminated the least, even showing outgroup favoritism (Sachdev & Bourhis, 1991). These results fit justifying the status quo hierarchy, especially under threat. That is, both high- and low-status minorities both gave resources to the higher-status groups. The twist here is that even subordinates justify the hierarchy; we return to this point.

Other research implicates the unconstrained use of power as begetting social ills. Sexual harassment constitutes a form of power abuse. Among men with a proclivity to sexually harass (individual differences in likelihood to sexually harass), subliminally priming power caused them to find a female subordinate more attractive (Bargh, Raymond, Pryor, & Strack, 1995). The link between organizational power (i.e., control over economic outcomes) and sexual harassment has long been known. The first and most widely recognized form of sexual harassment—quid pro quo harassment—involves an organizational superior promising or threatening work-related outcomes to a subordinate in exchange for sexual cooperation. In addition to control over economic outcomes, however, sexual harassers can use control over physical and social ones (i.e., their relative physical might and social status) to harass targets. This helps to explain why "contra-power harassment" is possible: Subordinate males sexually harass their female superiors (DeSouza & Fansler, 2003; McKinney, 1992). It also explains why men who are particularly concerned with protecting their status as men are most likely to harass (Maass, Cadinu, Guarnieri, & Grasselli, 2003), and why women (and men) who threaten male supremacy are most likely to be targeted (Berdahl, in press; Dall'Ara & Maass, 2000; Maass et al., 2003).

Regardless of its intent, the behavior of powerholders can have enormous impact on subordinates. In cross-gender interactions, powerful men praise powerless women but withhold resources from them; women treated this way feel angry, and their performance suffers. Male subordinates treated similarly by a powerful man also feel angry, but it apparently motivates them to perform better (Vescio, Gervais, Snyder, & Hoover, 2005). The fit between the stereotypes and the power positions seems to matter. That is, because female stereotypes include weakness, their stereotype fits the subordinate position as well as influence strategies that patronize them as subordinate (Vescio et al.,

2005). Power interacts with gender once again, in this case reinforcing male dominance.

In dyads, regardless of gender, the powerful tend to influence the subordinate, as suggested by the conflation of power and influence in the literature. For example, powerholders are more able to influence subordinates to confirm their expectancies about them because subordinates are motivated to have an agreeable interaction (Copeland, 1994). Arguably, subordinates are motivated to get along, and powerholders are motivated to get to know (Snyder & Kiviniemi, 2001). When getting to know entails a stable impression, such powerholder motivations merely confirm expectations, both perceptually and behaviorally (Snyder & Haugen, 1994, 1995). In close relationships, less powerful partners adjust to more powerful partners, becoming more emotionally similar over time (Anderson, Keltner, & John, 2003). In decision-making contexts, the less powerful tend to go along with the preferences of the more powerful (Agnew, 1999; Anderson & Berdahl, 2002), unless going along violates a strongly-held belief (Berdahl & Martorana, 2006).

System Justification

More generally, societal theories of power particularly emphasize the system-justifying nature of dominant ideology, or the ideology of the powerful group that also may get adopted by the powerless (Bourdieu, 1972/1977; Engels, 1902/1942; Jost & Banaji, 1994; Jost, Burgess, & Mosso, 2001; Marx, 1867/1987). High-status groups can justify the system and maintain their position, whereas low-status groups justify the system only at their own expense. Evidence indicates that they do so. Not surprisingly, low-status groups exhibit ambivalence toward their own group, especially when they are told or they believe status differences are legitimate (Jost & Burgess, 2000). Low-status groups favor high-status groups and attribute intelligence and responsibility to them, and they do this especially when explanations (even meaningless) are provided for the power difference (Haines & Jost, 2000). The most socially disadvantaged groups are most likely to defend and justify the status quo (Jost, Pelham, Sheldon, & Sullivan, 2003). What is more, low-status groups favor high-status groups even on implicit measures (Jost, Pelham, Brett, & Carvallo, 2002).

Clearly, people do systematically defend the societal bases for hierarchy. For example, people overwhelming report that high-status groups are more competent than low-status groups, around the world (Cuddy, Fiske, & Glick, in press; Cuddy, Fiske, Kwan, et al., in press; Fiske, Cuddy, Glick, & Xu, 2002). Stereotypes of poor people as incompetent and not especially nice, perhaps even exploiting the system, occur in over a dozen cultures (Cuddy, Fiske, & Glick, in press), and they predict observers' disgust, passive harm (exclusion), and active harm (attack) (Cuddy, Fiske, Kwan, et al., in press). Although most poor people elicit a mixture of pity and contempt, subtypes of poor people may prompt pity, for example, the poor-but-happy or poor-but-honest subtypes (Kay & Jost, 2003). There are many routes to system justification.

People endorse a variety of legitimating myths, according to social dominance theory (Sidanius & Pratto, 1999). Believing that some groups naturally dominate other groups in a dog-eat-dog world, which is high SDO, predicts favoring high-status groups, regardless of one's own status (Levin, Federico, Sidanius, & Rabinowitz, 2002). Believing that race-based affirmative action, but decidedly not legacy admission, is preferential treatment constitutes another kind of legitimating myth; a cloaked policy, benefiting the already-advantaged, allegedly differs from a publicized policy benefiting the disadvantaged (Chen & Tyler, 2001).

In sum, a variety of findings show that lack of constraints enable powerholders to justify their disproportionate control over outcomes. Power might create status on a societal level, as noted earlier when control over physical outcomes leads to control over economic ones, leading to control over social ones (i.e., social ideology, stereotypes, and status).

Power as Responsibility

Having relative control over others' valued outcomes means that powerholders are more independent of others, but the flip side is that their behavior and choices have a disproportionate impact on others. This introduces a potential tension in the subjective experience of having power: between independence from others and responsibility for others; between using control over outcomes to pursue one's self-interest and using it to help others pursue theirs. Of these two sides of power, the responsibility side has received much less attention in the literature.

If powerholders are unconstrained by dependency on others, that constraint can operate for ill if the default or more explicit motivation harms other people. But, similarly, lack of constraints can operate for the greater good, if the default motivation helps other people. If powerholders can operate however they will, it follows that their behavior will be more variable than that of subordinates, and indeed it is (Guinote, Judd, & Brauer, 2002). This variability may be experienced as arbitrary by subordinates, who need to predict and try to influence those who control their outcomes. Ironically, or perhaps in some cases deliberately, variability serves to focus attention on dominants (a phenomenon noted earlier). A Machiavellian powerholder would behave randomly just to maintain a sense of power and influence. Regardless of intent, the variability of powerholders' behavior doubtless contributes to their being then viewed as unique individuals (Dépret & Fiske, 1999; Fiske & Dépret, 1996).

"Choice" being the operative term for those unconstrained by other people's control over valued resources, what determines when powerholders choose to behave for good or ill? Both personality and situations can move powerholders in specific directions. First, powerholders are freer to express their personalities. Thus, for example, powerholders with a tendency to be communal (oriented to mutual needs) or have interdependent self-construals treat power as an occasion for social responsibility, whereas exchange-oriented powerholders and those with independent self-construals treat it as an occasion for self-interest (Chen, Lee-Chai, & Bargh, 2001; Chen & Welland, 2002). The Misuse of Power (MOP) scale correlates positively (convergently) with SDO, right-wing authoritarianism, a cynical philosophy of human nature, Machiavellianism, likelihood to sexually harass, and attractiveness of sexual aggression (Lee-Chai, Chen, & Chartrand, 2001). The MOP scale predicts a low communal orientation, consistent with the earlier results (Chen et al., 2001), as well as power-abusing responses to hypothetical scenarios.

In the opposite vein, a Helping Power Motivation scale (Frieze & Boneva, 2001) assesses responsible nurturance, in contrast to egoistic dominance associated with traditional measures of power (McClelland, 1976; McClelland & Teague, 1975; McClelland & Watson, 1973; Winter, 1973, 1975). The Helping Power Motivation scale reassuringly predicts wanting a job that helps others. Communal orientations could theoretically combine with unequal power to create paternalistic exploitation in the name of caring (Pratto & Walker, 2001). An example of such caretaking that demeans the other is the benevolent sexism component of Ambivalent Sexism Inventory (Glick & Fiske, 1996, 2001); so-called protection can lead to denying the protected group education, health care, and economic participation (Glick et al., 2000). Thus, a variety of individual differences can prime more pro- or antisocial responses to power.

Similarly, situations that prime interpersonal concerns can orient powerholders toward individuating and caring for their subordinates, whereas priming an organizational (exchange)-related focus orients them more toward self-interest (Overbeck & Park, 2001). Effectively, we need a person × position × situation understanding of power, whereby personality, structural control over resources, and situational cues all conspire to explain the use of power for good or ill.

To illustrate this *person × position × situation* account, consider research on adult–child interactions (e.g., Bugental & Lin, 2001), where the adult is in a position of structural power (older, bigger, more actual resource control) but can experience self as low power (chronically accessible low-perceived-power schema). The third component, the current situation, then has an inordinate effect on the adult's behavior, as if the person were obsessed with control issues. Under circumstances of clear, veridical high control, the low-perceived-power adult exercises assertiveness and engagement. Under veridically low-control circumstances, the low-perceived-power adult responds with excessive submissiveness and disengagement. In ambiguous circumstances, the same adult exhibits variable patterns. This polarized behavior thus leads someone with structural power to behave in highly variable fashion, but not exactly because of freedom and lack of constraint. On the contrary, the changeable, seemingly arbitrary behavior comes from an insecure powerholder. This style might well generalize beyond adult–child settings as a distinct explanation for the variability of some powerholders' behavior.

DIRECTIONS FOR FUTURE RESEARCH

Research on power in social psychology is in the midst of a revival. Not since the postwar era have so many social psychologists turned their attention to this fundamental aspect of individual experience and social dynamics. Current research on power tends to focus on how power affects individual perceptions, emotions, and behaviors, and the research emphasizes power as independence or freedom that yields positive benefits for the powerful and negative repercussions for the powerless. More research needs to study power in context—not as an individual attribute or orientation that exists independent of a social relationship but as a socially relative and situated phenomenon. By studying power in the context of actual interacting social systems, be they dyads, groups, or organizations, research can also shed light on how the effects of power on individual tendencies manifest themselves in the presence of others. We might also learn about how power affects social outcomes, such as the well-being, decisions, and performance of the social systems under study.

Particularly interesting but understudied aspects of power are its effects over time. More longitudinal designs could address questions about how power emerges, is maintained, and gets challenged (e.g., Fiol, O'Connor, & Aguinis, 2001; Martorana, Galinsky, & Rao, 2005), as well as how fleeting or permanent are its effects for individuals and groups.

Finally, more attention should be paid to the double-edged sword of power, or the ambivalent motives presented by the simultaneous experience of greed and guilt that are likely to follow from having relative control over valued outcomes. Western individualistic ideologies—ideologies perpetrated by the powerful and the independent—focus on power as an opportunity to pursue self-interest and downplay the responsibility for others that comes with it. This may help to explain why this perspective has been the emphasis of the literature on power by scholars in North America and Europe. An interesting avenue for future research would be to study just how individuals, and social systems, manage the dual implications of power, and how they come to emphasize power as an opportunity to maximize selfish desires versus how they come to emphasize power as an obligation to help others.

CONCLUSION

Work on power always begins with long definitional discussion, uncharacteristic of most other social psychology. One can only wonder why, but perhaps the ubiquity of power, status, and influence in social life explains the obsession to understand these concepts, as well as the variations on a theme that defines them. Our assessment is that the field knows more now than it did 10 years ago, and certainly more now than before we started studying power phenomena in rigorous scientific fashion. But plenty remains, particularly in simultaneously assessing personality differences, structural effects, and short-term cues, a research strategy that might capture the complex challenge of the phenomenon of social power.

NOTES

1. We limit our discussion to definitions of what has been called social power, or the power one person may have over another, as this is the focus of this chapter. There is also intrapersonal power, or power over oneself, with which we do not concern ourselves here. The term "power" in this chapter is therefore used with "power over" another in mind.
2. If two people have equal control over each other's outcomes, interdependence increases with this control. That is, when two people both have high control over each other's outcomes, interdependence is high; when they have low or no control over each other's outcomes, interdependence is low. The motive for cooperation and the possibility of conflict increases with interdependence, so individuals with high but equal levels of power may witness the highest levels of cooperation or conflict, but individuals with low but equal levels of power may witness little of either.
3. For the powerless, the tension lies between dependence on others and a lack of responsibility for them. System-justifying ideologies flip responsibility around, absolving the powerholder of responsibility and placing responsibility for self (and even others) on the powerless. To be systematic, one would discuss implications for lacking power as well as having it.

REFERENCES

Adler, A. (1966). The psychology of power. *Journal of Individual Psychology, 22*(2), 166–172.

Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality.* New York: Harper.

Agnew, C. R. (1999). Power over interdependent behavior within the dyad: Who decides what a couple does? In L. J. Severy & W. B. Miller (Eds.), *Advances in population: Psychosocial perspectives* (Vol. 3, pp. 163–188). London: Jessica Kingsley.

Aguinis, H., & Adams, S. K. R. (1998). Social-role versus structural models of gender and influence use in organizations. *Group and Organization Management, 23,* 414–446.

Aguinis, H., & Henle, C. A. (2001). Effects of nonverbal behavior on perceptions of a female employee's power bases. *Journal of Social Psychology, 141,* 537–549.

Aguinis, H., Nesler, M. S., Hosoda, M., & Tedeschi, J. T. (1994). The use of influence tactics in persuasion. *Journal of Social Psychology, 134,* 429–438.

Aguinis, H., Nesler, M. S., Quigley, B. M., Lee, S-J., & Tedeschi, J. T. (1996). Discussion in the college classroom: Triangulating observational and survey results. *Journal of Higher Education, 67,* 267–297.

Aguinis, H., Nesler, M. S., Quigley, B. M., & Tedeschi, J. T. (1994). Perceptions of power: A cognitive perspective. *Social Behavior and Personality, 22,* 377–384.

Aguinis, H., Simonsen, M. M., & Pierce, C. A. (1998). Effects of nonverbal behavior on perceptions of power bases. *Journal of Social Psychology, 138,* 455–469.

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism.* San Francisco: Jossey-Bass.

Anderson, C., & Berdahl, J. L. (2002). The experience of power: Examining the effects of power on approach and inhibition tendencies. *Journal of Personality and Social Psychology, 83,* 1362–1377.

Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, *81*, 116–132.

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology*, *84*, 1054–1068.

Arendt, H. (1960). Society and culture. *Daedalus*, *89*, 278–287.

Argyle, M. (1988). *Bodily communication* (2nd ed.). London: Methuen.

Aries, E. J., Gold, C., & Weigel, R. H. (1983). Dispositional and situational influences on dominance behavior in small groups. *Journal of Personality and Social Psychology*, *44*(4), 779–786.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bargh, J. A., Raymond, P., Pryor, J. B., & Strack, F. (1995). Attractiveness of the underling: An automatic power– sex association and its consequences for sexual harassment and aggression. *Journal of Personality and Social Psychology*, *68*, 768–781.

Bass, B. M. (1956). Leadership opinions as forecasts of supervisory success. *Journal of Applied Psychology*, *40*, 345–346.

Bass, B. M. (1960). *Leadership, psychology, and organizational behavior*. Oxford, UK: Harper.

Bass, B. M. (1981). From leaderless group discussions to the cross-national assessment of managers. *Journal of Management*, 7(2), 63–76.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Beckham, B., & Aronson, H. (1978). Selection of jury foremen as a measure of the social status of women. *Psychological Reports*, *43*, 475–478.

Bennett, J. B. (1988). Power and influence as distinct personality traits: Development and validation of a psychometric measure. *Journal of Research in Personality*, *22*(3), 361–394.

Berdahl, J. L. (in press). The sexual harassment of uppity women. *Journal of Applied Psychology*.

Berdahl, J. L., & Anderson, C. (2005). Gender and hierarchy in groups over time. *Group Dynamics: Theory, Research, and Practice*, *9*, 45–57.

Berdahl, J. L., & Martorana, P. (2006). Effects of power on emotion and expression during a controversial group discussion. *European Journal of Social Psychology [Special Issue on Social Power]*, *36*, 497–510.

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255.

Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge, UK: Cambridge University Press. (Original work published 1972)

Brown, J. H., Dovidio, J. F., & Ellyson, S. L. (1990). Reducing sex differences in visual displays of dominance: Knowledge is power. *Personality and Social Psychology Bulletin*, *16*, 358–368.

Brown, J. H., & Raven, B. H. (1994). Power and compliance in doctor/patient relationships. *Journal of Health Psychology*, *6*, 3–22.

Bugental, D. B., & Lin, E. K. (2001). The many faces of power: The strange case of Dr. Jekyll and Mr. Hyde. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 115–132). Philadelphia: Taylor & Francis.

Cartwright, D. (1959). Power: A neglected variable in social psychology. In D. Cartwright (Ed.), *Studies in social power* (pp. 1–14). Ann Arbor: University of Michigan, Institute for Social Research.

Cartwright, D. (1965). Influence, leadership, control. In J. G. March (Ed.), *Handbook of organizations* (pp. 1–47). Chicago: Rand McNally.

Chaiken, S., Wood, W., & Eagly, A. H. (1996). Principles of persuasion. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 702–742). New York: Guilford Press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The

perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Chen, E. S., & Tyler, T. R. (2001). Cloaking power: Legitimatizing myths and the psychology of the disadvantaged. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 241–261). Philadelphia: Taylor & Francis.

Chen, S., Lee-Chai, A. Y., & Bargh, J. A. (2001). Relationship to orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology*, *80*, 173–187.

Chen, S., & Welland, J. (2002). Examining the effects of power as a function of self-construals and gender. *Self and Identity*, *1*, 251–269.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 151–192). New York: McGraw-Hill.

Copeland, J. T. (1994). Prophecies of power: Motivational implications of social power for behavioral confirmation. *Journal of Personality and Social Psychology*, *67*, 264–277.

Cropanzano, R., Aguinis, H., Schminke, M., & Denham, D. L. (1999). Disputant reactions to managerial conflict resolution tactics. *Group and Organization Management*, *24*, 124–154.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (in press). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*.

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Leyens, J.-P., et al. (in press). Is the stereotype content model culture-bound?: A cross-cultural comparison reveals systematic similarities and differences. *British Journal of Social Psychology*.

Dahl, R. A. (1957). The concept of power. *Behavioral Science*, *2*, 201–218.

Dall'Ara, E., & Maass, A. (2000). Studying sexual harassment in the laboratory: Are egalitarian women at higher risk? *Sex Roles*, *41*, 681–704.

deCharmes, R. (1968). *Personal causation: The internal affective determinants of behavior*. New York: Academic Press.

DeDreu, C. K. W., & Van Kleef, G. A. (2004). The influence of power on the information search, impression formation, and demands in negotiation. *Journal of Experimental Social Psychology*, *40*(3), 303–319.

Dépret, E. F., & Fiske, S. T. (1993). Social cognition and power: Some cognitive consequences of social structure as a source of control deprivation. In G. Weary, F. Gleicher, & K. Marsh (Eds.), *Control motivation and social cognition* (pp. 176–202). New York: Springer-Verlag.

Dépret, E. F., & Fiske, S. T. (1999). Perceiving the powerful: Intriguing individuals versus threatening groups. *Journal of Experimental Social Psychology*, *35*, 461–480.

DeSouza, E., & Fansler, A. G. (2003). Contrapower sexual harassment: A survey of students and faculty members. *Sex Roles*, *48*, 519–542.

Dovidio, J. F., & Ellyson, S. L. (1982). Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly*, *45*, 106–113.

Dovidio, J. F., Ellyson, S. L., Keating, C. F., Heltman, K., & Brown, C. E. (1988). The relationship of social power to visual displays of dominance between men and women. *Journal of Personality and Social Psychology*, *54*, 233–242.

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. In E. M. Hetherington (Ed.), *Socialization, personality, and social development* (pp. 643–691). New York: Wiley.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Mahwah, NJ: Erlbaum.

Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, *129*, 569–591.

Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, *108*, 233–256.

Eagly, A. H., Karau, S. J., Miner, J. B., & Johnson, B. T. (1994). Gender and motivation to manage in hierarchic organizations: A meta-analysis. *Leadership Quarterly*, *5*(2), 135–159.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*, 3–22.

Ebenbach, D. H., & Keltner, D. (1998). Power, emotion, and judgmental accuracy in social conflict: Motivating the cognitive miser. *Basic and Applied Social Psychology*, *20*, 7–21.

Eibl-Eibesfeldt, I. (1989). *Human ethology. Foundations of human behavior*. Hawthorne, NY: Aldine de Gruyter.

Elias, S. M. (2004). Means of assessing ordinal interactions in social psychology: The case of sexism in judgments of social power. *Journal of Applied Social Psychology*, *34*, 1857–1877.

Elias, S. M., & Loomis, R. J. (2004). The effect of instructor gender and race/ethnicity on gaining compliance in the classroom. *Journal of Applied Social Psychology*, *34*, 937–958.

Ellyson, S. L., Dovidio, J. F., Corson, R. L., & Vinicur, D. L. (1980). Visual dominance behavior in female dyads: Situational and personality factors. *Social Psychology Quarterly*, *43*, 328–336.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, *27*, 31–41.

Engels, F. (1942). *The origin of the family, private property and the state*. New York: International. (Original work published 1902)

Erber, R., & Fiske, S. T. (1984). Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology*, *47*, 709–726.

Erchul, W. P., Raven, B. H., & Ray, A. G. (2001). School psychologists' perceptions of social power bases in teacher consultation. *Journal of Educational and Psychological Consultation*, *12*, 1–23.

Erchul, W. P., Raven, B. H., & Whichard, S. M. (2001). School psychologist and teacher perceptions of social power in consultation. *Journal of School Psychology*, *39*, 483–497.

Erchul, W. P., Raven, B. H., & Wilson, K. E. (2004). The relationship between gender of consultant and social power perceptions within school consultation. *School Psychology Review*, *33*, 582–590.

Eyuboglu, N., & Atac, O. A. (1991). Informational power: A means for increased control in channels of distribution. *Psychology and Marketing*, *8*, 197–213.

Fiol, C. M., O'Connor, E. J., & Aguinis, H. (2001). All for one and one for all? The development and transfer of power across organizational levels. *Academy of Management Review*, *26*, 224–242.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.

Fiske, S. T. (2004). *Social beings: A core motives approach to social psychology*. New York: Wiley.

Fiske, S. T., & Dépret, E. (1996). Control, interdependence, and power: Understanding social cognition in its social context. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 7, pp. 31–61). New York: Wiley.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902.

Fodor, E. M. (1984). The power motive and reactivity to power stresses. *Journal of Personality and Social Psychology*, *47*, 853–859.

Fodor, E. M. (1985). The power motive, group conflict, and physiological arousal. *Journal of Personality and Social Psychology*, *49*, 1408–1415.

Fodor, E. M., & Farrow, D. L. (1979). The power motive as an influence on use of power. *Journal of Personality and Social Psychology*, *37*(11), 2091–2097.

Fodor, E. M., & Riordan, J. M. (1995). Leader power motive and group conflict as influences on leader behavior and group member self-affect. *Journal of Research in Personality*, *29*, 418–431.

French, J. R. P., Jr., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Ann Arbor: University of Michigan, Institute for Social Research.

Frieze, I. H., & Boneva, B. (2001). Power motivation and motivation to help others. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 75–89). Philadelphia: Taylor & Francis.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*, 453–466.

Georgesen, J. C., & Harris, M. J. (1998). Why's my boss always holding me down? A meta-analysis of power effects on performance evaluations. *Personality and Social Psychology Review*, *2*, 184–195.

Georgesen, J. C., & Harris, M. J. (2000). The balance of power: Interpersonal consequences of differential power and expectation. *Personality and Social Psychology Bulletin*, *26*, 1239–1257.

Gerstner, C. R., & Day, D. V. (1997). Meta-analytic review of leader-member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, *82*, 827–844.

Gifford, R. (1991). Mapping non-verbal behavior on the interpersonal circle. *Journal of Personality and Social Psychology*, *61*, 279–288.

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491–512.

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, *56*, 109–118.

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., et al. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, *79*, 763–775.

Gold, G. J., & Raven, B. H. (1992). Interpersonal influence strategies in the Churchill–Roosevelt bases-for-destroyers exchange. *Journal of Social Behavior and Personality*, *7*, 245–272.

Goodwin, S. A., Gubin, A., Fiske, S. T., & Yzerbyt, V. (2000). Power can bias impression formation: Stereotyping subordinates by default and by design. *Group Processes and Intergroup Relations*, *3*, 227–256.

Gough, H. G. (1990). Testing for leadership with the California Psychological Inventory. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 355–379). West Orange, NJ: Leadership Library of America.

Groos, K. (1901). *The play of man* (E. L. Baldwin, Trans.). New York: Appleton.

Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: Evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology*, *82*, 708–721.

Haines, E. L., & Jost, J. T. (2000). Placating the powerless: Effects of legitimate and illegitimate explanation on effect, memory, and stereotyping. *Social Justice Research*, *13*, 219–236.

Halberstadt, A. G., & Saitta, M. B. (1987). Gender, nonverbal behavior, and perceived dominance: A test of the theory. *Journal of Personality and Social Psychology*, *53*(2), 257–272.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2001). Status roles and recall of nonverbal cues. *Journal of Nonverbal Behavior*, *25*, 79–100.

Imai, Y. (1989). The relationship between perceived social power and the perception of being influenced. *Japanese Psychological Research*, *31*, 97–107.

Imai, Y. (1991). Effects of influence strategies, perceived social power and cost on compliance with requests. *Japanese Psychological Research*, *33*, 134–144.

Imai, Y. (1993). Perceived social power and power motive in interpersonal relationships. *Journal of Social Behavior and Personality*, *8*, 687–702.

Imai, Y. (1994). Effects of influencing attempts on the perceptions of powerholders and the powerless. *Journal of Social Behavior and Personality*, *9*, 455–468.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*, 1–27.

Jost, J. T., & Burgess, D. (2000). Attitudinal ambivalence and the conflict between group and system justification motives in low status groups. *Personality and Social Psychology Bulletin*, *26*, 293–305.

Jost, J. T., Burgess, D., & Mosso, C. O. (2001). Conflicts of legitimation among self, group and system: The integrative potential of

system justification theory. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 363–388). New York: Cambridge University Press.

Jost, J. T., Pelham, B. W., Brett, W., & Carvallo, M. R. (2002). Nonconscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, *38*, 586–602.

Jost, J. T., Pelham, B. W., Sheldon, O., & Sullivan, B. Ni. (2003). Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged. *European Journal of Social Psychology*, *33*, 13–36.

Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of "Poor but Happy" and "Poor but Honest" stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology*, *85*, 823–837.

Keating, C. F., & Doyle, J. (2002). The faces of desirable mates and dates contain mixed social status cues. *Journal of Experimental Social Psychology*, *38*(4), 414–424.

Keating, C. F., Mazur, A., Segall, M. H., Cysneiros, P. E., Divak, W. T., Killside, J. E., et al. (1981). Culture and the perception of social dominance from facial expression. *Journal of Personality and Social Psychology*, *40*(4), 615–626.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, *2*, 51–60.

Keltner, D., & Robinson, R. J. (1997). Defending the status quo: Power and bias in social conflict. *Personality and Social Psychology Bulletin*, *23*, 1066–1077.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*, 265–284.

Kipnis, D. (1972). Does power corrupt? *Journal of Personality and Social Psychology*, *24*, 33–41.

Kipnis, D. (1976). *The powerholders*. Chicago: University of Chicago Press.

Kipnis, D. (1984). The use of power in organizations and in interpersonal settings. *Applied Social Psychology Annual*, *5*, 179–210.

Kipnis, D. (2001). Using power: Newton's second law. In A. Y. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 3–17). Philadelphia: Psychology Press.

Koslowsky, M., & Schwarzwald, J. (2001). The power interaction model. In A. Y. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 195–214). Philadelphia: Taylor & Francis.

Koslowsky, M., Schwarzwald, J., & Ashuri, S. (2001). On the relationship between subordinates' compliance to power sources and organisational attitudes. *Applied Psychology: An International Review*, *50*, 455–476.

Lee-Chai, A. Y., Chen, S., & Chartrand, T. (2001). From Moses to Marcos: Individual differences in the use and abuse of power. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 57–74). Philadelphia: Psychology Press.

Levin, S., Federico, C. M., Sidanius, J., & Rabinowitz, J. L. (2002). Social dominance orientation and intergroup bias: The legitimation of favoritism for high-status groups. *Personality and Social Psychology Bulletin*, *28*, 144–157.

Lewin, K. (1944). Constructs in psychology and psychological ecology. *University of Iowa Studies in Child Welfare*, *20*, 1–29. (Reprinted under title Constructs in field theory. In D. Cartwright (Ed.), *Field theory in social science* (pp. 30–42). London: Tavistock.

Lewin, K. (1997). Behavior and development as a function of the total situation. In D. Cartwright (Ed.), *Field theory in social psychology* (pp. 337–381). Washington, DC: American Psychological Association. (Original work published 1951)

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformation and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly*, *7*, 385–425.

Lukes, S. (1986). *Power: Readings in social and political theory*. New York: University Press.

Maass, A., Cadinu, M., Guarnieri, G., & Grasselli, A. (2003). Sexual harassment under social identity threat: The computer harassment paradigm. *Journal of Personality and Social Psychology*, *85*, 853–780.

Manz, C. C., & Gioia, D. A. (1983). The interrelationship of power and control. *Human Relations*, *36*, 459–476.

Martorana, P. V., Galinsky, A. D., & Rao, H. (2005). From system justification to system condemnation: Antecedents of attempts to change power hierarchies. In M. Neale, E. A. Mannix, & M. Thomas-Hunt (Eds.), *Research in managing teams and groups* (Vol. 7, pp. 285–315). Oxford, UK: Elsevier Science.

Marx, K. (1987). *Capital* (Vol. 1.) Moscow: Progress. (Original work published 1867)

Maslow, A. (1970). *Motivation and personality* (2nd ed.). New York: Harper & Row.

Mazur, A. (1985). A biosocial model of status in face-to-face primate groups. *Social Forces*, *64*, 377–402.

McClelland, D. C. (1976). *The achieving society*. Oxford, UK: Irvington.

McClelland, D. C., & Teague, G. (1975). Predicting risk preferences among power related tasks. *Journal of Personality*, *43*, 266–285.

McClelland, D. C., & Watson, R. I. (1973). Power motivation and risk-taking behavior. *Journal of Personality*, *41*, 121–139.

McKinney, K. (1992). Contrapower sexual harassment: The effects of student sex and type of behavior on faculty perceptions. *Sex Roles*, *27*, 627–643.

Mehrabian, A. (1972). *Nonverbal communication*. Oxford, UK: Aldine-Atherton.

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*, 237–254.

Morling, B., & Fiske, S. T. (1999). Defining and measuring harmony control. *Journal of Research in Personality*, *33*, 379–414.

Mullen, B. (1991). Group composition, salience, and cognitive representations: The phenomenology of being in a group. *Journal of Experimental Social Psychology*, *27*, 297–323.

Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.

Nesler, M. S., Aguinis, H., Quigley, B. M., Lee, S-J., & Tedeschi, J. T. (1999). The development and validation of a scale measuring global social power based on French and Raven's power taxonomy. *Journal of Applied Social Psychology*, *29*, 750–771.

Nesler, M. S., Aguinis, H., Quigley, B. M., & Tedeschi, J. T. (1993). The effect of credibility and perceived power. *Journal of Applied Social Psychology*, *23*, 1407–1425.

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*, 431–444.

Ng, S. H. (1980). *The social psychology of power*. New York: Academic Press.

Operario, D., & Fiske, S. T. (2001). Effects of trait dominance on powerholders' judgments of subordinates. *Social Cognition*, *19*, 161–180.

Overbeck, J. R., & Park, B. (2001). When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology*, *81*, 549–565.

Pettigrew, A. (1972). Information control as a power resource. *Sociology*, *6*, 187–204.

Pratto, F., Stallworth, L. M., & Sidanius, J. (1997). The gender gap: Differences in political attitudes and social dominance orientation. *British Journal of Social Psychology*, *36*(1), 49–68.

Rahim, M. A., Antonioni, D., & Psenicka, C. (2001). A structural equations model of leader power, subordinates' styles of handling conflict, and job performance. *International Journal of Conflict Management*, *12*, 191–211.

Raven, B. H. (1974). The comparative analysis of power and power preferences. In J. T. Tedeschi (Ed.), *Perspectives on social power* (pp. 172–198). Chicago: Aldine.

Raven, B. H. (1988). Social power and compliance in health care. In S. Maes, C. D. Spielberger, P. B. Defares, & I . G. Sarason (Eds.), *Topics in health psychology* (pp. 229–244). Oxford, UK: Wiley.

Raven, B. H. (1990). Political applications of the psychology of interpersonal influence and social power. *Political Psychology*, *11*, 493–520.

Raven, B. H. (1992). A power/interaction model of interpersonal influence: French and Raven thirty years later. *Journal of Social Behavior and Personality*, *7*, 217–244.

Raven, B. H. (1993). The bases of power: Origins and recent developments. *Journal of Social Issues*, *49*, 227–251.

Raven, B. H. (1999). Influence, power, religion, and the mechanisms of social control. *Journal of Social Issues*, *55*, 161–186.

Raven, B. H. (2001). Power/interaction and interpersonal influence. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 217–240). Philadelphia: Taylor & Francis.

Raven, B. H., Freeman, H. E., & Haley, R. W. (1982). Social power and compliance in hospital infection control. In A. W. Johnson, O. Grusky, & B. H. Raven (Eds.), *Contemporary health services: Social science perspectives* (pp. 139–176). Boston: Auburn House.

Raven, B. H., & Kruglanski, A. W. (1970). Conflict and power. In P. G. Swingle (Ed.), *The structure of conflict* (pp. 69–109). New York: Academic Press.

Raven, B. H., Schwarzwald, J., & Koslowsky, M. (1998). Conceptualizing and measuring a power/interaction model of interpersonal influence. *Journal of Applied Social Psychology*, *28*, 307–332.

Rodriguez-Bailon, R., Moya, M., & Yzerbyt, V. (2000). Why do superiors attend to negative stereotypic information about their subordinates? Effects of power legitimacy on social perception. *European Journal of Social Psychology*, *30*, 651–671.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28.

Rotter, J. B., & Hochreich, D. J. (1975). *Personality*. Glenview, IL: Scott, Foresman.

Ruscher, J. B., & Fiske, S. T. (1990). Interpersonal competition can cause individuating processes. *Journal of Personality and Social Psychology*, *58*, 832–843.

Ruscher, J. B., Fiske, S. T., Miki, H., & van Manen, S. (1991). Individuating processes in competition: Interpersonal versus intergroup. *Personality and Social Psychology Bulletin*, *17*, 595–605.

Russell, B. (1938). *Power: A new social analysis*. New York: Norton.

Sachdev, I., & Bourhis, R. Y. (1985). Social categorization and power differentials in group relations. *European Journal of Social Psychology*, *15*, 415–434.

Sachdev, I., & Bourhis, R. Y. (1991). Power and status differentials in minority and majority group relations. *European Journal of Social Psychology*, *21*, 1–24.

Schmid Mast, M. (2001). Gender differences and similarities in dominance hierarchies in same-gender groups based on speaking time. *Sex Roles*, *44*, 537–556.

Schmid Mast, M. (2002a). Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research*, *28*, 420–450.

Schmid Mast, M. (2002b). Female dominance hierarchies: Are they any different from males? *Personality and Social Psychology Bulletin*, *28*, 29–39.

Schmid Mast, M. (2004). Men are hierarchical, women are egalitarian: An implicit gender stereotype. *Swiss Journal of Psychology*, *63*, 107–111.

Schmid Mast, M., & Hall, J. A. (2003). Anybody can be a boss but only certain people make good subordinates: Behavioral impacts of striving for dominance and dominance aversion. *Journal of Personality*, *71*, 871–891.

Schmid Mast, M., & Hall, J. A. (2004). When is dominance related to smiling? Assigned dominance, dominance preference, trait dominance, and gender as moderators. *Sex Roles*, *50*, 387–399.

Schmid Mast, M., & Hall, J. A. (2004). Who is the boss and who is not? Accuracy of judging status. *Journal of Nonverbal Behavior*, *28*, 145–165.

Schmid Mast, M., Hall, J. A., Murphy, N. A., & Colvin, C. R. (2003). Judging assertiveness. *Philosophy, Sociology and Psychology*, *2*, 731–744.

Schubert, T. W. (2004). The power in your hand: Gender differences in bodily feedback from making a fist. *Personality and Social Psychology Bulletin*, *30*, 757–769.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression.* New York: Cambridge University Press.

Simon, H. A. (1957). *Models of man*. New York: Wiley.

Snyder, M., & Haugen, J. A. (1994). Why does behavioral confirmation occur? A functional perspective on the role of the perceiver. *Journal of Experimental Social Psychology*, *30*(3), 218–246.

Snyder, M., & Haugen, J. A. (1995). Why does behavioral confirmation occur? A functional perspective on the role of the target. *Personality and Social Psychology Bulletin*, *21*(9), 963–974.

Snyder, M., & Kiviniemi, M. T. (2001). Getting what they came for: How power influences the dynamics and outcomes of interpersonal interaction. In A. Lee-Chai & J. A. Bargh (Eds.), *The use and abuse of power* (pp. 133–155). Philadelphia: Taylor & Francis.

Spiegel, J., & Machotka, P. (1974). *Messages of the body*. New York: Free Press.

Stevens, L. E., & Fiske, S. T. (2000). Motivated impressions of a powerholder: Accuracy under task dependency and misperception under evaluative dependency. *Personality and Social Psychology Bulletin*, *26*, 907–922.

Sullivan, H. S. (1947). *Conceptions of modern psychiatry*. Washington, DC: William Alanson White Psychiatric Foundation.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral. *Journal of Personality and Social Psychology*, *80*, 86–94.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementary in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*, 558–568.

Turner, J. C. (1995). Social influence. In A. S. R. Manstead & M. Hewstone (Eds.), *Blackwell encyclopedia of social psychology* (pp. 562–567). Cambridge, MA: Blackwell.

Vescio, T. K., Snyder, M., & Gervais, S. J. (in press). Power, stereotyping, and social influence: The dynamics of interactions between the powerful and the powerless. *European Review of Social Psychology*.

Vescio, T. K., Gervais, S. J., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology*, *88*, 658–672.

Wartenberg, T. E. (1990). *The forms of power: From domination to transformation*. Philadelphia: Temple University Press.

Weber, M. (1978). *Economy and society* (G. Roth & C. Wittich, Trans.). Berkeley: University of California Press. (Original work published 1914)

White, R. W. (1959). Motivation reconsidered: The concepts of competence. *Psychological Review*, *66*, 297–333.

Wilson, P. R. (1968). Perceptual distortion of height as a function of ascribed academic status. *Journal of Social Psychology*, *74*(1), 97–102.

Winter, D. G. (1973). The power motive. New York: Free Press.

Winter, D. G. (1975, August). *Power motives and power behavior in women*. Paper prepared for Symposium on Women: Studies of power and powerlessness. American Psychological Association Convention, Chicago.

Wofford, J. C., & Liska, L. Z. (1993). Path-goal theories of leadership: A meta-analysis. *Journal of Management*, *19*, 857–876.

Yukl, G. (1989). Managerial leadership: A review of theory and research. *Journal of Management*, *15*, 251–289.

Yukl, G., & Falbe, C. M. (1991). Importance of different power sources in downward and lateral relations. *Journal of Applied Psychology*, *76*(3), 416–423.

Zebrowitz-McArthur, L., & Montepare, J. M. (1989). Contributions of a babyface and a childlike voice to impressions of moving and talking faces. *Journal of Nonverbal Behavior*, *13*(3), 189–203.

GROUP AND CULTURAL SYSTEM

The Social Psychology
of Intergroup Relations

Social Categorization, Ingroup Bias,
and Outgroup Prejudice

MARILYNN B. BREWER

Our understanding of intergroup relations draws on research from all domains of social psychology. Research in this area encompasses intraindividual, interpersonal, and group processes and incorporates theory and data from the study of self and identity, social cognition and attribution and attitudes to competition, aggression, and conflict. An empirical review of this vast literature is beyond the scope of this chapter. Rather than an exhaustive review of empirical findings, my purpose here is to extract some basic principles that have emerged from social psychological research on intergroup relations that illustrate how social psychology can inform both basic research and public policy in this important arena of human behavior. The organization of the chapter begins with an elaboration of three basic principles and concludes with the application of these three principles to reducing prejudice and intergroup conflict.

In brief, I argue that the social psychology of intergroup relations rests on the following fundamental propositions:

1. Group-based attitudes, perceptions, and behavior arise from basic cognitive categorization processes that partition the social world into *ingroups* and *outgroups.*
2. Attachment to and preference for ingroups is the primary driver of intergroup relations. *Ingroup favoritism* gives rise to intergroup discrimination, irrespective of attitudes toward specific outgroups.
3. Attitudes and emotions toward specific outgroups reflect appraisals of the nature of the relationships be-

tween ingroup and outgroup that have implications for the maintenance or enhancement of ingroup resources, values, and well-being. *Outgroup prejudices* both reflect and justify the existing structure of intergroup relations.

In the following pages I summarize the research literature that supports these basic propositions and then show how these principles have been integrated in social psychology's contributions to social programs designed to reduce intergroup prejudice and conflict.

PRINCIPLE 1: SOCIAL CATEGORIZATION

At least since the 1950s (cf. Allport, 1954; Campbell, 1956; Tajfel, 1969), social psychologists have recognized that intergroup prejudice and stereotyping arise in part from normal processes of categorization of the social world. Categorization (1) partitions the multidimensional variability among human beings into discrete subsets, accompanied by (2) accentuation of perceived intracategory similarities and intercategory differences (Doise, Deschamps, & Meyer, 1978; Tajfel, 1969; Tajfel & Wilkes, 1963).

The basic processes of categorization and category accentuation are presumed to be the same whether we are talking about individuals' partitioning of the world of physical objects and events or the social world. Individuals learn to classify objects as functionally interchangeable and develop concepts that distinguish members of

one category from those of another as a fundamental tool for negotiating the physical and social environment. Just as category learning and category representations are functional necessities for dealing with objects and events in the environment generally, discrete social categories serve to simplify, structure, and regulate our understandings of and interactions with other people. By carving variability among individuals into discrete groupings, categorization reduces complexity and leads to enhanced perceived similarity within categories and contrast (differentiation) between categories. Category distinctions influence both perception of and behavior toward category members, individually and collectively.

Social categorization underlies the phenomena of group cognition and social stereotypes. Social categorization also underlies ingroup–outgroup differentiation and associated intergroup processes. Because of this common origin, there has been a tendency in the social psychological literature to conflate studies of stereotyping, prejudice, and ingroup bias—as if *group* cognition and *intergroup* relations were one and the same. By contrast, I think it is important to maintain a distinction between these two consequences of social categorization, as represented in Figure 30.1.

On the one hand, as represented on the left side of Figure 30.1, social categorization gives rise to social stereotypes in the form of category prototypes, perceived trait distributions, and implicit theories about the social meaning of the category. Category stereotypes in turn have both evaluative and behavioral implications. Evaluations of individual category members and the category as a whole are derived from the positivity and negativity of the category stereotype, which may be predominantly positive, predominantly negative, or a mixture of both positive and negative characteristics.

Category stereotypes also regulate behavior toward category members. Stereotype-based discrimination arises when persons are treated differentially as a function of their category membership based on beliefs about the category as a whole. Gender discrimination provides a good illustrative case for the disassociation between stereotype-based discrimination and negative prejudice. Stereotypes about women are predominantly positive in

evaluative connotation, yet women are often disadvantaged in employment selection and promotion because of the implications of those stereotypes for expectations regarding competency and power. An important feature here is that such stereotype-based discrimination can follow from beliefs about *any* social category, *regardless of whether the perceiver is a member of that category or not*. In that sense, stereotyping is the product of "cold cognition," uninfluenced by self-referencing and associated emotional significance.

It is precisely this element of self-referencing or self-involvement that distinguishes stereotyping from prejudice, as represented on the right side of Figure 30.1. Prejudice arises when ingroup–outgroup differentiation is engaged in connection with particular social categorizations. In other words, prejudice involves thinking of social groups or categories in me–not me terms. Category membership alone is not sufficient to engage this differentiation between self and not-self. Even though I recognize that I am a member of the category "women," I can think about women as a social category (with associated stereotypes) without thinking of women as an "ingroup." Ingroup differentiation involves an additional process of *self*-categorization (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) or social identification whereby my sense of self is extended to the group as a whole. Similarly, I can think about a category (e.g., "librarians") to which I do not belong without invoking "outgroup" feelings. A category becomes an outgroup only when the self is actively disassociated from the group, in a "not-me" sense.

As depicted in Figure 30.1, prejudice arising from ingroup–outgroup differentiation processes can have three different loci. One form is the *ingroup* favoritism that has been the focus of much of the research on intergroup relations conducted within the social identity theory tradition. Here the focus is on differentiation of the ingroup from everyone else (the "us"–"not us"distinction). In this case there may or may not be any explicit outgroup; just the generalized "others" is sufficient. Prejudice and discrimination arise from differential favorability/positivity toward those who share this ingroup identity, but without any corresponding negativity or hostility toward non-ingroup members. Discrimination results from withholding from others favors and benefits that are extended only to the ingroup.

A second form of prejudice—perhaps the most virile form—is focused on the outgroup, without the necessity of any explicit ingroup identification being involved. The important distinction here is between "them" and "me," an explicit dissociation of the self from the target outgroup usually accompanied by negativity and hostility toward that group. In many cases, prejudice against gays or against specific stigmatized groups such as "skinheads" is probably outgroup prejudice of this type. Discrimination that is derived from outgroup prejudice (hate) is actively directed at harming or disadvantaging members of the outgroup, whether or not any personal benefit is gained in the process.

Finally, a third form of prejudice derives from the relationship *between* an ingroup and specific outgroups—the
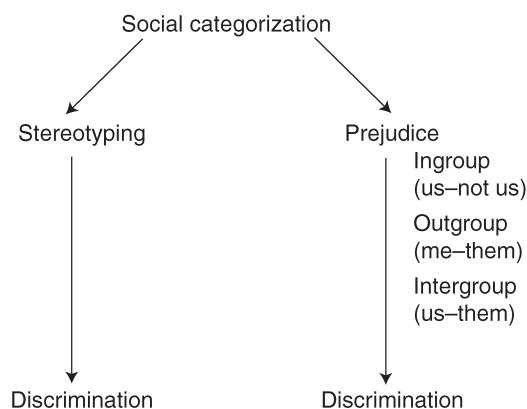
FIGURE 30.1. Consequences of social categorization.

classic "us"–"them" distinction. This is the type of prejudice that is aroused when intergroup comparison and competition is activated, with the consequence that ingroup benefits come at the expense of the outgroup and vice versa. It is the type of prejudice aroused when the outgroup is perceived as a threat, not only to the self but to the integrity, interests, or identity of the ingroup as a whole. Discrimination derived from this form of prejudice is motivated more by ingroup protection (rather than enhancement) as well as antagonism toward the outgroup.

These forms of ingroup–outgroup prejudice are, of course, not mutually exclusive. All share the characteristic of strong emotional and affective investment ("hot" cognition) associated with self-involvement. Many chronic prejudices (and associated discrimination) probably involve elements of all three types of ingroup–outgroup prejudice. In some cases, however, claims of ingroup preservation and protection may be used to justify what is really outgroup prejudice or hate. White supremists, for instance, frequently claim that their real purpose is enhancing and maintaining the identity and interests of the white race (as opposed to being anti-outgroups). But close scrutiny of their rhetoric and internal communications reveals a heavy dominance of outgroup hate speech relative to ingroup promotion (Von Hippel, Brewer, & Polifroni, 2001).

Separating Intergroup Relations from Group Cognition

My emphasis here on the conceptual and empirical distinction between group cognition as represented by stereotyping, and intergroup processes as represented by ingroup–outgroup prejudice is not meant to imply that the two are unrelated. On the contrary, there is probably a great deal of "leakage" between group cognition and ingroup–outgroup feelings and vice versa. Strong negative stereotypes of a particular social category (to which one does not belong) are very likely to lead to negative affect and prejudice directed against that group. On the other hand, ingroup favoritism and outgroup antagonism can color the evaluation of characteristics associated with that social category. Even positive stereotypical traits can become negatively valued when they are attributed to an outgroup, and ingroup traits can be positively valued even when they are stigmatizing characteristics (Campbell, 1967). Nonetheless, even though there is certainly this type of mutual influence between stereotypes and prejudices, it is important to understand and recognize the difference between discrimination that is based on group cognition and discrimination that is driven by group identification and emotion.

Some compelling evidence for the distinction between stereotype-based processes and ingroup–outgroup processes comes from recent research on implicit cognition. Initial research on behavioral priming (Bargh, Chen, & Burrows, 1996; Dijksterhuis, Bargh, & Miedema, 2000) indicated that priming the label or content of a social category automatically activates stereotype-related concepts, which in turn elicit behavioral responses that are assimilated to the activated concept. This assimilation to the category stereotype occurs even for persons who are not themselves members of the primed category. Thus, young college students who have been primed with terms associated with the elderly have been found to walk more slowly (Bargh et al., 1996), and intelligent college students primed with the category "hooligans" perform more poorly on a subsequent cognitive task than students primed with the category "professor" (Dijksterhuis & van Knippenberg, 1998) Such assimilative behavioral priming is apparently elicited by the mere activation of the mental representation of the social category, independent of its self-relevance.

More recent experiments, however, demonstrate that these priming effects are significantly altered when ingroup–outgroup categorization is made salient prior to the category priming experience. In this case, priming the outgroup category label elicits automatic behavioral *contrast* rather than assimilation (Schubert & Hafner, 2003; Spears, Gordijn, Dijksterhuis, & Stapel, 2004). Apparently, engaging the self-concept and social identities significantly changes responses to the category representation, even at this nonconscious level. Thus, the most current work on group cognition serves to validate historical differences within social psychology between the study of social cognition and the study of prejudice and intergroup relations. Although I agree that we have much to gain from better integration of these two literatures (e.g., Mackie & Smith, 1998) we also have something to lose if we fail to understand the differences as well.

For purposes of this chapter, I delimit the study of *intergroup* relations to research on attitudes, emotions, and behavior that are influenced by the salience of an ingroup–outgroup differentiation. This is consistent with Sherif's (1966) classic characterization of intergroup situations: "Whenever individuals belonging to one group interact, collectively or individually, with another group or its members in terms of their group identification, we have an instance of intergroup behavior" (p. 12). In line with this definition, this chapter discusses theories and research on intergroup relations under conditions in which awareness of ingroup membership and intergroup differentiation is a salient factor.

The Role of Mere Categorization in Intergroup Behavior

In a laboratory setting in Bristol, England, Henri Tajfel and his colleagues undertook initial experiments with the so-called "minimal intergroup situation" (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971). The results provided a powerful demonstration that merely classifying individuals into arbitrary distinct social categories was sufficient to produce ingroup–outgroup discrimination and bias, even in the absence of any interactions with fellow group members or any history of competition or conflict between the groups.

These laboratory studies succeeded in confirming the power of we–they distinctions to produce differential evaluation, liking, and treatment of other persons

depending on whether or not they are identified as members of the ingroup category. The laboratory experiments with the minimal intergroup situation demonstrated that ethnocentric loyalty and bias clearly do not depend on kinship or an extensive history of interpersonal relationships among group members but can apparently be engaged readily by symbolic manipulations that imply shared attributes or common fate. What appears to be critical for ingroup attachment is that there be a basis for distinctive identification of who is "us" and who is "them"—a rule of exclusion as well as inclusion.

Differentiation between ingroups and outgroups appears to be a necessary condition for intergroup discrimination, but categorization alone does not necessarily carry evaluative implications. Acknowledging one's ingroup membership is not equivalent to identification with that social group. This is clear from Tajfel's (1981) definition of social identity as "that part of an individual's self-concept which derives from his knowledge of his membership of a social group . . . *together with* the value and emotional significance attached to that membership" (p. 255, emphasis added). Social identification, then, entails affective and evaluative processes that are above and beyond mere cognitive classification of the self and others into a shared social category. The affective significance of social identification arises from the felt attachment between the self and the ingroup as a whole (cf. Oyserman, Chapter 18, this volume).

Social identification represents the extent to which the ingroup has been incorporated into the sense of self and, at the same time, that the self is experienced as an integral part of the ingroup. With high levels of social identification, the group's outcomes and welfare become closely connected to one's own sense of well-being (Brewer, 1991). According to social identity theory, it is this engagement of the self that accounts for the positive valuation of the ingroup and positive orientations toward fellow ingroup members. Self-worth is both projected onto and derived from positive ingroup evaluation. The distinction between ingroup membership and ingroup identification implies that social identification is not an automatic by-product of categorization. Identification implies some *motivation* to define oneself in terms of that group membership and to achieve and maintain inclusion in the ingroup category. A deeper understanding of the nature and consequences of this attachment of individuals to their ingroups is essential to an understanding of relationships between groups.

PRINCIPLE 2: THE PRIMACY OF INGROUP IDENTIFICATION

> Although we could not perceive our own in-groups excepting as they contrast to out-groups, still the in-groups are psychologically primary. . . . Hostility toward out-groups helps strengthen our sense of belonging, but it is not required. . . . The familiar is *preferred*. What is alien is regarded as somehow inferior, less "good," but there is not necessarily hostility against it. . . . Thus, while a certain amount of predilection is inevitable in all in-group memberships, the reciprocal atti-

tude toward out-groups may range widely. (Allport, 1954, p. 42, original emphasis)

Allport's (1954) chapter, "Ingroup Formation" (from which the foregoing quotation is taken), is one of the less cited sections of his classic book *The Nature of Prejudice*, but it warrants closer attention as a precursor to later research on ingroup bias and intergroup discrimination. In this chapter, Allport postulated that ingroups are "psychologically primary" in the sense that familiarity, attachment, and preference for one's ingroups come prior to development of attitudes toward specific outgroups.

Recognizing that ingroup attitudes and intragroup relationships may be independent of attitudes and behavior toward outgroups leads to a new approach to research on intergroup relations. If we take Allport's insight about the primacy of ingroup orientations seriously, we must first come to a better understanding of how and why ingroups are formed and why individuals exhibit ingroup loyalty, identification, and attachment in the first place. The second question is why and under what conditions does the formation and maintenance of ingroups and ingroup loyalty lead to negative relationships with outgroups.

Ethnocentrism and Ingroup Positivity

Along with proclaiming the psychological primacy of ingroup membership as the foundation of intergroup behavior, Allport (1954) also noted that "a certain amount of predilection is inevitable in all in-group memberships" (p. 42). This idea that ingroups are inevitably positively regarded accords with the concept of "ethnocentrism" as introduced by Sumner (1906) several decades earlier. Ethnocentrism was described by Sumner as a universal characteristic of human social groups whereby

> a differentiation arises between ourselves, the we-group, or in-group, and everybody else, or the others-group, outgroups. The insiders in a we-group are in a relation of peace, order, law, government, and industry, to each other. . . . Ethnocentrism is the technical name for this view of things in which one's own group is the center of everything, and all others are scaled and rated with reference to it. . . . Each group nourishes its own pride and vanity, boasts itself superior, exalts its own divinities, and looks with contempt on outsiders. . . . (pp. 12–13)

Experimental evidence from research in the minimal group paradigm demonstrated just how powerfully mere social categorization can influence differential thinking, feeling, and behaving toward ingroup versus outgroup members. Upon social categorization of individuals into ingroups and outgroups, people spontaneously experience more positive affect toward the ingroup. They also favor ingroup members directly in terms of evaluations and resource allocations (Tajfel et al., 1971), as well as indirectly in valuing the products of their work (Dustin & Davis, 1970; Ferguson & Kelley, 1964). In addition, ingroup membership increases the psychological bond and feelings of "oneness" that facilitate the arousal of

promotive tension or empathy in response to others' needs or problems (Hornstein, 1976) so that prosocial behavior is offered more readily to ingroup than to outgroup members (Piliavin, Dovidio, Gaertner, & Clark, 1981). Moreover, people are more likely to be cooperative and exercise more personal restraint when using endangered common resources when these are shared with ingroup members than with others (Brewer & Kramer, 1986; DeCremer & van Dijk, 2002; DeCremer & van Vugt, 1999), and they work harder for groups they identify more as their ingroup (Worchel, Rothgerber, Day, Hart, & Butemeyer, 1998).

Engaging social identities also influences perceptions of social justice and justice motives in general. Recent research provides evidence that group identification bounds the scope of people's concerns for distributive justice. For instance, Wenzel (2000) found that only East and West Germans who strongly identified with the inclusive category "German" were likely to apply norms of equality to entitlements of all citizens (east and west) of the country. Similarly, Tyler and his colleagues have found that individuals who are highly identified with a group or organization are more likely to defer to authorities and to be concerned with procedural fairness (Huo, Smith, Tyler, & Lind, 1996; Tyler, Degoey, & Smith, 1996). And people appear to be *less* concerned about distributive or procedural justice for actions related to outgroups in intergroup contexts than they are in interpersonal contexts (Bruins, Platow, & Ng, 1995; Platow et al., 1997).

When a particular social identity is made salient, individuals are likely to think of themselves as having characteristics that are representative of that social category. Social identity, in other words, leads to self-stereotyping (Hogg & Turner, 1987; Simon & Hamilton, 1994). Group stereotypical traits that may not be particularly relevant to personal identity become central to the self-concept when social identity is activated. When a woman's female identity is not salient, she may think of herself in terms of personal traits that are not relevant to the masculinity–femininity distinction (e.g., as organized, neat, and politically conservative). When her identity as a member of the female category is made salient, however, this same woman may think about herself in terms of those characteristics that make her more like other women and distinct from most men (e.g., nurturant and dependent). When an ingroup identity is salient, traits for which the self-concept and the group stereotype overlap become particularly accessible (Coats, Smith, Claypool, & Banner, 2000; Smith & Henry, 1996).

The hallmark of ingroup identification is ingroup positivity. There is even ample evidence that positive affect and evaluation are activated automatically by an ingroup label or whenever a group (even a minimal group) is associated with the self (Farnham, Greenwald, & Banaji, 1999; Otten & Wentura, 1999; Purdue, Dovidio, Gurtman, & Tyler, 1990; Rudman, Greenwald, & McGhee, 2001). This does not mean, however, that ingroup evaluations are indiscriminately positive on all dimensions of assessment. When there is objective evidence of outgroup achievement or a consensual status hierarchy in which the outgroup is recognized to be of higher status than the ingroup, then some degree of outgroup positivity (relative to the ingroup) is frequently obtained (Jost, 2001). However, ingroup positivity is consistently found on traits or attributes that are self-defining or self-relevant (Otten, 2002), and on traits reflecting basic social values (e.g., warmth, trustworthiness, and cooperativeness) ingroup positivity appears to be essentially universal (Brewer, 2001; LeVine & Campbell, 1972).

Ingroup Formation: The Boundaries of Cooperation and Trust

Ingroup identification implies some *motivation* to define oneself in terms of that group membership and to achieve and maintain inclusion in the ingroup category. Because identification entails some sacrifice of an autonomous self-concept, the question arises as to why individuals would attach some measure of their self-worth and well-being to the fate of a collective.

Theoretical understanding of ingroup identification starts from the recognition that group living represents the fundamental survival strategy that characterizes the human species. In the course of our evolutionary history, humans abandoned most of the physical characteristics and instincts that make possible survival and reproduction as isolated individuals or pairs of individuals, in favor of other advantages that require cooperative interdependence with others in order to survive in a broad range of physical environments. In other words, as a species we have evolved cooperation rather than strength and social learning rather than instinct as basic adaptations (cf. Caporael, Chapter 1, this volume).

Given the morphology and ecology of evolving hominids, the interface between hominids and their habitat must have been a group process. Finding food, defense from predation, moving across a landscape—these matters of coping with the physical habitat—are largely group processes. Over time, if exploiting a habitat is more successful as a collective group process than as an individual process, then not only would more successful groups persist but so also would individuals better adapted to group living. The result would be a shift to cooperative groups as the selective context for uniquely human mental systems. The result of selection *in* groups would be the evolution of perceptual, affective, and cognitive processes that support the development and maintenance of group membership (Caporael, Dawes, Orbell, & van de Kragt, 1989). Without a group, the probability of reproduction and survival to reproductive age is lowered for humans.

The result is that as a species, human beings are characterized by *obligatory interdependence* (Brewer, 1997; Caporael, 1997). For long-term survival, we must be willing to rely on others for information, aid, and shared resources, and we must be willing to give information and aid and to share resources with others. At the individual level, the potential benefits (receiving resources from others) and costs (giving resources to others) of mutual cooperation go hand in hand and set natural limits on cooperative interdependence. The decision to cooperate

(to expend resources to another's benefit) is a dilemma of trust since the ultimate benefits depend on everyone else's willingness to do the same (Brewer, 1981). A cooperative system requires that trust dominate over distrust. But indiscriminate trust (or indiscriminate altruism) is not an effective individual strategy; cooperation must be contingent upon the probability that others will cooperate as well.

Social categorization and clear group boundaries provide one mechanism for achieving the benefits of cooperative interdependence without the risk of excessive costs. Ingroup membership is a form of contingent altruism. By limiting aid to mutually acknowledged ingroup members, total costs and risks of nonreciprocation can be contained (see Takagi, 1996, for a related argument). Thus, ingroups can be defined as bounded communities of mutual trust and obligation that delimit mutual interdependence and cooperation.

An important aspect of this mutual trust is that it is *depersonalized* (Brewer, 1981), extended to any member of the ingroup whether personally related or not. Psychologically, expectations of cooperation and security promote positive attraction toward other ingroup members and motivate adherence to ingroup norms of appearance and behavior that ensure that one will be recognized as a good or legitimate ingroup member. Symbols and behaviors that differentiate the ingroup from local outgroups become particularly important here, to reduce the risk that ingroup benefits will be inadvertently extended to outgroup members, and to ensure that oneself is recognized as a member of the ingroup and entitled to those benefits. Assimilation within and differentiation between groups is thus mutually reinforcing, along with ethnocentric preference for ingroup interactions and institutions.

A consequence of ingroup identification and intergroup boundaries is that individuals modify their social behavior depending on whether they are interacting with ingroup or outgroup members. Ingroup behavior is governed by norms and sanctions that reinforce expectations of mutual cooperation and trustworthiness. Depersonalized trust is supported by implicit understandings that ingroup members will monitor the behavior and interactions of other group members, sanctioning deviations from group expectations about appropriate ingroup attitudes and behavior. Thus, shared ingroup membership may be taken as prima facie evidence that other members of the group will live by the codes of conduct that bind them together as a group (Kramer, Brewer, & Hanna, 1996).

Optimal Distinctiveness Theory

Cooperative groups must meet certain structural requirements in order to exist, just as organisms must have certain structural properties in order to be viable. For community-sized groups these organizational imperatives include mobilization and coordination of individual effort, communication, internal differentiation, optimal group size, and boundary definition. The benefits to individuals of cooperative arrangements cannot be achieved unless prior conditions have been satisfied that make the behavior of other individuals predictable and coordinated. Group survival depends on successful solution to these problems of internal organization and coordination.

The advantage of extending social interdependence and cooperation to an ever wider circle of conspecifics comes from the ability to exploit resources across an expanded territory and buffer the effects of temporary depletions or scarcities in any one local environment. But expansion comes at the cost of increased demands on obligatory sharing and regulation of reciprocal cooperation and free riding. Both the carrying capacity of physical resources and the capacity for distribution of resources, aid, and information inevitably constrain the potential size of cooperating social networks. Thus, effective social groups cannot be either too small or too large. To function, social collectives must be restricted to some optimal size—sufficiently large and inclusive to realize the advantages of extended cooperation, but sufficiently exclusive to avoid the disadvantages of spreading social interdependence too thin.

Based on this analysis of one structural requirement for group survival, Brewer (1991) hypothesized that the conflicting benefits and costs associated with expanding group size would have shaped social motivational systems at the individual level. If humans are adapted to live in groups and depend on group effectiveness for survival, our motivational systems should be tuned to the requirements of group effectiveness. We should be uncomfortable depending on groups that are too small to provide the benefits of shared resources but also uncomfortable if group resources are distributed too widely. A unidirectional drive for inclusion would not have been adaptive without a counteracting drive for differentiation and exclusion. Opposing motives hold each other in check, with the result that human beings are not comfortable either in isolation or in huge collectives. These social motives at the individual level create a propensity for adhering to social groups that are both bounded and distinctive. As a consequence, groups that are optimal in size are those that will elicit the greatest levels of member loyalty, conformity, and cooperation, reflecting the fit between individual psychology and group structure.

The optimal distinctiveness model of social identity (Brewer, 1991) was based on this evolutionary perspective on the functions of ingroup formation and differentiation. The theory assumes that psychological mechanisms at the individual level (opposing needs for differentiation and inclusion) coevolved with structural requirements at the group level for successful group functioning and coordination (size and boundedness). The model posits that humans are characterized by two opposing needs that govern the relationship between the self-concept and membership in social groups. The first is a need for assimilation and inclusion, a desire for belonging that motivates immersion in social groups. The second is a need for differentiation from others that operates in opposition to the need for immersion. As group membership becomes more and more inclusive, the need for inclusion is satisfied but the need for differenti-

ation is activated; conversely, as inclusiveness decreases, the differentiation need is reduced but the need for assimilation is activated. These competing drives ensure that interests at one level are not consistently sacrificed to interests at the other. According to the model, the two opposing motives produce an emergent characteristic—the capacity for social identification with distinctive groups that satisfy both needs simultaneously.

The relative strength of inclusion and differentiation needs at the individual level interacts with group properties such as size and permeability to determine social identification and ingroup attachment and loyalty. The psychology of assimilation and differentiation limits the extent to which strong social identification can be indefinitely extended to highly inclusive, superordinate social groups or categories. Thus, one implication of optimal distinctiveness theory is that ingroup loyalty, and its concomitant depersonalized trust and cooperation, is most effectively engaged by relatively small, distinctive groups or social categories. Groups that satisfy these needs are those that provide the necessary boundaries on mutual obligation and trust. Optimal group identities become essential to a secure self-concept. Once group identification has been achieved, maintaining a secure sense of inclusion and boundedness becomes tantamount to protecting one's own existence. When optimal inclusion within a distinctive ingroup is not achieved, the sense of self is threatened and vulnerable.

Ingroup Preference Needs No Outgroups

The arguments in the previous sections make the case that there is no need to require intergroup conflict to account for ingroup formation and exclusion of outgroups. In fact, in light of both paleoanthropological and archaeological evidence, it makes little sense to see conflict as the source of ingroup formation. There is no reason to believe that early hominids lived under dense population conditions in which bands of people lived in close proximity with competition over local resources. Estimates of the total human population during the Middle Paleolithic are less than 1.5 million (Hassan, 1981). Group living was well established much earlier—2.5 million years ago by human ancestors—and complex sociality evolved early among primate ancestors (Foley, 1996). Early evidence of population packing occurs around 15,000 years ago (Alexander, 1989; Stiner, 2002), too recently to have been relevant to the origins of human sociality. As Alexander himself admits, there is no evidence of intergroup conflict in early human evolutionary history (Alexander, 1989). Given the costs of intergroup fighting combined with low population density, flight rather than fight would seem to be the strategy of choice for our distant ancestors.

The idea that ingroup cooperation is born of intergroup conflict is also inconsistent with contemporary research on social identity and intergroup relations (Brewer, 1999, 2001). Despite widespread belief that ingroup positivity and outgroup derogation are reciprocally related, empirical research demonstrates little consistent relation between the two. Indeed, results from both laboratory experiments and field studies indicate that variations in ingroup positivity and social identification do not systematically correlate with degree of bias or negativity toward outgroups (Brewer, 1979; Hinkle & Brown, 1990; Kosterman & Feshbach, 1989; Struch & Schwartz, 1989). For example, in a study of the reciprocal attitudes among 30 ethnic groups in East Africa, Brewer and Campbell (1976) found that almost all the groups exhibited systematic differential positive evaluation of the ingroup over all outgroups on dimensions such as trustworthiness, obedience, friendliness, and honesty. However, the correlation between degree of positive ingroup regard and social distance toward outgroups was essentially .00 across the 30 groups.

Experiments with the minimal intergroup situation also provided additional evidence that ingroup favoritism is prior to, and not necessarily associated with, outgroup negativity or hostility. Brewer (1979) reported that most minimal group studies that assessed ratings of the ingroup and outgroup separately found that categorization into groups leads to enhanced ingroup ratings in the absence of decreased outgroup ratings. Furthermore, the positive ingroup biases exhibited in the allocation of positive resources in the minimal intergroup situation (Tajfel et al., 1971) are essentially eliminated when allocation decisions involve the distribution of negative outcomes or costs (e.g., Mummendey et al., 1992), suggesting that individuals are willing to differentially benefit the ingroup compared to outgroups but are reluctant to harm outgroups more directly. In a more recent review of developmental studies on intergroup attitudes, Cameron, Alvarez, Ruble, and Fuligni (2001) similarly concluded that children tend to display a positivity bias toward their ingroup but no negativity toward the outgroup.

Subsequent research in both laboratory and field settings has come to acknowledge the important distinction between ingroup bias that reflects beneficence and positive sentiments toward the ingroup that are withheld from outgroups ("subtle" prejudice) and discrimination that reflects hostility, derogation, and intent to harm the outgroup ("blatant" prejudice) (Pettigrew & Meertens, 1995). This is not to say that ingroup-based discrimination is benign or inconsequential. Indeed, many forms of institutional racism and sexism are probably attributable to discrimination based on ingroup preference rather than prejudice against outgroups (Brewer, 1996). Nonetheless, the absence of positive regard and lack of trust for outgroups that is characteristic of most ingroup–outgroup differentiation can be conceptually and empirically distinguished from the presence of active hostility, distrust, and hate for outgroups that characterizes virulent prejudice.

PRINCIPLE 3: WHAT ACCOUNTS FOR OUTGROUP HATE?

In most social psychology textbook definitions of "prejudice," the concept is associated with negative attitudes or behaviors toward specific outgroups. In other words,

social psychological research on prejudice is focused on outgroup negativity, rather than ingroup positivity (though that is, too, a form of prejudice).

If ingroup identification and attachment does not inevitably entail disdain for or conflict with outgroups, then explaining outgroup hate and hostility requires processes that go beyond social identity per se (Turner & Reynolds, 2001). As Allport noted, although ingroup positivity appears to be inevitable, attitudes toward relevant outgroups can vary considerably—from generally positive to neutral to negative or hostile.

Systematic variation in the relationship between ingroup identification and outgroup attitudes has been demonstrated in recent survey studies of mutual attitudes among ethnic subgroups in South Africa (Duckitt, Callaghan, & Wagner, 2005) and New Zealand (Duckitt & Parra, 2004). The findings from these studies demonstrate clearly that the nature of the relationship between ingroup identification/positivity and outgroup negativity varies considerably depending on which outgroup is being assessed. Among ethnic subgroups, the ingroup positivity–outgroup positivity correlation ranged from positive (particularly between ethnic minorities) to zero to negative (particularly between dominant and subordinate groups). Furthermore, the nature of these correlations tended to be reciprocal for pairs of groups, suggesting that attitudes toward specific outgroups are shaped by the structure of intergroup relations within a particular societal context.

Types of Ingroup–Outgroup Relationships

Duckitt's findings are consistent with the idea that the relationship between ingroup identification and outgroup attitudes depends on the implications that the outgroup has for the ingroup. Consistent with the principle of primacy of the ingroup, attitudes toward specific outgroups should be shaped by the motivation to protect and enhance the ingroup. Within this framework we can distinguish three primary patterns of ingroup–outgroup relationships:

1. *Ingroup autonomy.* In some contexts, particular outgroups are essentially irrelevant to the well-being of the ingroup. Outgroups that are psychologically or physically very distant from the ingroup are a source of neither comparison nor threat. Under these circumstances, we would expect to find zero relationship between ingroup attachment and outgroup attitudes.

2. *Intergroup comparison.* When an outgroup is a relevant standard of comparison for ingroup evaluation, this gives rise to intergroup social competition (Turner, 1975). If ingroup worth is evaluated *relative to* an outgroup, then ingroup outcomes improve only as outgroup benefits decrease. In that case, it is not clear whether the motivation for discriminatory behavior is to benefit the ingroup (i.e., discrimination *for* us) or to disadvantage the outgroup (discrimination *against* them).

3. *Intergroup conflict.* Intergroup comparison and social competition provide an initial connection between valuing the ingroup and devaluing relevant outgroups.

In the literature on social motives (McClintock, 1972), however, a clear distinction is made between *competition*–the motivation to seek relative gain for the ingroup over others—and *aggression*–the motivation to harm the other as an end in itself. Following this distinction, we can distinguish between intergroup discrimination that is based solely on ingroup favoritism and discrimination or prejudice that entails an active component of outgroup derogation and aggression (Levin & Sidanius, 1999; Struch & Schwartz, 1989).

Social identity theory provides an adequate framework for understanding ingroup bias and intergroup discrimination of the positive types described earlier. Assimilation of the self to the ingroup accounts for ingroup positivity, and from there, the extension to relative enhancement of the ingroup over outgroups under conditions of intergroup comparison is not a big motivational leap. The story becomes more complicated, however, when one tries to use the same framework to account for more virulent outgroup hate and intergroup hostility. To justify aggression against outgroups in the interest of the ingroup, the very existence of the outgroup, or its goals and values, must be seen as a threat to the maintenance of the ingroup and to one's own social identity. Thus, understanding the relationship between ingroup identification and outgroup hostility requires understanding how the interests of the ingroup and those of the outgroup come to be perceived as in conflict.

One approach to conceptualizing how perceptions of ingroup–outgroup relations may lead to outgroup negativity is integrated threat theory (Stephan & Stephan, 2000). This model distinguishes four different sources of experienced threat from a specific outgroup: *realistic threats* (threats to the existence, power, or material well-being of the ingroup or ingroup members), *symbolic threats* (threats to the ingroup worldview arising from perceived group differences in morals, values, and standards), *intergroup anxiety* (personal fear or discomfort experienced in connection with actual or anticipated interactions with members of the outgroup), and *negative stereotypes* (beliefs about outgroup characteristics that imply unpleasant or conflictual interactions and negative consequences for the self or the ingroup). In a field test of this model, Stephan and colleagues (2002) found that ratings of realistic threat, symbolic threat, and intergroup anxiety were significant predictors of negative interracial attitudes and that these threat perceptions mediated the effects of other predictor variables such as ingroup identification, intergroup contact, and status differences.

Stephan's taxonomy of intergroup threat delineates the ways in which the very existence of a particular outgroup may be perceived as a danger to the ingroup, but it begs the question of how ingroup members come to see outgroups in this way. Except under conditions of realistic group conflict (i.e., life-and-death competition for scarce resources or open warfare), the perception that an outgroup constitutes a symbolic or reputational threat to the ingroup is highly subjective, and we still need more explicit theory of how these perceptions arise.

Next I suggest two factors that may promote or exacerbate the perception of outgroup threat.

Common Goals

The presence of realistic competition over scarce resources or other group goals is clearly a strong basis for intergroup conflict and hostility. In contrast, the presence of superordinate goals or common threat is widely believed to provide the conditions necessary for intergroup cooperation and reduction of conflict (e.g., Sherif, 1966). This belief is an extrapolation of the general finding that *intragroup* solidarity is increased in the face of shared threat or common challenge.

It may be true that loosely knit ingroups become more cohesive and less subject to internal factioning when they can be rallied to the demands of achieving a common goal. The dynamics of interdependence are quite different, however, in the case of highly differentiated social groups. Among members of the same ingroup, engaging the sense of trust necessary for cooperative collective action is essentially nonproblematic. In an intergroup context, however, perceived interdependence and the need for cooperative interaction makes salient the absence of mutual trust. Without the mechanism of depersonalized trust based on common identity, the risk of exploited cooperation looms large and distrust dominates over trust in the decision structure. It is for this reason that I have argued elsewhere (Brewer, 2000) that the *anticipation* of positive interdependence with an outgroup, brought on by perceptions of common goals or common threat, actually promotes intergroup conflict and hostility. When negative evaluations of the outgroup such as contempt are also already present, common threat in particular may promote scapegoating and blame rather than mutual cooperation.

Perceived positive interdependence with the outgroup also threatens intergroup differentiation (Branscombe, Ellemers, Spears, & Doosje, 1999; Jetten, Spears, & Postmes, 2004). To the extent that feelings of secure inclusion, ingroup loyalty, and optimal identity are dependent on the clarity of ingroup boundaries and intergroup distinctions, shared experiences and cooperation with the outgroup threaten the basis for social identification. Particularly for individuals who are exclusively vested in a single group identity, the threat of lost distinctiveness may override the pursuit of superordinate goals and lead to resistance to cooperation (collaboration) even at the cost of ingroup self-interest.

Power Politics

Moral superiority, distrust of outgroups, and social comparison are all processes that emerge from ingroup maintenance and favoritism and can lead to hostility and conflict between groups even in the absence of realistic conflict over material resources or power. When groups are political entities, however, these processes may be exacerbated through deliberate manipulation by group leaders in the interests of mobilizing collective action to secure or maintain political power. Social category differ-

entiation provides the fault lines in any social system that can be exploited for political purposes. When trust is ingroup based, it is easy to fear control by outsiders; perceived common threat from outgroups increases ingroup cohesion and loyalty; appeals to ingroup interests have greater legitimacy than appeals to personal self-interest. Thus politicization—an important mechanism of social change—can be added to the factors that may contribute to a correlation between ingroup love and outgroup hate.

Intergroup Emotions and Image Theory

The general idea that intergroup attitudes are shaped by the perceived relationship between the ingroup and outgroup (in particular whether the existence of the outgroup poses a threat to the ingroup or not) is consistent with recent theories of prejudice as intergroup emotion (Smith, 1993). It has long been recognized that intergroup attitudes are a complex mixture of affective, evaluative, and cognitive responses, but recent social psychological research has been dominated by a focus on the cognitive component (stereotypes) to the relative exclusion of the more emotional underpinnings of prejudice. Although feelings and emotions toward specific outgroups are certainly connected to beliefs about those groups, these components are conceptually and empirically separable (Esses, Haddock, & Zanna, 1993).

Zanna and Rempel (1988) have suggested that an attitude can be defined as an overall evaluation of a social group that is based on a combination of cognitive information (beliefs and stereotypes) and affective information (feelings and emotions). Attitudes toward different groups may reflect different weightings of affective and cognitive components. For some individuals, prejudice toward a specific group may be primarily emotion driven, whereas for other individuals or other groups, prejudice is primarily a matter of beliefs and values (Esses et al., 1993). Furthermore, emotional states may directly affect the beliefs and evaluations associated with an outgroup.

Emotional reactions to a particular outgroup can include positive emotions (e.g., admiration and respect) as well as a range of negative emotions (e.g., fear, disgust, anxiety, and hate). In Dijker's (1987) examination of the relation between emotions and attitudes toward two minority groups in the Netherlands, although both types of emotion predicted evaluation of the outgroups, positive emotions were more predictive of attitudes toward one group and negative emotions were more predictive of attitudes toward the other. Similarly, in an investigation of the relation between positive and negative emotional responses toward seven minority groups in the United States, both types of emotion predicted prejudice toward these groups (Stangor, Sullivan, & Ford, 1991, study 1).

Attitude Ambivalence

The emotional component of prejudice is made more complex when one realizes that positive and negative affect may be held independently (Cacioppo & Berntson,

1994; Larsen & Diener, 1992). If positive and negative emotions are potentially independent, the reduction of extreme negative affect toward a particular outgroup does not necessarily result in increased positive affect. Recently, Pettigrew and Meertens (1995) have suggested that, although "blatant" forms of prejudice against outgroups involve strong negative affect, more "subtle" forms of prejudice involve the absence of positive emotional reactions toward the outgroup rather than the presence of negative emotions. For many outgroups, prejudice comes in the form of relative indifference and exclusion from the benefits of positive affect and leniency accorded to ingroup members. Antipathy-based prejudice and discrimination occur only against outgroups that arouse strong negative emotions.

Prejudice may also be associated with emotional ambivalence when arousal of negative affect occurs in situations in which positive attitudes are desirable. The theory of ambivalence amplification (Hass, Katz, Rizzo, Bailey, & Eisenstadt, 1991; Katz & Hass, 1988) refers to a complex motivational state that applies particularly to the relation between a dominant, politically powerful group (such as Whites in the United States) and a less powerful or disadvantaged minority group (such as Blacks). It emphasizes the conflict between egalitarian values and antipathy toward the outgroup. On the one hand, for example, Whites hold pro-Black feelings that stem from the endorsement of egalitarianism and fairness as societal values and they simultaneously perceive that Blacks are disadvantaged. Concomitantly, these same Whites possess anti-Black feelings that arise from early socialization and stereotypes, including the perception that Blacks deviate from the traditional cultural values, such as the Protestant work ethic. In turn, these ambivalent feelings motivate exaggerated evaluations of Blacks. Thus, when a Black individual and a White individual exhibit the same positive behavior, the Black will be evaluated more positively than the White; conversely, a Black behaving negatively will be evaluated more harshly than a White who behaves in that same manner. And the more strongly a White American simultaneously holds these conflicting motives toward Blacks, the more extreme these evaluations will be.

The ambivalence associated with internal conflicts between egalitarian values and negative affect toward outgroups plays a role in a number of theories of prejudice and racism. Theories of "aversive" (Gaertner & Dovidio, 1986), "modern" (McConahay, 1986; McConahay & Hough, 1976), "symbolic" (Sears, 1988; Sears & Funk, 1991), or "regressive" racism (Rogers & Prentice-Dunn, 1981) argue that there is a conflict between egalitarian values and the emotional antipathy toward outgroup members that is often automatically activated. Among the propositions of aversive racism (Gaertner & Dovidio, 1986, 2000) is that the salience of egalitarian values influences the expression of racial bias. When the salience of egalitarian values is high, there will be no overt display of racial discrimination. When it is low, discrimination may appear, particularly in subtle forms. According to the model, prejudice influences behavior via an indirect process whereby the race of the other person enhances the salience of elements in a situation that would justify or rationalize a negative response to that individual. In interactions between ingroup and outgroup members, then, egalitarian values may typically exert pressure to behave in a nonprejudiced manner. However, when there is a justifying cause that can rationalize a negative response, that will legitimize negative behavior toward outgroup members, resulting in amplified negative responses and prejudice that is not recognized as such by the actor (see Dovidio & Gaertner, 2004, for review).

Another explanation for the consequences of ambivalence or aversive racism is that conflict of values affects resources available for cognitive processing (Richeson & Shelton, 2003). Devine (1989), for instance, argues that because prejudicial responses are well learned from early childhood, two processes are necessary for suppression of prejudiced behavior: (1) inhibition of automatically activated negative stereotypes and (2) activation of "controlled information processing" that makes salient those associations that are consistent with egalitarian values. Thus, circumstances in which egalitarian values are salient will allow an individual to inhibit negative associations and behave in a manner consistent with them. If, however, either time or processing capacity are limited or constrained, biased behavior emerges as a consequence of overlearned negative racial associations. This suggests that biased behavior can occur unintentionally, even in one who sincerely embraces egalitarianism.

Differentiating Negative Emotions: Appraisal Theory of Intergroup Emotions

In addition to distinguishing between positive and negative emotions as components of intergroup attitudes, researchers have begun to recognize the importance of distinguishing among different types of negative emotions in intergroup contexts. Distinct emotions reflect different underlying causes and lead to different types of behavior. Smith (1993) has suggested that five specific emotions are most likely to be aroused in intergroup situations: fear, disgust, contempt, anger, and jealousy. Of these, fear and disgust can be distinguished as emotions that imply avoidance or movement away from the outgroup, whereas contempt and anger imply movement against the outgroup (although fear can also elicit the attack response if the perceiver feels trapped or cornered and unable to effectively flee the source of fear). Attitudes that are driven by the former emotional states are likely to have different cognitive contents and behavioral implications than attitudes that are associated with the latter forms of emotion.

Mackie, Devos, and Smith (2000) demonstrated, across three empirical studies, that (1) for groups that are defined by a basic value conflict, anger and fear can be differentiated as distinct negative emotional responses to the outgroup, (2) appraisals of relative ingroup strength determine the degree of reported anger toward the outgroup, and (3) level of felt anger mediates the relationship between strength appraisals and participants' desire to confront, oppose, or attack members of the

outgroup. Based on these findings, Mackie and colleagues conclude that intergroup attitudes and behavior are channeled by the specific emotions that are elicited in response to appraisals of a particular outgroup in relation to the ingroup.

According to appraisal theories of emotion, the type of emotion directed toward outgroups may be a function of the degree of conflict of interest that is perceived to exist between the outgroup and the ingroup. When perceived conflict or threat is relatively low, negative emotions toward outgroups are likely to be associated with appraisals of status and legitimacy. The perception that the outgroup is different from the ingroup in ways that are devalued or illegitimate gives rise to feelings of moral superiority, intolerance, and concomitant emotions of contempt and disgust toward relevant outgroups. The emotions associated with moral superiority may justify some negative discrimination against outgroups but do not necessarily lead directly to hostility or conflict. In various contexts, groups have managed to live in a state of mutual contempt over long periods without going to war over their differences. The emotions of contempt and disgust are associated with avoidance rather than attack, so intergroup peace may be maintained through segregation and mutual avoidance.

As perceived conflict increases, avoidant emotions such as anxiety and disgust may be replaced by emotions such as anger, which instigate active hostility and aggression. This relationship was indicated by results of a study of the attitudes of Israeli citizens toward members of an ultraorthodox sect (Struch & Schwartz, 1989). Two types of intergroup attitudes were assessed, one measuring aggressive intent against the outgroup, the other measuring more conventional ingroup–outgroup evaluative ratings. The strongest predictors of ratings on the aggression measure were perceived conflicts of interests and values between the respondent's own religious group and the outgroup sect. However, these conflict measures did not predict the more traditional evaluative ratings, and these two measures of outgroup attitude were essentially uncorrelated.

Image Theory

One theoretical perspective that links perceptions and emotions in intergroup contexts is represented by "image theories" of international relations as developed by Cottam (1977) and Herrmann (1985). These theorists argue that images form as a *consequence* of strategic relationships between nations and serve a functional purpose. Working from Heider's (1958) balance theory, Herrmann suggested that images in international relations serve to balance a positive self-image and behavioral inclinations toward another nation. Herrmann argued that the nature of relationships between nations is represented in terms of perceived threats and perceived opportunities that give rise to emotions and associated behavioral implications. The cognitive system is balanced when the evaluative and behavioral implications of the image of the other nation matches the behavioral inclination generated from the threat or opportunity in a way

that maintains consistency with a positive, moral self-image (see also Sande, Goethals, Ferrari, & Worth, 1989). Thus, appraisals of the nature of the relationship between ingroup and outgroup (threat vs. opportunity) give rise to specific images of the outgroup nation that serve to account for and justify affective and behavioral orientations toward that group.

Herrmann (1985) identified three critical dimensions of intergroup relationships that give rise to sentiments and subsequent images of the other. These appraisal dimensions are (1) evaluations of competitive versus cooperative goal interdependence, (2) assessments of relative power, and (3) evaluations of relative cultural status. Together, these three assessments are the primary determinants of perceived threat and opportunity in intergroup relationships. When relationships are perceived to be characterized by extreme values on these three dimensions, emotions are aroused that determine both specific behavioral orientations toward the outgroup and specific content of the cognitive image of that group. Among the generic images that can arise from different configurations of intergroup relationship appraisals, five were identified as particularly relevant to international relations. These were labeled the "ally" image, the "enemy" image, the "barbarian" image, the "dependent (colonial)" image, and the "imperialist" image (Herrmann & Fischerkeller, 1995). Evidence from both laboratory experiments (Alexander, Brewer, & Herrmann, 1999) and field studies of interracial attitudes (Alexander, Brewer, & Livingston, 2005) support the hypothesis that relational appraisals can account for differences in outgroup perceptions and attitudes.

The influence of specific perceptions of the intergroup situation to create images and reactions is consistent with appraisal theories of intergroup emotions. In fact, the mediating role of emotions is implicit in the premises of image theory. First, the theory holds that appraisals of the pattern of intergroup relationships give rise to "sentiments" (Herrmann, 1985) that drive the behavioral orientations and consequent image activation. While the sentiments were not identified in terms of specific emotions, the idea that emotions correspond to specific action tendencies is integral to appraisal theories of emotion.

Based on this, a case can be made for incorporating an appraisal theory of emotions into image theory more formally (Brewer & Alexander, 2002). Information about the goals, status, and power relationships between ingroup and outgroup gives rise to appraisals along the dimensions of goal congruence, coping (relative ingroup strength and resources), and legitimacy (status evaluations). The resulting appraisal configuration elicits specific emotions (e.g., fear, anger, or resentment) that are associated with specific action tendencies (e.g., retreating, attacking, or resisting). In the context of intergroup relations, these emotional responses and associated action orientations are translated into cognitive representations of the outgroup (images) and preferences for behavior and policies toward that group. Table 30.1 summarizes this emotion appraisal model with respect to five specific intergroup configurations that have been drawn

TABLE 30.1. Image Theory of Intergroup Emotions

Relationship pattern	Intergroup emotion	Outgroup image	Action tendency
Goal compatibility Status equal Power equal	Admiration Trust	Ally	Cooperate
Goal incompatibility Status equal Power equal	Anger	Enemy	Contain or compete
Goal independent Status lower Power lower	Disgust Contempt	Dependent	Exploit or protect
Goal incompatibility Status lower Power higher	Fear Intimidation	Barbarian	Defend
Goal independent Status higher Power higher	Jealousy Resentment	Imperialist	Rebel

Note. From Brewer and Alexander (2002). Copyright 2002 by Taylor & Francis. Adapted by permission.

from image theory in international relations. By combining political theory with recent advances in the social psychology of intergroup emotions, we have a much more nuanced picture of the cognitive and affective underpinnings of different forms of intergroup behavior.

APPLYING THE PRINCIPLES: PREJUDICE REDUCTION THROUGH CONTACT AND COOPERATION

When ingroup identification leads to outgroup derogation, hostility, or overt conflict, principles for understanding the origins of intergroup prejudice must be turned toward the issue of how to reduce the negative consequences of ingroup–outgroup differentiation. Broadly speaking, there are two approaches to changing intergroup attitudes that have been derived from social psychological theory and research. One approach follows from principles of social categorization and focuses on altering ingroup–outgroup categorization. A second approach follows from intergroup emotion theory and involves changing appraisals of the outgroup vis-à-vis the ingroup.

Social Categorizaion and the Contact Hypothesis

The "contact hypothesis" is a general set of ideas about reducing intergroup prejudice and discrimination that developed among social scientists in the 1940s in the context of interracial relations in the United States (Allport, 1954; Watson, 1947; Williams, 1947). The basic idea behind the hypothesis is that hostility between groups is fed by unfamiliarity and separation and that *under the right conditions*, contact among members of different groups will reduce hostility and promote more positive intergroup attitudes. According to Allport (1954), the four most important of these qualifying conditions were that (1) integration has the support of authority, fostering *social norms* that favor intergroup acceptance; (2) the situation has high "acquaintance potential," promoting *intimate contact among members of both groups*; (3) the contact situation promotes *equal status* interactions among members of the social groups; and (4) the situation creates conditions of *cooperative interdependence* among members of both groups. Each of these conditions was derived from results of early research on racial desegregation and intergroup contact in the United States, on which the hypothesis was initially based.

Although the principle underlying the contact hypothesis was based on both empirical research and practical experience, research on the effects of contact tended to be largely atheoretical and driven by the exigencies of real-world desegregation settings. One advance toward a more integrative theory of prejudice reduction was achieved when contact research was combined with concepts of social categorization and social identity theory to provide a theoretical framework for understanding the cognitive mechanisms by which cooperative contact is presumed to work (see Brewer & Miller, 1984; Brown & Hewstone, 2005; Gaertner, Mann, Murrell, & Dovidio, 1989; Hewstone, 1996; Hewstone & Brown, 1986; Wilder, 1986). From the social categorization perspective, the issue to be addressed is how intergroup contact and cooperation can be structured to alter cognitive representations in ways that would eliminate one or more of the basic features of the negative intergroup schema.

Based on the premises of social identity theory, three alternative models for contact effects have been developed and tested in experimental and field settings, namely, decategorization, recategorization, and intergroup contact. Each of these models can be described in terms of (1) the structural representation of the contact situation that is recommended, (2) the psychological processes that promote attitude change within the contact setting, and (3) the mechanisms by which contact experiences are generalized to changed attitudes toward the outgroup as a whole. The first two models seek to change attitudes and perceptions by altering the salience of ingroup–outgroup social categorization in the contact situation. The third model addresses how intergroup attitudes can be changed while ingroup–outgroup differentiation remains salient.

Decategorization: The Personalization Model

The first model is essentially a formalization and elaboration of the assumptions implicit in the contact hypothesis itself (Brewer & Miller, 1984). A primary consequence of salient ingroup–outgroup categorization is the deindividuation of members of the outgroup. Social behavior in category-based interactions is characterized by a tendency to treat individual members of the outgroup as undifferentiated representatives of a unified social cate-

gory, ignoring individual differences within the group. The personalization perspective on the contact situation implies that intergroup interactions should be structured to reduce the salience of category distinctions and promote opportunities to get to know outgroup members as individual persons.

The conditional specifications of the contact hypothesis (equal status, intimate, cooperative interaction) can be interpreted as features of the situation that reduce category salience and promote more differentiated and personalized representations of the participants in the contact setting. Attending to personal characteristics of group members not only provides the opportunity to disconfirm category stereotypes but also breaks down the monolithic perception of the outgroup as a homogeneous unit (Wilder, 1978). In this scheme, the contact situation encourages attention to information at the individual level that replaces category identity as the most useful basis for classifying participants.

Repeated personalized contacts with a variety of outgroup members should, over time, undermine the value and meaningfulness of the social category stereotype as a source of information about members of that group. This is the process by which contact experiences are expected to generalize—via reducing the salience and meaning of social categorization in the long run (Brewer & Miller, 1988).

A number of experimental studies provide evidence supporting this perspective on contact effects (Bettencourt, Brewer, Croak, & Miller, 1992; Marcus-Newhall, Miller, Holtz, & Brewer, 1993). Miller, Brewer, and Edwards (1985), for instance, demonstrated that a cooperative task that required personalized interaction with members of the outgroup resulted not only in more positive attitudes toward outgroup members in the cooperative setting but also toward other outgroup members shown on a videotape, compared to cooperative contact that was task focused rather than person focused.

The personalization model is also supported by the early empirical evidence for the effects of extended, intimate contact on racial attitudes, reviewed previously. More recently, extensive data on effects of intergroup friendships have been derived from surveys in Western Europe regarding attitudes toward minority immigrant groups (Hamberger & Hewstone, 1997; Pettigrew, 1997; Pettigrew & Meertens, 1995). Across samples in France, Great Britain, the Netherlands, and Germany, Europeans with outgroup friends scored significantly lower on measures of prejudice, particularly affective prejudice (Pettigrew, 1998). This positive relationship did not hold for other types of contact (work or residential) that did not involve formation of close personal relationships with members of the outgroup. Although there is clearly a bidirectional relationship between positive attitudes and extent of personal contact, path analyses indicate that the path from friendship to reduction in prejudice is stronger than the other way around (Pettigrew, 1998).

Other recent research also reveals two interesting extensions of the personalized contact effect. One is evidence (again from European survey data) that personal friendships with members of one outgroup may lead to tolerance toward outgroups in general and reduced nationalistic pride—a process that Pettigrew (1997) refers to as "deprovincialization." A second extension is represented by evidence that contact effects may operate indirectly or vicariously. Although interpersonal friendship across group lines leads to reduced prejudice, even knowledge that an ingroup member befriended an outgroup member has potential to reduce bias (Wright, Aron, McLaughlin-Volpe, & Ropp, 1997).

Recategorization: The Common Ingroup Identity Model

The second social categorization model of intergroup contact and prejudice reduction is also based on the premise that reducing the salience of ingroup–outgroup category distinctions is key to positive effects. In contrast to the decategorization approaches described earlier, recategorization is not designed to reduce or eliminate categorization but rather to structure a definition of group categorization at a higher level of category inclusiveness in ways that reduce intergroup bias and conflict (Allport, 1954, p. 43). Specifically, the common ingroup identity model (Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993; Gaertner, Dovidio, Nier, Ward, & Banker, 1999) proposes that intergroup bias and conflict can be reduced by factors that transform participants' representations of memberships from two groups to one more inclusive group. With common ingroup identity, the cognitive and motivational processes that initially produced ingroup favoritism are redirected to benefit the former outgroup members.

Among the antecedent factors proposed by the common ingroup identity model are the features of contact situations (Allport, 1954) that are necessary for intergroup contact to be successful (e.g., interdependence between groups, equal status, and equalitarian norms). From this perspective, cooperative interaction, for example, enhances positive evaluations of outgroup members, at least in part, because cooperation transforms members' representations of the memberships from "us" and "them" to a more inclusive "we."

To test this hypothesis directly, Gaertner, Mann, Dovidio, Murrell, and Pomare (1990) conducted a laboratory experiment that brought two three-person laboratory groups together under conditions designed to vary independently the members' representations of the aggregate as one group or two groups (by varying factors such as seating arrangement) and the presence or absence of intergroup cooperative interaction. Supportive of the hypothesis concerning how cooperation reduces bias, among participants induced to feel like two groups, the introduction of cooperative interaction increased their perceptions of one group and also reduced their bias in evaluative ratings relative to those who did not cooperate during the contact period. Also supportive of the common ingroup identity model, reduced bias associated with introducing cooperation was due to enhanced favorable evaluations of outgroup members. In further support for the common ingroup identity model, this ef-

fect of cooperation was mediated by the extent to which members of both groups perceived themselves as one group.

Outside the laboratory, survey studies conducted in natural settings across very different intergroup contexts offered converging support for the proposal that the features specified by the contact hypothesis can increase intergroup harmony in part by transforming members' representations of the memberships from separate groups to one more inclusive group. Participants in these studies included students attending a multiethnic high school (Gaertner, Rust, Dovidio, Bachman, & Anastasio, 1994), banking executives from a wide variety of institutions across the United States who had experienced a corporate merger (Bachman, 1993), and college students who were members of blended families whose households were composed of two formerly separate families trying to unite into one (Banker & Gaertner, 1998). Results of these studies showed that the more the aggregate felt like one group, the lower the bias in affective reactions in the high school, the less the intergroup anxiety for the bankers, and the greater the amount of step-family harmony (Gaertner, Dovidio, & Bachman, 1996).

Challenges to the Decategorization/ Recategorization Models

Although the structural representations of the contact situation advocated by the decategorization (personalization) and recategorization (common ingroup identity) models are different, the two approaches share common assumptions about the need to reduce category differentiation and associated processes. Because both models rely on reducing or eliminating the salience of intergroup differentiation, they involve structuring contact in a way that will challenge or threaten existing social identities. Both cognitive and motivational factors conspire to create resistance to the dissolution of category boundaries or to reestablish category distinctions across time. Although the salience of a common superordinate identity or personalized representations may be enhanced in the short run, these may be difficult to maintain across time and social situations.

Preexisting social–structural relationships between groups may also create strong forces of resistance to changes in category boundaries. Cognitive restructuring may be close to impossible (at least as a first step) for groups already engaged in deadly hostilities. Even in the absence of overt conflict, asymmetries between social groups in size, power, or status create additional sources of resistance. When one group is substantially numerically smaller than the other in the contact situation, the minority category is especially salient and minority group members may be particularly reluctant to accept a superordinate category identity that is dominated by the other group. Another major challenge is created by preexisting status differences between groups, where members of both high- and low-status groups may be threatened by contact and assimilation (Mottola, 1996).

The Intergroup Contact Model: Changing Appraisals of the Ingroup–Outgroup Situation

These challenges to processes of decategorization/ recategorization led Hewstone and Brown (1986; Brown & Hewstone, 2005) to recommend an alternative approach to intergroup contact wherein cooperative interactions between groups are introduced without degrading the original ingroup–outgroup categorization. Specifically, this model favors encouraging groups working together to perceive complementarity by recognizing and valuing mutual superiorities and inferiorities within the context of an interdependent cooperative task or common, superordinate goals. This strategy allows group members to maintain their social identities and positive distinctiveness while avoiding insidious intergroup comparisons. Thus, the intergroup contact model does not seek to change the basic category structure of the intergroup contact situation but to change the intergroup affect from negative to positive interdependence and evaluation.

To promote positive intergroup experience, Hewstone and Brown recommend that the contact situation be structured so that members of the respective groups have distinct but complementary roles to contribute toward common goals. In this way, both groups can maintain positive distinctiveness within a cooperative framework. Evidence in support of this approach comes from the results of an experiment by Brown and Wade (1987) in which work teams composed of students from two different faculties engaged in a cooperative effort to produce a two-page magazine article. When the representatives of the two groups were assigned separate roles in the team task (one group working on figures and layout, the other working on text), the contact experience had a more positive effect on intergroup attitudes than when the two groups were not provided with distinctive roles (see also Deschamps & Brown, 1983; Dovidio, Gaertner, & Validzic, 1998).

Hewstone and Brown (1986) argued that generalization of positive contact experiences is more likely when the contact situation is defined as an *intergroup* situation rather than an interpersonal interaction. Generalization in this case is direct rather than requiring additional cognitive links between positive affect toward individuals and representations of the group as a whole. This position is supported by evidence that cooperative contact with a member of an outgroup leads to more favorable generalized attitudes toward the group as a whole when category membership is made salient during contact (e.g., Brown, Vivian, & Hewstone, 1999: Hewstone et al., 2005; van Oudenhoven, Groenewoud, & Hewstone, 1996).

Although ingroup–outgroup category salience is usually associated with ingroup bias and the negative side of intergroup attitudes, cooperative interdependence is assumed to override the negative intergroup schema, particularly if the two groups have differentiated, complementary roles to play. Because it capitalizes on needs for distinctive social identities, the intergroup contact model

provides a solution that is highly stable in terms of the cognitive–structural aspects of the intergroup situation. The affective component of the model, however, is likely to be more unstable. Salient intergroup boundaries are associated with mutual distrust (Insko & Schopler, 1987), which undermines the potential for cooperative interdependence and mutual liking over any length of time. By reinforcing perceptions of group differences, the differentiation model risks reinforcing negative beliefs about the outgroup in the long run and intergroup anxiety (Greenland & Brown, 1999; Islam & Hewstone, 1993), and the potential for fission, and conflict along group lines remains high.

Hybrid Models: An Integration of Approaches

As reviewed previously, each of the cognitive–structural models of intergroup contact and prejudice reduction has its weaknesses and limitations, particularly when one seeks to generalize beyond small group interactions in laboratory settings. These criticisms have led a number of writers to suggest that some combination of all three models may be necessary to create conditions for long-term attitude change (e.g., Brewer, 1996; Brown & Hewstone, 2005; Gaertner et al., 2000; Hewstone, 1996; Pettigrew, 1998).

More integrative models of intergroup contact take advantage of the fact that individuals are members of multiple social groups, which imply different social identities and ingroup loyalties. Traditionally, social identities have been treated as if they were mutually exclusive, with only one social categorization (ingroup–outgroup differentiation) salient at any one time. New research has begun to challenge this assumption of exclusivity and to explore the implications of holding multiple group identities, or identities at different levels of inclusiveness, simultaneously.

Hierarchical Dual Identities

In recent work regarding the development of a common ingroup identity, it has been proposed that embracing a more inclusive superordinate identity does not necessarily require each group to forsake its original group identity completely (Gaertner et al., 1990, 1994). In many contexts this may be impossible or undesirable. In some intergroup contexts, however, when members simultaneously perceive themselves as members of different groups but also as part of the same team or superordinate entity, intergroup relations between these subgroups are more positive than if members only considered themselves separate groups (Brewer & Schneider, 1990). For example, minority students in the multiethnic high school who identified themselves using both a minority subgroup and an American superordinate identity had lower intergroup bias than those students who identified themselves using only their minority group identity (Gaertner et al., 1994). Also, the greater the extent to which majority and minority students perceived the study body as "different groups . . . all playing on the

same team" (the dual identity item), the lower their degree of intergroup bias. By contrast, the more they conceived of the student body as "belonging to different groups" the higher the intergroup bias.

Other research also supports the value of a dual identity for reducing bias and improving intergroup relations. Two studies further suggest that the intergroup benefits of a strong superordinate identity remain relatively stable even when the strength of the subordinate identity becomes equivalently high (Huo et al., 1996; Smith & Tyler, 1996). This suggests that identification with a more inclusive social group does not require individuals to deny their ethnic identity. In addition, a dual identity can also lead to even more positive outgroup attitudes than those associated with a superordinate identity alone (Hornsey & Hogg, 2000). In terms of promoting more harmonious intergroup interactions, a dual identity capitalizes on the benefits of common ingroup membership as well those accrued from mutual differentiation between the groups.

On the other hand, dual identities are not always associated with positive relations between subgroups within the superordinate category. Mummendey and Wenzel (1999) make a convincing case that under some circumstances, making a shared superordinate category salient can lead to enhanced derogation of other subgroups when both subgroup and superordinate group identities are salient. This can happen if the values and attributes of the ingroup are projected onto the superordinate group, in which case subgroups that differ from these attributes come to be seen as *deviant* (rather than just "different") and a potential source of symbolic threat to the ingroup and the superordinate. In studies of national groups in the European Union, Mummendey and Waldzus (2004) have demonstrated that individuals who profess dual identification also exhibit higher ingroup projection, which in turn mediates negative attitudes toward other subgroup nations.

Cross-Cutting Identities

Embedded categories at different levels of inclusiveness represent only one form of multiple ingroup identities. Individuals may also be members of social categories that overlap only partially, if at all. Many bases of social category differentiation—gender, age, religion, ethnicity, occupation—represent cross-cutting cleavages. From the standpoint of a particular person, other individuals may be fellow ingroup members on one dimension of category differentiation but outgroup members on another. (For instance, for a woman business executive, a male colleague is an ingroup member with respect to occupation but an outgrouper with respect to her gender identification.) It is possible that such orthogonal social identities are kept isolated from each other so that only one ingroup–outgroup distinction is activated in a particular social context. But there are reasons to expect that simultaneous activation of multiple ingroup identities is both possible and has potential for reducing prejudice and discrimination based on any one category distinction.

Evidence from both anthropology (e.g., Gluckman, 1955) and political sociology (e.g., Coser, 1956) has long suggested that societies characterized by cross-cutting loyalty structures are less prone to schism and internal intergroup conflict than societies characterized by a single hierarchical loyalty structure. More recently, social psychologists have also begun to consider the implications of such multiple cross-cutting social identities for reduction of ingroup bias at the individual level (Brown & Turner, 1979; Deschamps & Doise, 1978; Marcus-Newhall et al., 1993; Vanbeselaere, 1991).

Experimental studies with both natural and artificial categories have demonstrated that adding a cross-cutting category distinction reduces ingroup bias and increases positive attitudes toward crossed category members compared to simple ingroup–outgroup differentiation (Vanbeselaere, 1991) or compared to situations in which category distinctions are convergent or superimposed (Bettencourt & Dorr, 1998; Marcus-Newhall et al., 1993; Rust, 1996). In these studies, cooperative interaction in the context of cross-cutting social identities and roles increases intracategory differentiation and reduces perceived intercategory differences, resulting in less category-based evaluations of individual group members. Furthermore, the benefits of cross-categorization may be enhanced when both category distinctions are embedded in a common superordinate group identity (Gaertner et al., 1999; Rust, 1996). Thus, crossed categorization and recategorization may work together to produce enhanced inclusiveness and reduced intergroup discrimination.

Cross-cutting category memberships, however, do not always result in reduced category salience and greater intergroup acceptance (Vanbeselaere, 1991). If one category distinction is more socially meaningful or functionally important than others, intergroup discrimination based on that categorization may be unaffected by the existence of cross-cutting memberships in other, less important groups. More important, multiple group identities may be combined into a single ingroup (e.g., categorization based on shared ethnicity *and* gender), which is more exclusive than either category membership considered separately. This possibility shifts the focus to the *phenomological* rather than the objective categorization structure. To understand when cross-cutting categories will be effective in reducing prejudice, we need to know how the individual subjectively represents his or her ingroups and the interrelationships among them.

As one step toward understanding the subjective structure of multiple social identities, Roccas and Brewer (2002) introduced the concept of *social identity complexity*. The idea behind the complexity construct is that it is not only how many social groups an individual identifies with that matters but, more important, how those different identities are subjectively combined to determine the overall inclusiveness of the individual's ingroup memberships. Specifically, Roccas and Brewer proposed that multiple social identities can be represented along a continuum of complexity and inclusiveness, reflecting the degree to which different identities are both differenti-

ated and integrated in the individual's cognitive representation of his or her group memberships. At the low end of the complexity dimension, the individual defines the ingroup as the *intersection* of all of his or her group identities, creating a single, highly exclusive identity category (e.g., female Republican college professors) whereby others who do not share all the same memberships are effectively outgroup members. At the high end of the social identity complexity dimension, the individual recognizes that each of his or her group memberships incorporates a different set of people as ingroup members and the combined representation is the *sum* of all of these group identities—more inclusive than any one ingroup identity considered alone.

Roccas and Brewer (2002) also speculated that social identity complexity (as represented by perceived overlap among ingroup memberships) would be associated with tolerance for outgroups in general. Social identity complexity is based on chronic awareness of cross-categorization in one's own social group memberships and those of others. A simple social identity is likely to be accompanied by the perception that any individual who is an outgroup member on one dimension is also an outgroup member on all others. In contrast, if an individual is aware that one of his ingroups only partly overlaps with any other of his ingroups, then we assume that he is also aware that some of his ingroup members have crossed group memberships: They are ingroup members on one dimension but are simultaneously outgroup members on others. Making salient that an outgroup member on one category dimension is an ingroup member on another decreases bias by comparison with instances in which the latter information is not available (Gaertner et al., 1993).

There are a number of theoretical reasons why a complex representation of ingroup categorization should influence intergroup attitudes and behavior in ways that reduce bias and discrimination. First, cross-cutting distinctions make social categorization more complex and reduce the magnitude of ingroup–outgroup distinctions. According to social categorization theory (Deschamps & Doise, 1978; Doise, 1978; Vanbeselaere, 1991), processes of intracategory assimilation and intercategory contrast counteract each other when categories are cross-cutting. Thus, the effects of intercategory accentuation are reduced or eliminated, and differences between groups are minimized (or no greater than perceived differences within groups). This undermines the cognitive basis of ingroup bias. Second, partially overlapping group memberships reduce the evaluative significance for the self of intergroup comparisons, thereby undermining the motivational base for intergroup discrimination (Vanbeselaere, 1991). Third, multiple group memberships reduce the importance of any one social identity for satisfying an individual's need for belonging and self-definition (Brewer, 1991), again reducing the motivational base for ingroup bias.

Finally, principles of cognitive balance (Heider, 1958; Newcomb, 1963) are also brought into play when ingroups and outgroups have overlapping membership. When another person is an ingroup member on one cat-

egory dimension but belongs to an outgroup in another categorization, cognitive inconsistency is introduced if that individual is evaluated positively as an ingroup member but is also associated with others who are evaluated negatively as outgroup members. In an effort to resolve such inconsistencies, interpersonal balance processes should lead to greater positivity toward the outgroup based on overlapping memberships. The processes here are similar to those underlying the "extended contact effect" (Wright et al., 1997). Having someone connected to the self (i.e., a friend or an ingroup member) associated with an outgroup should reduce negativity toward that outgroup.

In sum, both cognitive and motivational factors lead us to predict that complex social identities will be associated with reduced ingroup favoritism and increased tolerance and positivity toward outgroups in general. Results from a telephone interview survey of adult residents of the state of Ohio (Brewer & Pierce, 2005) supported this hypothesis. Individual differences in complexity of perception of their national, religious, occupational, political, and recreational social identities were systematically related to their attitudes toward ethnic outgroups and diversity. Individuals with a more complex representation of their own multiple ingroups were more tolerant of diversity and more accepting of outgroups in general than were individuals with less social identity complexity. Thus, both the actual complexity of the structure of the society and the individual's subjective representation of his or her networks of social group memberships contribute to the reduction of negative ingroup–outgroup relations.

CONCLUSIONS: IMPLICATIONS FOR MULTICULTURAL SOCIETIES

The principles of social categorization, ingroup favoritism, and outgroup prejudice discussed in this chapter have important implications for promoting positive intergroup relations within a context in which groups must live together interdependently. The same basic principles apply whether we are considering two nuclear families joining into a common household, departments or companies combined within an organization, diverse ethnic or religious groups within a nation, or nation-states within an international community. In any of these contexts, the goals of contact and cooperation compete with natural tendencies toward ingroup–outgroup differentiation, separation, and exclusion. Personalization across category boundaries and formation of common superordinate identities—processes that reduce the social meaning of category boundaries—are in tension with pluralistic values that seek to maintain cultural variation and distinct social identities.

The tension between differentiation and integration must be recognized and acknowledged in any complex social system. Exclusive focus on either assimilation or separation as the solution to intergroup discrimination and conflict is neither desirable nor realistic. Proponents of multiculturalism assert that alternatives to these extremes are possible, that groups can maintain distinct identities at the same time that their members participate in a shared, superordinate group structure. Berry (1984), for instance, has argued that there are four different forms of interethnic relations possible in a pluralistic society, depending on how members of the diverse ethnic groups relate to their own ethnic identity and to their role in the society at large. In Berry's classification system, integration is the form of intercultural relations in which identification with ethnic subgroups and identification with the larger society are both engaged. The reciprocal relations among processes of personalization, recategorization, and mutual differentiation discussed earlier are compatible with this view of social integration and constitute the necessary underpinnings for an equalitarian multicultural society.

REFERENCES

Alexander, M. G., Brewer, M. B., & Herrmann, R. K. (1999). Images and affect: A functional analysis of out-group stereotypes. *Journal of Personality and Social Psychology, 77,* 78–93.

Alexander, M. G., Brewer, M. B., & Livingston, R. W. (2005). Putting stereotype content in context: Image theory and interethnic stereotypes. *Personality and Social Psychology Bulletin, 31,* 781–794.

Alexander, R. D. (1989). Evolution of the human psyche. In P. Mellars & C. Stringer (Eds.), *The human revolution* (pp. 455–513). Princeton, NJ: Princeton University Press.

Allport, G. (1954). *The nature of prejudice.* Cambridge, MA: Addison-Wesley.

Bachman, B. A. (1993). *An intergroup model of organizational mergers.* Unpublished doctoral dissertation, University of Delaware.

Banker, B. S., & Gaertner, S. L. (1998). Achieving stepfamily harmony: An intergroup relations approach. *Journal of Family Psychology, 12,* 310–325.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71,* 230–244.

Berry, J. W. (1984). Cultural relations in plural societies: Alternatives to segregation and their sociopsychological implications. In N. Miller & M. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 11–27). New York: Academic Press.

Bettencourt, B. A., Brewer, M. B., Croak, M. R., & Miller, N. (1992). Cooperation and the reduction of intergroup bias: The role of reward structure and social orientation. *Journal of Experimental Social Psychology, 28,* 301–319.

Bettencourt, B. A., & Dorr, N. (1998). Cooperative interaction and intergroup bias: Effects of numerical representation and crosscut role assignment. *Personality and Social Psychology, 24,* 1276–1293.

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 35-58). Oxford, UK: Blackwell.

Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307–324.

Brewer, M. B. (1981). Ethnocentrism and its role in interpersonal trust. In M. Brewer & B. Collins (Eds.), *Scientific inquiry and the social sciences* (pp. 345–360). San Francisco: Jossey-Bass.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17,* 475–482.

Brewer, M. B. (1996). In-group favoritism: The subtle side of intergroup discrimination. In D. Messick & A. Tenbrunsel (Eds.),

Codes of conduct: Behavioral research into business ethics (pp. 160–170). New York: Russell Sage.

Brewer, M. B. (1997). On the social origins of human nature. In G. McGarty & S. A. Haslam (Eds.), *The message of social psychology: Perspectives on mind and society* (pp. 54–62). Oxford, UK: Blackwell.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, *55*(3), 429–444.

Brewer, M. B. (2000). Superordinate goals versus superordinate identity as bases for intergroup cooperation. In D. Capozza & R. Brown (Eds.), *Social identity processes* (pp. 117–132). London: Sage.

Brewer, M. B. (2001). Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate? In R. Ashmore, L. Jussim, & D. Wilder (Eds.), *Social identity, intergroup conflict, and conflict reduction* (pp. 17–41). New York: Oxford University Press.

Brewer, M. B., & Alexander, M. G. (2002). Intergroup emotions and images. In D. Mackie & E. Smith (Eds.), *From prejudice to intergroup emotions* (pp. 209–225). New York: Psychology Press.

Brewer, M. B., & Campbell, D. T. (1976). *Ethnocentrism and intergroup attitudes: East African evidence.* New York: Halsted-Press (Sage).

Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology*, *50*, 543–549.

Brewer, M., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281–302). New York: Academic Press.

Brewer, M. B., & Miller, N. (1988). Contact and cooperation: When do they work? In P. Katz & D. Taylor (Eds.), *Eliminating racism: Means and controversies* (pp. 315–326). New York: Plenum Press.

Brewer, M. B., & Pierce, K. P. (2005). Social identity complexity and outgroup tolerance. *Personality and Social Psychology Bulletin*, *31*, 428–437.

Brewer, M. B., & Schneider, S. (1990). Social identity and social dilemmas: A double-edged sword. In D. Abrams & M. Hoggs (Eds.), *Social identity theory: Constructive and critical advances* (pp. 169–184). London: Harvester-Wheatsheaf.

Brown, R. J., & Hewstone, M. (2005). An integrative theory of intergroup contact. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 256–343). New York: Elsevier.

Brown, R. J., & Turner, J. C. (1979). The criss-cross categorization effect in intergroup discrimination. *British Journal of Social and Clinical Psychology*, *18*, 371–383.

Brown, R. J., Vivian, J., & Hewstone, M. (1999). Changing attitudes through intergroup contact: The effects of group membership salience. *European Journal of Social Psychology*, *29*, 741–764.

Brown, R. J., & Wade, G. (1987). Superordinate goals and intergroup behaviour: The effect of role ambiguity and status on intergroup attitudes and task performance. *European Journal of Social Psychology*, *17*, 131–142.

Bruins, J., Platow, M., & Ng, S. H. (1995). Distributive and procedureal justice in interpersonal and intergroup situations. *Social Justice Research*, *8*, 103–121.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*, 401–423.

Cameron, J. A., Alvarez, J. M., Ruble, D. N., & Fuligni, A. J. (2001). Children's lay theories about ingroups and outgroups: Reconceptualizing research on prejudice. *Personality and Social Psychology Review*, *5*(2), 118–128.

Campbell, D. T. (1956). Enhancement of contrast as a composite habit. *Journal of Abnormal and Social Psychology*, *53*, 350–355.

Campbell, D. T. (1967). Stereotypes and the perception of outgroup differences. *American Psychologist*, *22*, 812–829.

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*, 276–298.

Caporael, L. R., Dawes, R. M., Orbell, J. M., & van de Kragt, A. J. C. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, *12*, 683–739.

Coats, S., Smith, E. R., Claypool, H. M., & Banner, M. J. (2000). Overlapping mental representations of self and ingroup: Reaction time evidence and its relationship with explicit measures of group identification. *Journal of Experimental Social Psychology*, *36*, 304–315.

Coser, L. A. (1956). *The functions of social conflict.* New York: Free Press.

Cottam, R. (1977). *Foreign policy motivation: A general theory and a case study.* Pittsburgh, PA: University of Pittsburgh Press.

DeCremer, D., & van Dijk, E. (2002). Reactions to group success and failure as a function of identification level: A test of the goal-transformation hypothesis in social dilemmas. *Journal of Experimental Social Psychology*, *38*, 435–442.

DeCremer, D., & van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology*, *29*, 871–893.

Deschamps, J.-C., & Brown, R. J. (1983). Superordinate goals and intergoup conflict. *British Journal of Social Psychology*, *22*, 189–195.

Deschamps, J-C., & Doise, W. (1978). Crossed category memberships in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 141–158). Cambridge, UK: Cambridge University Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Dijker, A. J. M. (1987). Emotional reaction to ethnic minorities. *European Journal of Social Psychology*, *17*, 305–325.

Dijksterhuis, A., Bargh, J. A., & Miedema, J. (2000). Of men and mackerels: Attention and automatic social behavior. In H. Bless & J. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 37–51). Philadelphia: Psychology Press/Taylor & Francis.

Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, *74*, 865–877.

Doise, W. (1978). *Groups and individuals: Explanations in social psychology.* Cambridge, UK: Cambridge University Press.

Doise, W., Deschamps, J-C., & Meyer, G. (1978). The accentuation of intracategory similarities. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 159–170). London: Academic Press.

Dovidio, J. F., & Gaertner, S. L. (2004). Aversive racism. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 36, pp. 1–52). San Diego, CA: Elsevier.

Dovidio, J. F., Gaertner, S. L., & Validzic, A. (1998). Intergroup bias: Status differentiation and a common ingroup identity. *Journal of Personality and Social Psychology*, *75*, 109–120.

Duckitt, J., Callaghan, J., & Wagner, C. (2005). Group identification and outgroup attitudes in four South African ethnic groups: A multidimensional approach. *Personality and Social Psychology Bulletin*, *31*, 633–646.

Duckitt, J., & Parra, C. (2004). Dimensions of group identification and outgroup attitudes in four ethnic groups in New Zealand. *Basic and Applied Social Psychology*, *26*, 237–247.

Dustin, D. A., & Davis, H. P. (1970). Evaluative bias in group and individual competition. *Journal of Social Psychology*, *80*, 103–108.

Esses, V. M., Haddock, G., & Zanna, M. P. (1993). Values, stereotypes and emotions as determinants of intergroup attitudes. In D. Mackie & D. Hamilton (Eds.), *Affect, cognition and stereotyping* (pp. 137–166). San Diego, CA: Academic Press.

Farnham, S. D., Greenwald, A. G., & Banaji, M. R. (1999). Implicit self-esteem. In D. Abrams & M. Hogg (Eds.), *Social identity and social cognition* (pp. 230–248). Oxford, UK: Blackwell.

Ferguson, C. K., & Kelley, H. H. (1964). Significant factors in overevaluation of own groups' products. *Journal of Abnormal and Social Psychology*, *69*, 223–228.

Foley, R. (1996). The adaptive legacy of human evolution: A search

for the environment of evolutionary adaptedness. *Evolutionary Anthropology*, *4*, 194–203.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Philadelphia: Psychology Press.

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, *4*, 1–26.

Gaertner, S. L., Dovidio, J. F., & Bachman, B. A. (1996). Revisiting the contact hypothesis: The induction of a common ingroup identity. *International Journal of Intercultural Relations*, *20*, 271–290.

Gaertner, S. L., Dovidio, J. F., Banker, B., Houlette, M., Johnson, K., & McGlynn, E. (2000). Reducing intergroup conflict: From superordinate goals to decategorization, recategorization, and mutual differentiation. *Group Dynamics*, *4*, 98–114.

Gaertner, S. L., Dovidio, J. F., Nier, J. A., Ward, C. M., & Banker, B. S. (1999). Across cultural divides: The value of a superordinate identity. In D. Prentice & D. Miller (Eds.), *Cultural civides: Understanding and overcoming group conflict* (pp. 173–212). New York: Russell Sage.

Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology*, *59*, 692–704.

Gaertner, S. L., Mann, J. A., Murrell, A. J., & Dovidio, J. F. (1989). Reduction of intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology*, *57*, 239–249.

Gaertner, S. L., Rust, M. C., Dovidio, J. F., Bachman, B. A., & Anastasio, A. (1994). The contact hypothesis: The role of a common ingroup identity on reducing intergroup bias. *Small Groups Research*, *25*, 224–290.

Greenland, K., & Brown, R. J. (1999). Categorization and intergroup anxiety in contact between British and Japanese nationals. *European Journal of Social Psychology*, *29*, 503–521.

Gluckman, M. (1955). *Customs and conflict in Africa*. London: Blackwell.

Hamberger, J., & Hewstone, M. (1997). Inter-ethnic contact as a predictor of prejudice: Tests of a model in four West European nations. *British Journal of Social Psychology*, *36*, 173–190.

Hass, R., Katz, I., Rizzo, N., Bailey, J., & Eisenstadt, D. (1991). Cross-racial appraisal as related to attitude ambivalence and cognitive complexity. *Personality and Social Psychology Bulletin*, *17*, 83–92.

Hassan, F. A. (1981). *Demographic archaeology*. New York: Academic Press.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Herrmann, R. K. (1985). *Perceptions and behavior in Soviet foreign policy*. Pittsburgh, PA: University of Pittsburgh Press.

Herrmann, R. K., & Fischerkeller, M. (1995). Beyond the enemy image and spiral model: Cognitive-strategic research after the Cold War. *International Organization*, *49*, 415–450.

Hewstone, M. (1996). Contact and categorization: Social psychology interventions to change intergroup relations. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 323–368). New York: Guilford Press.

Hewstone, M., & Brown, R. J. (1986). Contact is not enough: An intergroup perspective on the "contact hypothesis." In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1–44). Oxford, UK: Basil Blackwell.

Hewstone, M., Cairns, E., Voci, A., Paolini, S., McLernon, F., Crisp, R. J., et al. (2005). Intergroup contact in a divided society: Challenging segregation in Northern Ireland. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 265–292). Philadelphia: Psychology Press.

Hinkle, S., & Brown, R. (1990). Intergroup comparisons and social identity: Some links and lacunae. In D. Abrams & M. Hogg (Eds.), *Social identity theory: Construction and critical advances* (pp. 48–70). London: Harvester Wheatsheaf.

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping, and the salience of social categories. *British Journal of Social Psychology*, *26*, 325–340.

Hornsey, M. J., & Hogg, M. A. (2000). Subgroup relations: A comparison of mutual intergroup differentiation and common ingroup identity models of prejudice reduction. *Personality and Social Psychology Bulletin*, *26*, 242–256.

Hornstein, H. A. (1976). *Cruelty and kindness: A new look at aggression and altruism.* Englewood Cliffs, NJ: Prentice Hall.

Huo, Y., Smith, H., Tyler, T. R., & Lind, E. A. (1996). Superordinate identification, subgroup identification, and justice concerns: Is separatism the problem; is assimilation the answer? *Psychological Science*, *7*, 40–45.

Insko, C. A., & Schopler, J. (1987). Categorization, competition, and collectivity. In C. Hendrick (Ed.), *Group processes. Review of personality and social psychology* (Vol. 8, pp. 213–251). Beverly Hills, CA: Sage.

Islam, M. R., & Hewstone, M. (1993). Dimensions of contact as predictors of intergroup anxiety, perceived outgroup variability and outgroup attitude: An integrative account. *Personality and Social Psychology Bulletin*, *19*, 700–710.

Jetten, J., Spears, R., & Postmes, T. (2004). Intergroup distinctiveness and differentiation: A meta-analytic integration. *Journal of Personality and Social Psychology*, *86*, 862–879.

Jost, J. T. (2001). Outgroup favoritism and the theory of system justification: An experimental paradigm for investigating the effects of socio-economic success on stereotype content. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the legacy and future of social cognition* (pp. 89–102). Hillsdale, NJ: Erlbaum.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893–905.

Kosterman, R., & Feshbach, S. (1989). Toward a measure of patriotic and nationalistic attitudes. *Political Psychology*, *10*, 257–274.

Kramer, R. M., Brewer, M. B., & Hanna, B. A. (1996). Collective trust and collective action: The decision to trust as a social dilemma. In R. Kramer & T. Tyler (Eds.), *Trust in organizations* (pp. 357–389). Thousand Oaks, CA: Sage.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. *Review of Personality and Social Psychology*, *13*, 25–59.

Levin, S., & Sidanius, J. (1999). Social dominance and social identity in the United States and Israel: Ingroup favoritism or outgroup derogation? *Political Psychology*, *20*, 99–126.

LeVine, R. A., & Campbell, D. T. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes and group behavior*. New York: Wiley.

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, *79*, 602–616.

Mackie, D. M., & Smith, E. R. (1998). Intergroup relations: Insights from a theoretically integrative approach. *Psychological Review*, *105*, 499–529.

Marcus-Newhall, A., Miller, N., Holtz, R., & Brewer, M. B. (1993). Crosscutting category membership with role assignment: A means of reducing intergroup bias. *British Journal of Social Psychology*, *32*, 125–146.

McClintock, C. (1972). Social motivation: A set of propositions. *Behavioral Science*, *17*, 438–454.

McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. Dovidio & S. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Orlando, FL: Academic Press.

McConahay, J. B., & Hough, J. C. (1976). Symbolic racism. *Journal of Social Issues*, *32*, 23–45.

Miller, N., Brewer, M. B., & Edwards, K. (1985). Cooperative interaction in desegregated settings: A laboratory analogue. *Journal of Social Issues*, *41*(3), 63–79.

Mottola, G. (1996). *The effects of relative group status on expectations of merger success*. Unpublished doctoral dissertation, University of Delaware, Newark.

Mummendey, A., Simon, B., Dietze, C., Grunert, M., Haeger, G., Kessler, S., et al. (1992). Categorization is not enough: Inter-

group discrimination in negative outcome allocations. *Journal of Experimental Social Psychology*, *28*, 125–144.

Mummendey, A., & Waldzus, S. (2004). National differences and European plurality: Discrimination or tolerance between European countries. In R. Herrmann, T. Risse, & M. Brewer (Eds.), *Transnational identities: Becoming European in the EU* (pp. 59–72). Lanham, MD: Rowman & Littlefield.

Mummendey, A., & Wenzel, M. (1999). Social discrimination and tolerance in intergroup relations. *Personality and Social Psychology Review*, *3*, 158–174.

Newcomb, T. M. (1963). Stabilities underlying changes in interpersonal attraction. *Journal of Abnormal and Social Psychology*, *66*, 376–386.

Otten, S. (2002). I am positive and so are we: The self as determinant of favoritism toward novel ingroups. In J. Forgas & K. Williams (Eds.), *The social self: Cognitive, interpersonal, and intergroup processes* (pp. 273–291). New York: Psychology Press.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology*, *29*, 1049–1071.

Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin*, *23*, 173–185.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*, 65–85.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, *25*, 57–75.

Piliavin, J. A., Dovidio, J. F., Gaertner, S. L., & Clark, R. D., III. (1981). *Emergency intervention*. New York: Academic Press.

Platow, M. J., Harley, K., Hunter, J. A., Hanning, P., Shave, R., & O'Connell, A. (1997). Interpreting in-group favoring allocations in the minimal group paradigm. *British Journal of Social Psychology*, *36*, 107–117.

Purdue, C., Dovidio, J., Gurtman, M., & Tyler, R. (1990). Us and them: Social categorization and the process of ingroup bias. *Journal of Personality and Social Psychology*, *59*, 475–486.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287–290.

Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review*, *6*, 88–106.

Rogers, R. W., & Prentice-Dunn, S. (1981). Deindividuation and anger-mediated interracial aggression: Unmasking regressive racism. *Journal of Personality and Social Psychology*, *41*, 63–73.

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin*, *27*, 1164–1178.

Rust, M. C. (1996). *Social identity and social categorization*. Unpublished doctoral dissertation, University of Delaware, Newark.

Sande, G., Goethals, G., Ferrari, L., & Worth, L. (1989). Value-guided attributions: Maintaining the moral self-image and the diabolical enemy-image. *Journal of Social Issues*, *45*, 91–118.

Schubert, T. W., & Hafner, M. (2003). Contrast from social stereotypes in automatic behavior. *Journal of Experimental Social Psychology*, *39*, 577–584.

Sears, D. O. (1988). Symbolic racism. In P. Katz & D. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53–84). New York: Plenum Press.

Sears, D. O., & Funk, C. L. (1991). The role of self-interest in social and political attitudes. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 1–91). New York: Academic Press.

Sherif, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation*. New York: Houghton Mifflin.

Simon, B., & Hamilton, D. H. (1994). Social identity and self-stereotyping: The effects of relative group size and group status. *Journal of Personality and Social Psychology*, *66*, 699–711.

Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. In D. Mackie & D. Hamil-ton (Eds.), *Affect, cognition, and stereotyping* (pp. 297–315). San Diego, CA: Academic Press.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin*, *22*, 635–642.

Smith, H. J., & Tyler, T. R. (1996). Justice and power: When will justice concerns encourage the advantaged to support policies which redistribute economic resources and the disadvantaged to willingly obey the law? *European Journal of Social Psychology*, *26*, 171–200.

Spears, R., Gordijn, E., Dijkersterhuis, A., & Stapel, D. (2004). Reaction in action: Intergroup contrast in automatic behavior. *Personality and Social Psychology Bulletin*, *30*, 605–616.

Stangor, C., Sullivan, L., & Ford, T. (1991). Affective and cognitive determinants of prejudice. *Social Cognition*, *9*, 359–380.

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., et al. (2002). The role of threats in the racial attitudes of blacks and whites. *Personality and Social Psychology Bulletin*, *28*, 1242–1254.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Mahwah, NJ: Erlbaum.

Stiner, M. C. (2002). Carnivory, coevolution, and the geographic spread of the genus *Homo*. *Journal of Archeological Research*, *10*, 1–63.

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*, 364–373.

Sumner, W. G. (1906). *Folkways*. New York: Ginn.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, *25*, 79–97.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(2), 96–102.

Tajfel, H. (1981). *Human groups and social categories*. Cambridge, UK: Cambridge University Press.

Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*, 149–178.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, *54*, 101–114.

Takagi, E. (1996). The generalized exchange perspective on the evolution of altruism. In W. Liebrand & D. Messick (Eds.), *Frontiers in social dilemma research* (pp. 311–336). Berlin: Springer-Verlag.

Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology*, *5*, 5–34.

Turner, J. C., Hogg, M., Oakes, P., Reicher, S., & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford, UK: Blackwell.

Turner, J. C., & Reynolds, K. (2001). The social identity perspective in intergroup relations: Theories, themes, and controversies. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology. Intergroup processes* (pp. 133–152). Oxford, UK: Blackwell.

Tyler, T. R., Degoey, P., & Smith, H. (1996). Understanding why the justice of group procedures matters: A test of the psychological dynamics of the group-value model. *Journal of Personality and Social Psychology*, *69*, 482–497.

Vanbeselaere, N. (1991). The different effects of simple and crossed categorizations: A result of the category differentiation process or of differential category salience? *European Review of Social Psychology*, *2*, 247–278.

Van Oudenhoven, J. P., Groenewoud, J. T., & Hewstone, M. (1996). Cooperation, ethnic salience and generalisation of interethnic attitudes. *European Journal of Social Psychology*, *26*, 649–661.

Von Hippel, W., Brewer, M. B., & Polifroni, M. (2001, May). *Nazi's on the Web: Support for image theory in the Internet communications of White Supremacists*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Watson, G. (1947). *Action for unity*. New York: Harper.

Wenzel, M. (2000). Justice and identity: The significance of inclu-

sion for perceptions of entitlement and the justice motive. *Personality and Social Psychology Bulletin*, *26*, 157–176.

Wilder, D. A. (1978). Reduction of intergroup discrimination through individuation of the outgroup. *Journal of Personality and Social Psychology*, *36*, 1361–1374.

Wilder, D. A. (1986). Social categorization: Implications for creation and reduction of intergoup bias. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 291–355). New York: Academic Press.

Williams, R. M., Jr. (1947). *The reduction of intergroup tensions*. New York: Social Science Research Council.

Worchel, S., Rothgerber, H., Day, E. A., Hart, D., & Butemeyer, J. (1998). Social identity and individual productivity with groups. *British Journal of Social Psychology*, *37*, 389–413.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*, 73–90.

Zanna, M., & Rempel, J. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). New York: Cambridge University Press.

CHAPTER 31

Social Psychology of Leadership

MICHAEL A. HOGG

Leadership is a huge, rambling behemoth of a topic. Because it so profoundly affects all of our lives, leadership is the focus not only of pure and applied academic disciplines such as philosophy, history, management and psychology but also of current affairs, literature and the arts, and of course the popular self-help and personal growth industry. Biography is frequently about leadership, and most classic accounts of history are mainly accounts of the actions of leaders. Our day-to-day life is pervaded by the impact of leadership—for example, leadership in the political, governmental, corporate, work, educational, and artistic spheres—and we all, to varying degrees, occupy leadership roles ourselves. Not surprisingly, people take a keen interest in leadership and we all have our own views on leaders and leadership. Incompetent leadership and leadership in the service of evil, in particular, are of great concern to us all (e.g., Kellerman, 2004). An excellent illustration of the scope of the topic of leadership is Goethals, Sorenson, and Burns's (2004) recent four-volume *Encyclopedia of Leadership*—1,927 pages, 1.2 million words, and 373 substantive (1,000–6,000 word) entries written by 311 scholars.

Because leadership has many facets and dimensions and occurs in many different contexts, it is a legitimate focus for many disciplines. For example, political science focuses on political and governance contexts (e.g., Gergen, 2000), philosophy focuses on the general principles of good leadership (e.g., Cawthon & Clark, 2002), history has often focused on the actions of "great people" (e.g., Simonton, 1994), and management science and or-

ganizational psychology, which are the main domains of leadership research, focus primarily on business leadership in corporate and organizational contexts (e.g., Avolio, 1999; Bass & Avolio, 1994; Bennis, 1989; Burns, 2003; Kellerman, 1999; Northhouse, 2004; Yukl, 2002).

Perhaps partly because of the dominance of research on business leadership, the scientific study of public leadership is somewhat less prolific—and yet public leadership has a huge impact on our lives. The Center for Public Leadership at Harvard University recently convened a 35-member committee of experts to name America's best currently living leaders (presidents and ex-presidents were excluded from the list). Published in *U.S. News & World Report* (October 31, 2005), the final list of 25 leaders included political leaders such as Colin Powell and Condolezza Rice; industry innovators such as Bill Gates, Steve Jobs, and Howard Schutz; media personalities such as Oprah Winfrey; and scientific leaders such as Francis Collins and Craig Venter, who mapped the human genome.

SOCIAL PSYCHOLOGY AND LEADERSHIP

Where does social psychology fit into all this? Historically, leadership was a significant and major focus of social psychology. This was the case through the glory days of research in the small-group dynamics tradition (e.g., Cartwright & Zander, 1968; Shaw, 1981). Indeed, leadership research has been a focus of some of social

psychology's classic research programs (e.g., Bales, 1950; Fiedler, 1964; Hollander, 1958; Lippitt & White, 1943; Sherif, 1966; Stogdill, 1974).

However, during the 1970s social psychology became preoccupied by an interest in social perception, social inference, and social cognitive processes (e.g., Devine, Hamilton, & Ostrom, 1994; Fiske & Taylor, 1991), and, in Europe, with large-scale intergroup relations and processes of social change (e.g., Moscovici, 1976; Tajfel, 1984). Research on small interactive groups lost its shine (e.g., Steiner, 1974) and there was an exodus of small-group research and small-group researchers to neighboring disciplines, notably organizational psychology and management science (Levine & Moreland, 1995; McGrath, 1997; Sanna & Parks, 1997; Tindale & Anderson, 1998).

Leadership research suffered the same fate—the third edition of *The Handbook of Social Psychology* featured a chapter on leadership, by Hollander (1985), but the most recent, fourth edition did not (Gilbert, Fiske, & Lindzey, 1998). For a period of almost a quarter of a century, since the mid-1970s most significant social psychological research on leadership has been conducted outside the social psychology mainstream—mainly in organizational and management disciplines where it found a natural and receptive home. This scientific focus on business leadership may have come at a cost—a somewhat restricted leadership research agenda, directing attention away from public leadership, nonprofit leadership, social movement leadership, celebrity leadership, scientific leadership, philanthropic leadership, and other manifestations of leadership as a general phenomenon in which individuals or groups are able to provide leadership to others.

Over the past 5–10 years things have changed. There has been a growing revival of interest within the mainstream of social psychology in leadership. This is attested to by many new empirical articles, and by handbook chapters (Chemers, 2001; Lord, Brown, & Harvey, 2001), an integrative position paper (van Knippenberg, van Knippenberg, De Cremer, & Hogg, 2004), a number of books (e.g., Chemers, 1997; Lord & Brown, 2004; Messick & Kramer, 2005; van Knippenberg & Hogg, 2003a), and of course this handbook chapter. There has also been a significant focus on gender and leadership (e.g., Eagly, Karau, & Makhijani, 1995; Eagly, Makhijani, & Klonsky, 1992).

One reason for this revival is that social psychology has once again taken an interest in studying groups—but this is a more wide-ranging and more cognitive interest in groups that incorporates both small-scale face-to-face groups and large-scale social categories and focuses on both what happens within and between groups (see Abrams & Hogg, 1998; Moreland, Hogg, & Hains, 1994).

DEFINING LEADERSHIP

The main reason, however, for this revival in interest in leadership is that leadership is a quintessentially social psychological phenomenon. Take Gordon Allport's (1968) classic definition of social psychology as "the scientific investigation of how the thoughts, feelings and behaviors of individuals are influenced by the actual, imagined or implied presence of others" (p. 3)—"influence" is the key term here, and what leaders do is influence other people. Chemers (2001) defines leadership as "a process of social influence through which an individual enlists and mobilizes the aid of others in the attainment of a collective goal" (p. 376). Leadership requires there to be an individual, or clique, which influences the behavior of another individual or group of individuals—where there are leaders there must be followers.

Not surprisingly, it is difficult to find a single simple definition of leadership that embraces all the research done on leadership—there are numerous definitions that reflect and are closely tied to the variety of different theories of, approaches to, and perspectives on leadership. From a social psychological point of view, however, leadership rests on the proximal or distal interaction between leader and follower that allows the former to influence the latter. The interaction is largely symbolic and is structured by an interaction of leader attributes and cognitions, follower attributes and cognitions, and properties of the immediate interactive context and wider sociostructural context.

This is a broad and complex characterization of the field that invites the question of what is *not* leadership. If I asked you to spend the weekend helping me clean my apartment and you agreed, either because you liked me or you were afraid of me, it would be influence but not leadership—this is a classic case of compliance (e.g., Cialdini & Trost, 1998). Related to this, the exercise of power is generally not considered to be leadership (e.g., Chemers, 2001; Lord, Brown, & Harvey, 2001; Moscovici, 1976; Raven, 1993). If you agreed because you knew that there was a neighborhood norm to clean at the weekend, that would be conformity to a norm (e.g., Turner, 1991)—not an example of leadership. If on the other hand I had first convinced you that we should develop a neighborhood cleaning norm, and you subsequently adhered to that norm, then that most definitely would be leadership. Leaders play a critical role in defining collective goals, and in this respect leadership is more typically a group processes than an interpersonal process—it is an influence process that plays out more noticeably in group than interpersonal contexts.

Another key question is, How do we evaluate leaders and leadership (e.g., Kellerman, 2004)? In answering this question it is useful to distinguish between the two evaluative dimensions of effective/ineffective leaders and good/bad leaders. An effective leader is someone who is successful in setting new goals and influencing others to achieve them. Here, the evaluation of leadership is largely an objective matter of fact—how much influence did the leader have in setting new goals and were the goals achieved? In contrast, evaluating whether the leader is good or bad is largely a subjective judgment based on one's preferences and perspective, and one's goals and group memberships. We evaluate leaders in terms of their character (e.g., nice, nasty, or charismatic), the ethics and morality of the means they use to influ-

ence others and achieve goals (e.g., persuasion, coercion, oppression, or democratic decision making), and the nature of the goals that they lead their followers toward (e.g., saving the environment, reducing starvation and disease, producing a commodity, combating oppression, waging war, and engaging in genocide). Here good leaders are those who have attributes we applaud, use means we approve of, and set and achieve goals we value. Thus, secular Westerners and supporters of al-Qaeda would disagree on whether Osama bin Laden is a good leader (they disagree on the value of his goals and the morality of his means) but may agree that he has been a relatively effective leader (agreeing that he has mobilized fundamentalist Muslims around his cause).

CHAPTER AIMS AND OUTLINE

Because leadership is such a huge subject it would be impossible to provide, in this single chapter, a comprehensive and detailed coverage of the entire field. I have had to be selective. In doing this I have tried to make sure that the main themes and ideas are covered, albeit sometimes briefly, but have focused more on social psychology than on management and organizational science, and I have tried to spend more time on some of the more recent contributions from and within social psychology.

I start by discussing the extent to which some people are more effective leaders because they have specific personalities that predispose them to leadership. This focus on personality and individual differences surfaces in a slightly different guise later when we discuss transformational and charismatic leadership. To counterbalance the personality perspective is a situational perspective that essential argues that we can all lead if the circumstances are right.

Generally, an interactionist perspective is most useful—particular kinds of leadership behavior are more effective in some situations, and other behaviors are more effective in other situations. This view spawned substantial research on the behavior of leaders, in an attempt to distinguish between different categories of leader behavior—one important distinction that surfaces in different guises is between behaviors that focus primarily on getting the task done and behaviors that focus primarily on how group members feel. The interactionist perspective finds its most elaborated form in contingency theories of leadership—the leadership effectiveness of certain types of behavior is contingent upon situational factors. I discuss Fiedler's contingency theory, normative decision theory, and path-goal theory.

Leadership involves interactions between leaders and followers. One perspective on these interactions is that they involve transactions in which both leaders and followers provide resources for each other—if the transaction is mutually beneficial then leadership is effective. I discuss Hollander's notion of idiosyncrasy credit, the vertical dyad linkage (VDL) model, and leader–member exchange (LMX) theory. Path-goal theory can also be classified as a transactional theory.

Undoubtably the most popular approach to leadership in recent years, particularly in management and organizational science, is the transformational model. Effective leaders are those who can instill in followers their vision for a new and better future for the group—effective leaders need to be able to transform groups. Charisma and charismatic leadership play a central role in transformational leadership.

Another, more social psychological, perspective on leadership is fueled by social cognition and what we know about the role of schemas in social perception. According to leader categorization theory, we all have schemas of leadership, and effective leaders are those whom we can categorize as a good fit to our schema of an effective leader in a particular situation. Schematic expectations also play a central role in expectations state theory, status characteristics theory, and role congruity theory.

I have dedicated a section to the social identity theory of leadership. Although only relatively recently formulated, this approach has attracted substantial attention (it crops up throughout the chapter), and because it links social cognition, self-conception, and group behavior, it represents a direct treatment of leadership as a group process. The basic idea is that where people identify with a group, leadership benefits if the leader is seen to embody the group's attributes—this builds popularity and trust and allows the leader latitude to be innovative.

The body of the chapter finishes with a discussion of trust and leadership, and gender and leadership. Trust is very basic to effective leadership. I discuss the role of justice perceptions and trust in leadership, and the role of leadership in building trust that may resolve social dilemmas. The study of gender differences in perceived leadership effectiveness has gathered momentum over the past 10 years—mainly through the work of Eagly and her associates. One of the main factors influencing gender and leadership is role incongruity—where gender stereotypes clash with leadership role schemas, leaders find it difficult to be effective.

The chapter concludes with a brief summary and overview in which in which I explicitly pull together some basic principles of leadership—things that we know pretty much for sure about leadership. I also identify some burgeoning directions for future leadership research in social psychology.

INDIVIDUAL DIFFERENCES
AND LEADERSHIP PERSONALITIES

In everyday life people tend to personify leadership—they focus on the person who leads rather than the context or process of leadership. We probably do this because the leader, as an individual, is figural against the background of the group and is therefore the focus of our attention. Under these conditions people are likely to explain the leader's behavior internally in terms of invariant personality dispositions (cf. Gilbert & Malone, 1995; Haslam, Rothschild, & Ernst, 1998; Ross, 1977)—there is evidence that this is indeed the case (e.g., Fiske & Dépret, 1996; Meindl, Ehrlich, & Dukerich, 1985).

Social psychologists are little different from people in everyday life and have therefore tried to explain leadership in terms of personality constellations that suit some people for effective leadership better than others. The "great person" analysis of leadership has a long and illustrious pedigree, going back to Plato and ancient Greece. One scholarly view, championed in the 19th century by Francis Galton (1892), was that leaders are born, not made. Most scholars, however, rejected the argument that leadership is innate, instead proposing that it was acquired very early in life, and therefore that an enduring constellation of personality attributes exists that imbues people with charisma and a predisposition to lead (e.g., Carlyle, 1841; House, 1977).

A prodigious amount of research has been conducted to identify reliable personality correlates of effective leadership. All that has been found is a handful of weak correlates, among which intelligence and talkativeness are the most reliable—leading Stogdill (1948), in his classic review, to conclude that leadership is not a "mere possession of some combination of traits" (p. 66), and others to proclaim that the search for a leadership personality is simplistic and futile (e.g., Conger & Kanungo, 1998). In general, correlations among traits, and between traits and effective leadership, are very low (Stogdill, 1974, reports an average correlation of .30).

Nevertheless, the belief that some people are better leaders than others because they have enduring traits that predispose them to effective leadership persists. This idea has reemerged, as we shall see later, in a different guise in modern theories of transformational leadership that place an emphasis on charisma (e.g., Avolio & Yammarino, 2003; Bass, 1985; Conger & Kanungo, 1998). Rather than focusing on specific traits, this tradition focuses on the wider Big Five personality dimensions of extraversion/surgency, agreeableness, conscientiousness, emotional stability, and intellect/openness to experience. A definitive meta-analysis by Judge, Bono, Ilies and Gerhardt (2002) reports a multiple correlation of .58 of these attributes with leadership—with extraversion/surgency, intellect/openness to experience, and conscientiousness as the best predictors of effective leadership.

SITUATIONAL PERSPECTIVES

In contrast to personality approaches that attribute effective leadership to possession of particular enduring trait constellations is the view that anyone can lead effectively if the situation is right. The most extreme form of this perspective is to deny any influence at all to the leader. For example, much of Tolstoy's epic novel *War and Peace* is a vehicle for his critique of the great person account of history: "To elicit the laws of history we must leave aside kings, ministers and generals, and select for study the homogeneous, infinitesimal elements which influence the masses" (Tolstoy, 1869, p. 977). Likewise, Karl Marx's theory of history places explanatory emphasis on the actions of collectivities, not individuals.

This perspective may be too extreme. For example, from an analysis of 300 military battles, Simonton (1980) found that situational factors did not account for all the outcomes—personal attributes of specific leaders were also significant correlates. In reality, effective leadership is a function of a match between one's behavioral style or personality attributes and the particular requirements of a specific leadership situation. Different situations call for different leadership properties, and therefore the most effective leader in a given context is the group member who is best equipped to assist the group in achieving its objectives (Bales, 1950). For example, Carter and Nixon (1949) had pairs of high school students perform three different tasks—an intellectual task, a clerical task and a mechanical assembly task. Those who took the lead in the first two tasks rarely led in the mechanical assembly task. In another example, Sherif, Harvey, White, Hood, and Sherif (1961; Sherif, 1966) documented in one of their boys' camp studies a change of leader in one of the groups when the situation changed to one of intergroup competition.

THE BEHAVIOR OF LEADERS

If effective leadership is an interaction between leader attributes and situational requirements, then we need to know about leader attributes. Based on dissatisfaction with the predictive reliability and utility of personality and on concerns about the validity of the construct of personality itself, another perspective on the study of leadership focuses on what leaders do—their actual behavior. This perspective spawned some of social psychology's classic leadership research.

One of the earliest studies was Lippitt and White's (1943) experiment on the effect of three different leadership styles (autocratic, democratic, and laissez-faire) on group atmosphere, morale, and effectiveness in afterschool activities clubs for young boys. They found that a democratic leadership style was most effective—it produced a friendly, group-centered, task-oriented atmosphere that was associated with relatively high group productivity, which was unaffected by whether or not the leader was physically present.

Another program of research, into interaction styles in groups (Bales, 1950), identified two key leadership roles: *task specialist* and *socioemotional specialist* (Slater, 1955). No one person could occupy both roles simultaneously, and the person occupying the task-specialist role was more likely to be the dominant leader. Task specialists tend to be centrally involved, often by offering opinions and giving directions, in the task-oriented aspects of group life, whereas socioemotional specialists tend to respond and pay attention to the feelings of other group members.

The Ohio State leadership studies constitute a third major leadership program (e.g., Fleishman, 1973; Stogdill, 1974). In this research, a scale for measuring leadership behavior was devised, the Leader Behavior Description Questionnaire (LBDQ) (e.g., Shartle, 1951), and a distinction was drawn between *initiating structure*

and *consideration*. Leaders rating high on initiating structure define the group's objectives and organize members' work toward the attainment of these goals: They are task oriented. Leaders rating high on consideration are concerned with the welfare of subordinates and seek to promote harmonious relationships in the group: They are relationship oriented. Unlike Bales (1950), who believed that task-oriented and socioemotional attributes were inversely related, the Ohio State researchers believed their dimensions to be independent—a single person could be high on both initiating structure (task-oriented) and consideration (socioemotional), and such a person would be a particularly effective leader. Research tends to support this latter view (e.g., Sorrentino & Field, 1986; Stogdill, 1974).

The general distinction between a leadership style that pays more attention to the group task and getting things done and one that pays attention to relationships among group members is quite pervasive in the leadership literature. For example, as we see later, it also appears in Fiedler's (1964) influential contingency theory of leadership, and in a slightly different guise in leader–member exchange (LMX) theory's emphasis on the quality of the leader's relationship with his or her followers (e.g., Graen & Uhl-Bien, 1995). Furthermore, it is a distinction that may hold across cultures, but with the caveat that what counts as task oriented or socioemotional leadership behavior may vary from culture to culture (e.g., Smith, Misumi, Tayeb, Peterson, & Bond, 1989)—for example, eating lunch with workmates is considered socioemotional leadership in some cultures but not others (Misumi & Peterson, 1985).

CONTINGENCY THEORIES

Contingency theories of leadership recognize that the leadership effectiveness of particular leadership behaviors or styles is contingent upon the properties of the leadership situation: Some styles are better suited to some situations or tasks than are others.

Fiedler's Contingency Theory

In social psychology the best known contingency theory is that proposed by Fiedler (1964, 1967). Fiedler, like Bales (1950), distinguished between task-oriented and relationship-oriented leaders—measuring leadership style in a rather unusual way with his Least Preferred Coworker (LPC) scale in which respondents rated their least preferred coworker on semantic differentials. High LPC scores meant that the respondent felt favorably inclined toward a fellow member even if he or she was not performing well (indicating a relationship-oriented style), and low LPC scores indicated a task-oriented style because the respondent was harsh on a poorly performing coworker.

Fiedler classified situations in terms of the quality of leader–member relations, the clarity of the structure of the task, and the intrinsic power and authority the leader had by virtue of his or her position as leader—good

leader–member relations in conjunction with a clear task and substantial position-power furnished maximal "situational control" (making leadership easy), whereas poor leader–member relations, a fuzzy task, and low position power furnished minimal "situational control" (making leadership difficult).

Fiedler's prediction was that low LPC, task-oriented, leaders would be most effective when situational control was low (the group needs a directive leader who focuses on getting things done) *and* when it was high (the group is doing just fine so there is little need to worry about morale and relationships within the group). High LPC, relationship-oriented leaders are more effective when situational control lies between these extremes. Against a background of some controversy and criticism (e.g., Peters, Hartke, & Pohlmann, 1985) focused, for example, on the measurement of situational control and on the characterization of leadership style as an invariant personal quality, Fiedler's predictions based on contingency theory have generally been supported (see meta-analyses by Strube & Garcia, 1981, and Schriesheim, Tepper, & Tetrault, 1994).

One troublesome finding, reported by Kennedy (1982), is that the 20% or so of leaders who have neither high nor low LPC scores are actually the most effective leaders of all, and their effectiveness is not affected by situational control. Another limitation of Fiedler's theory is that it is somewhat static—it does not focus on the dynamic interactive processes that occur within a group between leaders and followers and among followers, which characterize leadership situations.

Normative Decision Theory

Vroom and his associates have brought a contingency perspective to the explanation of leadership and the role of subordinate participation in group decision making (e.g., Vroom & Jago, 1988; Vroom & Yetton, 1973). Vroom's normative decision theory (NDT) identifies three decision-making strategies among which leaders can choose: autocratic (subordinate input is not sought), consultative (subordinate input is sought, but the leader retains the authority to make the final decision), and group decision making (leader and subordinate's are equal partners in a truly shared decision-making process). The relative efficacy of these strategies is contingent upon the quality of leader–subordinate relationships (which influences how committed and supportive subordinates are), and on task clarity and structure (which influences how much need the leader has for subordinate input).

In decision-making contexts autocratic leadership is fast and effective if subordinate commitment and support are high and the task is clear and well structured. When the task is less clear, greater subordinate involvement is needed and therefore consultative leadership is best. When subordinates are not very committed or supportive, group decision making is required to increase participation and commitment. Predictions from NDT are reasonably well supported empirically (e.g., Field & House, 1990)—leaders and managers report better deci-

sions and better subordinate ratings when they follow the prescriptions of the theory. However, there is a tendency for subordinates to prefer fully participative group decision making, even when it is not the most effective strategy.

Path-Goal Theory

Path-goal theory (PGT), developed by House and his colleagues (e.g., House, 1971; House & Mitchell, 1974), is another well-known contingency theory—although it can also be classified as a transactional leadership theory (see below). PGT is predicated on the assumption that a leader's main function is to motivate followers by clarifying the paths (i.e., behaviors and actions) that will help them attain their goals. It distinguishes between the two classes of leader behavior identified by the LBDQ, described earlier: *structuring* behaviors, whereby the leader directs task-related activities, and *consideration* behaviors, whereby the leader addresses followers' personal and emotional needs.

PGT predicts that structuring behaviors will be most effective when followers are unclear about their goals and how to reach them (e.g., when the task is new, difficult or ambiguous). When tasks are well understood structuring behaviors are less effective, or they can even backfire because they are viewed as undue meddling and micromanagement. Consideration behaviors are most effective when the task is boring or uncomfortable, but they can backfire when followers are already engaged and motivated because they are considered distracting and unnecessary.

Empirical support for PGT is mixed, and most scholars agree that tests of the theory tend to suffer from flawed methodology and from being incomplete and simplistic (Schriesheim & Neider, 1996). For these reasons research on PGT tapered off in the early 1980s. Recently, however, House (1996) has tried to reinvigorate the theory by addressing some of these concerns and by making it feel more contemporary—for example, the interpersonal focus of the original formulation has been expanded to include ways in which a leader can motivate an entire work group rather than just individual followers.

TRANSACTIONAL LEADERSHIP

One limitation of contingency theories is that in general they are somewhat static in not focusing on the dynamic interaction between leader and follower(s) that allows leaders to lead and encourages followers to follow. After all, effective leadership requires leaders and followers to provide support and gratification for one another (Messick, 2005). This lacuna is addressed by theories of transactional leadership.

Transactional leadership has been referred to as a "process of exchange that is analogous to contractual relations in economic life [and] contingent on the good faith of the participants" (Downton, 1973, p. 75). Leaders transact with followers to get things done, setting expectations and goals and providing recognition and re-

wards for task completion (Burns, 1978). Taking this idea further, Bass (1985) talked of contingent rewards and punishments—mutual benefits are exchanged (transacted) between leaders and followers against a background of contingent rewards and punishments that shape up cooperation and trust. Leader member transactions may also have an equity dimension (Walster, Walster, & Berscheid, 1978). Because effective leaders play a greater role in steering groups to their goals than do followers, followers may reinstate equity by rewarding the leader with social approval, praise, prestige, status, and power—in other words, with the trappings of effective leadership.

Although PGT, discussed above (e.g., House, 1971), is a contingency theory because of its focus on how the effectiveness of leadership behaviors is contingent upon situations, it can also be considered a transactional theory because it focuses on the transactions between leaders and followers that enhance motivation and lead to goal attainment.

Idiosyncrasy Credit

A well-known early approach to leadership that focused on leader–follower transactions is Hollander's (1958; Hollander & Julian, 1970) analysis of idiosyncrasy credit. Hollander argued that for leaders to be effective, they need to be allowed by the group to be innovative, to be able to experiment with new ideas and new directions—to be idiosyncratic. Drawing on the equity argument, Hollander wondered what circumstances of leadership would create a transactional relationship between leader and followers in which the followers would provide the leader with the resources to be able to be idiosyncratic. Hollander argued that leaders who had behaved in a highly conformist manner as they climbed the organizational ladder accumulated "idiosyncrasy credits" from the group. When the leader arrived at the top of the organization, followers would effectively hand over these credits (as a resource) to the leader, who could then "spend" them by behaving idiosyncratically, innovatively, and creatively.

Research provides some support for this analysis. Merei (1949) introduced older children who had shown leadership potential into small groups of younger children in a Hungarian nursery. The most successful leaders were those who initially complied with existing group practices and who only gradually and later introduced minor variations. In another study, Hollander and Julian (1970) found that leaders of decision-making groups who were ostensibly democratically elected enjoyed more support from the group, felt more competent at the task, and were more likely to suggest solutions that diverged from those of the group as a whole.

The idea that leaders who initially conform to group norms are ultimately allowed to be innovative and, paradoxically, nonnormative has recently been given a more group-oriented spin than allowed by a strictly interpersonal transaction approach. Drawing on the social identity approach in general (see Hogg, 2003, 2006, for an overview) and its more recent theory of leadership

(Hogg, 2001; Hogg & van Knippenberg, 2003; van Knippenberg & Hogg, 2003b; see below) it has been argued that group normative behavior on the part of a leader communicates to the group that the leader is "one of us"—a central member who identifies strongly with the group and embodies the norms and aspirations of the group and is therefore unlikely to do the group any harm (Platow, Hoar, Reid, Harley, & Morrison, 1997; Platow, Reid, & Andrew, 1998; Platow & van Knippenberg, 2001; van Knippenberg & van Knippenberg, 2005). If one identifies strongly with the group oneself, then such a person is to be trusted (e.g., Brewer, 1981; Hogg, in press; Tyler, 1997; Yamagishi & Kiyonari, 2000) and followed, because whatever he or she suggests or does is likely to be in the best interest of the group. The social identity analysis of leadership and the role of trust in leadership are both discussed more fully below.

Vertical Dyad Linkage Model and Leader–Member Exchange Theory

Leader–member transactions play a central role in the vertical dyad linkage (VDL) model of leadership (Dansereau, Graen, & Haga, 1975), which has evolved into leader–member exchange (LMX) theory (e.g., Graen & Uhl-Bien, 1995; Sparrowe & Liden, 1997). According to VDL researchers, leaders develop dyadic exchange relationships with different specific subordinates. In these dyadic relationships, the subordinate can either be treated as a close and valued "ingroup" member with the leader or in a more remote manner as an "outgroup" member who is separate from the leader.

As the VDL model evolved into LMX theory this dichotomous, ingroup versus outgroup, treatment of LMX relationships has disappeared and been replaced by a continuum of quality of exchange relationships ranging from ones that are based on mutual trust, respect and obligation (high-quality LMX relationships) to ones that are rather mechanically based on the terms of the formal employment contract between leader and subordinate (low-quality LMX relationships).

In high-quality LMX relationships subordinates are favored by the leader and receive many valued resources, which can include material benefits (e.g., money and privileges) as well as psychological benefits (e.g., trust and confidences). Leader–member exchanges go beyond the formal employment contract, with managers showing influence and support and giving the subordinate greater autonomy and responsibility. High-quality relationships should motivate subordinates to internalize the group's and the leader's goals. In low-quality LMX relationships subordinates are disfavored by the leader and receive fewer valued resources. Leader–member exchanges simply adhere to the terms of the employment contract, with little attempt by the leader to develop or motivate the subordinate. Subordinates will simply comply with the leader's goals, without necessarily internalizing them as their own.

LMX theory predicts that effective leadership hinges on the development of high-quality LMX relationships. These relationships enhance subordinates' well-being

and work performance and bind them to the group more tightly through loyalty, gratitude, and a sense of inclusion. Differentiated LMX relationships develop in groups because the leader usually has to relate to a large number of subordinates, and it is simply more efficient to select some subordinates in whom to invest a great deal of interpersonal energy and to treat the rest in a less intensive and personalized manner. The selection process takes time because it goes through a number of stages: role taking, in which the leader has expectations and tries out different roles on the subordinate; role making, in which mutual leader–member exchanges (e.g., of information and support) establish the subordinate's role; and role routinization, in which the leader–member relationship has become stable, smooth-running, and automatic.

Research confirms that differentiated LMX relationships do exist in most organizations; that high-quality LMX relationships are more likely to develop when the leader and the subordinate have similar attitudes, like one another, belong to the same sociodemographic groups, and both perform at a high level; and that high-quality LMX relationships are associated with (most studies are correlational, not causal) better performing and more satisfied workers who are more committed to the organization and less likely to leave (see Gerstner & Day, 1997; Graen & Uhl-Bien, 1995; Liden, Sparrowe, & Wayne, 1997; Schriesheim, Castro, & Cogliser, 1999). The stages of LMX relationship development are consistent with more general models of group development (e.g., Levine & Moreland, 1994; Tuckman, 1965).

The main general limitation of LMX theory is that it focuses on dyadic leader–member relations. It fails to consider that such dyadic relations are located in a wider context of shared group membership in which followers interact with one another as group members and are influenced by their perceptions of the leader's relations with other group members (Hogg & Martin, 2003; Hogg, Martin, & Weeden, 2004; Scandura, 1999). For example, from the perspective of the social identity theory of leadership (Hogg, 2001; see below), we would expect that when members identify strongly with a group they might find differentiated LMX relationships that favor some members over others to be uncomfortably personalized and fragmentary of the group and would not endorse such leaders. Instead, members might prefer a somewhat more depersonalized leadership style that treated all members relatively equally as group members—endorsing such leaders more strongly. This general hypothesis has recently been tested and supported in two field surveys of leadership perceptions in organizations in Wales and India (Hogg et al., 2005).

TRANSFORMATIONAL LEADERSHIP

Transactional theories of leadership represent a particular focus on leadership, but transactional leadership is itself a particular leadership style that can be contrasted to other leadership styles. In defining transactional leadership, Burns (1978) contrasted it sharply with transforma-

tional leadership. This contrast is put rather clearly by Judge and Bono (2000): Transactional leaders appeal to followers' self-interest, whereas transformational leaders inspire followers to adopt a vision that involves more than individual self-interest.

Three key components of transformational leadership are (1) individualized consideration (careful attention to followers' needs, abilities and aspirations, in order to help raise aspirations, improve abilities, and satisfy needs); (2) intellectual stimulation (challenging of followers' basic thinking, assumptions, and practices to help them develop newer and better mindsets and practices); and (3) charismatic/inspiring leadership, which provides the energy, reasoning, and sense of urgency that transforms followers (Avolio & Bass, 1987; Bass, 1985).

Transformational leadership theorists were quickly mortified that the charisma/inspiration component inadvertently admitted notorious dictators such as Hitler, Stalin, and Pol Pot into the hallowed club of transformational leaders—all were effective leaders insofar as they mobilized groups around their goals. So, a distinction was drawn between good charismatic leaders with socialized charisma that they use in a "morally uplifting" manner to improve society and bad charismatic leaders who use personalized charisma to tear down groups and society—the former are transformational, the latter are not (e.g., Avolio & Gibbons, 1988; Bass, 1990; O'Connor, Mumford, Clifton, Gessner, & Connelly, 1995).

In recent years the distinction between transactional and transformational leadership has been joined by a third type of leadership—laissez-faire (noninterfering) leadership, which involves not making choices or taking decisions and not rewarding others or shaping their behavior. Avolio and Bass (1991; Avolio, 1999) use laissez-faire leadership as a baseline anchor point in their full-range leadership model that has transformational leadership sitting at the apex (see Antonakis & House, 2003).

First published by Bass and Avolio in 1990, the Multifactor Leadership Questionnaire (MLQ) was designed to measure the components of transactional and transformational leadership. After revisions to address criticisms and problems, it is now in its fifth version, and it has been used in every conceivable organization, at every conceivable level, and on almost every continent. It has become the de facto leadership questionnaire of choice of the organizational and management research communities—producing numerous large-scale meta-analyses of findings (e.g., Lowe, Kroeck, & Sivasubramaniam, 1996; see also Avolio & Yammarino, 2003).

The Black Box of Transformation

One of the contemporary challenges for transformational leadership theory is to explicate what happens in the head of a follower to transform leader behavior into follower thought and behavior—what is the social psychology of transformation? Shamir, House, and Arthur (1993) suggest that followers personally identify with the leader and in this way make the leader's vision their own.

Dvir, Eden, Avolio, and Shamir (2002) suggest that the behavior of transformational leaders causes followers to identify more strongly with the organization's core values.

These ideas fit well with but are taken further by the more comprehensive social identity theory of leadership (e.g., Hogg, 2001) that is discussed later. Where group members identify strongly with a group, leaders who are considered central/prototypical group members are able to be innovative in defining a group's goals and practices. Strong identification is associated with internalization of group norms as one's own beliefs and actions. In this way leaders can transform groups.

Charisma and Charismatic Leadership

As we saw previously, the notion of charisma is so central to transformational leadership theory that a distinction was drawn between good and bad charisma, in order to distinguish between nontransformational villains (e.g., Hitler) and transformational heroes (e.g., Gandhi). Herein lies an initial problem with transformational leadership theory that can probably be traced to its grounding in organizational and management science (Jack Welch, former CEO of General Electric, is often held up as the apotheosis of a charismatic transformational leader) rather than social psychology or political science—one person's transformational leader can sometimes be another's war criminal or vice versa (much like one person's freedom fighter is another's terrorist). For example, whether Lenin and Mao are considered transformational leaders or not may rest more on one's political persuasion and ideological leanings than on transformational leadership theory's notion of good versus bad charisma (see earlier discussion of effective/ineffective vs. good/bad dimensions of leadership).

There is perhaps a more general issue concerning the role of charisma in transformational leadership. Scholars speak of charismatic leadership as a product of the leader's personal charisma and followers' reactions to the leader's charisma in a particular situation—personal charisma alone may not guarantee charismatic leadership (e.g., Bryman, 1992). However, it is difficult to escape the inference that charisma is a relatively enduring personality trait—in which case some of the problems of past personality theories of leadership may have been reintroduced (Haslam & Platow, 2001; Mowday & Sutton, 1993). Indeed, charismatic/transformational leadership has been linked to possession of the big five personality traits of extraversion/surgency, agreeableness, and intellect/openness to experience (e.g., Judge et al., 2002). Charismatic leadership is also linked to the related construct of visionary leadership (e.g., Conger & Kanungo, 1998) and the view that people differ in terms of how visionary they are as leaders. Visionary leaders are special people who are able to identify attractive future goals and objectives for a group and mobilize followers to internalize these as their own.

There is no doubt that charisma facilitates effective leadership; probably because charismatic individuals are emotionally expressive, enthusiastic, driven, eloquent, vi-

sionary, self-confident, and responsive to others (e.g., House, Spangler, & Woycke, 1991; Lindholm, 1990; Riggio & Carney, 2003). These attributes allow a person to be influential and persuasive and therefore able to make others buy their vision for the group and sacrifice personal goals for collective goals. Meindl and Lerner (1983; Meindl et al., 1985) talk about visionary leaders heightening followers' sense of shared identity, and how this shared identity produces a collective "heroic motive" that puts group goals ahead of personal goals.

The role of shared identity in charismatic leadership is given a different treatment by the social identity theory of leadership. Social identity researchers adhere to a meta-theory that views personality as playing at best a minor role in group processes (e.g., Abrams & Hogg, 2004; Billig, 1976). So, in the social identity theory of leadership (e.g., Hogg, 2001; Hogg & van Knippenberg, 2003; see below) charisma is seen more as a product of group processes than an ultimate cause of effective leadership. Social identity processes in salient groups that members identify strongly with render group prototypical (central) leaders influential and attractive, imbue them with trust, and allow them to be innovative. Followers attribute these qualities internally to the leader's personality—thus constructing a charismatic leadership personality. Empirical studies provide some support for the attributional construction of charisma (e.g., Fiske & Dépret, 1996; Meindl et al., 1985) and for the wider social identity perspective on charisma and leadership (see Haslam & Platow, 2001; Platow & van Knippenberg, 2001).

LEADER PERCEPTIONS AND LEADERSHIP SCHEMAS

Leader Categorization Theory

Keeping pace with the social cognition revolution in social psychology in the late 1970s (Taylor, 1998), and drawing on earlier work on implicit theories of leadership (Hollander & Julian, 1969), an approach to leadership has developed that focuses on the specific content of our leadership schemas and on the causes and consequences of categorization of someone as a leader. Leader categorization theory (LCT), or implicit leadership theory (e.g., Lord, Brown, Harvey, & Hall, 2001; Lord, Foti, & DeVader, 1984; Lord, Foti, & Phillips, 1982; Lord & Hall, 2003; Lord & Maher, 1991), assumes that leadership perceptions play a key role in leader selection decisions, in leader endorsement, and in a leader's power base, and thus in the extent to which people can effectively exercise leadership and influence others.

The theory rests on the notion that people have implicit leadership theories that shape their perceptions of (potential) leaders. In making leadership judgments, leadership schemas (called prototypes by Lord and colleagues) based on these implicit leadership theories are activated, and characteristics of the leader are matched against the relevant schema of effective leadership. Earlier conceptions of LCT (e.g., Lord et al., 1984) viewed leader schemas as relatively general and fixed. The contemporary version (e.g., Lord, Brown, Harvey, & Hall, 2001; Lord & Hall, 2003), which invokes a role for connectionist networks, views leadership schemas as flexible cognitive structures that are regenerated in situ to meet contextual demands.

The basic prediction, however, of LCT remains the same. The better the match between the leader's characteristics and the perceiver's leadership schema, the more favorable are leadership perceptions. For example, a perceiver whose leadership schema favors "intelligent," "organized," and "dedicated" as core leadership attributes is more likely to endorse a leader the more the leader is perceived to be intelligent, organized, and dedicated.

LCT focuses on categories and associated schemas of leadership and leaders (e.g., military generals, CEOs, outward bound leaders), not on social groups as categories (e.g., a psychology department, a corporation, a sports team). LCT's leader categories are tied to tasks and functions, and transcend groups—for example, a CEO schema applies similarly to companies such as Apple, Dell, GM, Toyota, Starbucks, Google, and so forth, whereas each company may have very different group norms and prototypes. LCT leaves unanswered the question of how schemas of group membership may influence leadership (but see Lord & Hall, 2003)—more specifically the role played by group identification and group prototypicality in effective leadership. This lacuna is addressed by the social identity theory of leadership (e.g., Hogg, 2001; see below)

Expectation States, Status Characteristics, and Role Congruity

Related to LCT are two other theories that also focus on leader categorization processes but do not go into social cognitive details quite so extensively—expectation states/status characteristics theory, and role congruity theory. Both theories suggest that the match between an individual's characteristics and abstracted conceptions of status and leadership affect leadership perceptions.

Expectation states theory or status characteristics theory (e.g., Berger, Fisek, Norman, & Zelditch, 1977; Berger, Wagner, & Zelditch, 1985; Ridgeway, 2001) attributes influence (and by implication leadership) within groups to possession of specific status characteristics (qualities that match what the group actually does) and diffuse status characteristics (stereotypical properties of high-status groups in society). Influence within an interactive group is a function of the extent to which a person possesses characteristics that suit him or her to effective task performance (i.e., specific status characteristics), and possesses characteristics that categorize him or her as a member of a high-status sociodemographic category (i.e., diffuse status characteristics). Influence, or leadership, is an additive function of perceived group task competence and perceived societal status (Ridgeway, 2003).

Role congruity theory focuses on gender and leadership (see below). According to Heilman (1983), gender-linked differential leadership potential may rest on a perceived congruity or incongruity between stereotype-based skills and attributes of an individual, and the per-

ceived nature of a job's requirements. In other words, because there is greater overlap between general leader schemas and male stereotypes than between leader schemas and female stereotypes, people tend to have more favorable perceptions of male leaders than of female leaders (Eagly, 2003; Eagly & Karau, 2002). Research provides some support for role congruity theory (e.g., Martell, Parker, Emrich, & Swerdlin Crawford, 1998; Shore, 1992).

SOCIAL IDENTITY AND LEADERSHIP

Leadership is a relationship in which some people are able to influence others to embrace, as their own, new values, attitudes, and goals and to exert effort on behalf of and in pursuit of those values, attitudes, and goals. The relationship is almost always configured by and played out within the parameters of a group—a small group like a team, a medium-size group like an organization, or a large group like a nation. Effective leadership inspires others to adopt group membership defining values, attitudes, and goals and to behave in ways that serve the group as a collective. Effective leaders are able to transform individual action into group action. Thus, leadership has an important identity function. People look to their leaders to express and epitomize their identity, to clarify and focus their identity, to forge and transform their identity, and to consolidate, stabilize, and anchor their identity.

Leadership is a core feature of social groups—it is difficult to think about groups without thinking about who leads or manages them, and about how well they are led and managed. Indeed, if one scratches below the surface of apparently leaderless groups a tacit leadership structure almost always emerges (e.g., Counselman, 1991). Given the pervasiveness and importance of leadership in group life, it seems odd that leadership is often not a core feature of the study of group processes and group processes are often not a core feature of the study of leadership (Chemers, 2001), and the identity function of leadership is very rarely center stage.

One way in which these issues have recently been addressed is through the development of the social identity theory of leadership (Hogg, 2001, 2005; Hogg & van Knippenberg, 2003; van Knippenberg & Hogg, 2003b). Firmly grounded in the social identity approach (e.g., Hogg & Abrams, 1988; Tajfel & Turner, 1986; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; see Hogg, 2003, 2006, for recent formulations), the social identity theory of leadership has, in a relatively short time, energized a significant amount of new leadership research in social psychology that focuses on the role of group membership and social identity (see Ellemers, de Gilder, & Haslam, 2004; van Knippenberg & Hogg, 2003a; van Knippenberg et al., 2004).

The central tenet is that as people identify more strongly with a group, they pay more attention to the group prototype and to what and who is more prototypical—this is because the prototype defines the group's membership attributes. In these salient group contexts being perceived to be highly prototypical makes one more influential. There are a number of social identity-related processes that cause this to happen. First, prototypical members best embody the group's attributes, and so they are seen to be the source rather than target of conformity processes—they are the ones to whom other members seem to align their behavior (e.g., Turner, 1991).

Second, prototypical members are liked as group members (a process of depersonalized social attraction), and, because there is usually significant agreement on the prototype, the group as a whole likes the leader—he or she is consensually popular in group terms (Hogg, 1993). This process not only facilitates influence (we are more likely to comply with requests from people we like—Berscheid & Reis, 1998) but also accentuates the evaluative (status) differential between leader and followers.

Third, prototypical leaders generally find the group more central and important to self-definition and therefore identify more strongly with it. They have a greater investment in the group and thus are more likely to behave in group-serving ways. They embody group norms more precisely, and they are more likely to favor the ingroup over outgroups, to treat ingroup members fairly, and generally to act in ways that promote the ingroup. These behaviors confirm their prototypicality and membership credentials and encourage group members to trust them to be acting in the best interest of the group even when it may not appear that they are—they are furnished with legitimacy (Tyler, 1997; Tyler & Lind, 1992; see Platow et al., 1998). One important consequence of this is that prototypical leaders are actually able to be innovative—they can, paradoxically, diverge from group norms and be less conformist than non- or less prototypical leaders. Innovation is, of course, a key component of effective leadership as it allows one to transform the group and steer it safely through the stormy waters of change (see earlier discussion of transformational leadership).

This analysis of how prototypical leaders can be innovative describes the processes that may account for Hollander's (1958) idea, discussed earlier, that to be effective a leader needs initially to conform to group norms to earn "idiosyncrasy credits" in order later to be able to diverge from such norms and be innovative. The social identity analysis is wider in that conformity and the temporal dimension are not always necessary—the key factor is that the leader behaves in ways that build trust based on shared identity and the perception that the leader is centrally invested in the group.

Finally, because the prototype is so central to group life, information related to the prototype is figural against the background of other information in the group. A prototypical leader is probably the most direct source of prototype information, and so is figural against the background of the group. Members pay close attention to the leader and, as in other areas of social perception and inference, attribute his or her behavior to invariant or essential properties of the leader's personality—they engage in the fundamental attribution

error (Ross, 1977), correspondence bias (Gilbert & Malone, 1995), or essentialism (Haslam et al., 1998).

This process readily constructs a charismatic leadership personality, because the behaviors that are being internally attributed include being the source of influence, being able to gain compliance from others, being popular, having higher status, being innovative, and being trusted. In this way charisma plays a role (cf. transformational leadership), but charisma is constructed by group processes rather than being a static personality attribute that is brought to the group (cf. Haslam & Platow, 2001). In any event, the construction of charisma further facilitates effective and innovative leadership on the part of a prototypical leader.

Social identity-based leadership processes confer on a leader considerable power to maintain his or her position of leadership. Because such leaders are invested with charisma, status, and so forth, they are very effective prototype managers. Largely through communication and talk they can construct, reconstruct, or change the group prototype in ways that protect or promote their highly prototypical central position in the group—a process that can be referred to as norm talk (Hogg & Tindale, 2005; see also Fiol, 2002; Gardner, Paulsen, Gallois, Callan, & Monaghan, 2001; Reid & Ng, 2000). Indeed, one of the key attributes of an effective leader is precisely this visionary and transformational activity—a leader is able to change what the group sees itself as being and can be considered an entrepreneur of identity (Reicher & Hopkins, 2003).

Specifically, prototypical leaders can talk up their own prototypicality and/or talk down aspects of their own behavior that are nonprototypical. They can identify deviants or marginal members to highlight their own prototypicality or to construct a particular prototype for the group that enhances their own prototypicality. They can secure their own leadership position by vilifying contenders for leadership and casting the latter as nonprototypical. They can identify as relevant comparison outgroups those outgroups that are most favorable to their own prototypicality—that is, they can manipulate the social comparative frame and thus the prototype and their own prototypicality. They can engage in a discourse that raises or lowers salience—if we are highly prototypical then raising salience will provide us with the leadership benefits of high prototypicality; if we are not very prototypical, then lowering salience will protect us from the leadership pitfalls of not being very prototypical. Research suggests that all these processes are used by leaders to manage their prototypicality (e.g., Reicher & Hopkins, 1996, 2001, 2003). Generally, leaders who feel they are not, or are no longer, prototypical, strategically engage in a range of group-oriented behaviors to strengthen their membership credentials (e.g., Platow & van Knippenberg, 2001).

As a final point on the social identity theory of leadership, it should be noted that social identity leadership processes only or more strongly occur in groups that members identify more strongly with. As group salience or strength of identification weakens, social identity leadership processes weaken and leadership becomes less based on group prototypicality and more based on other factors such as leadership schemas.

The social identity theory of leadership, although a relative newcomer to the leadership scene, has already attracted substantial empirical support from laboratory experiments and more naturalistic studies and surveys (for summaries of empirical work, see Hogg, 2001; Hogg & van Knippenberg, 2003; van Knippenberg & Hogg, 2003b; van Knippenberg et al., 2004).

TRUST AND LEADERSHIP

A key feature of leadership is trust (e.g., Dirks & Ferrin, 2002)—witness current concerns about corporate corruption and distrust of our business and government leaders (e.g., Boyle & Tkaczyk, 2004; Kellerman, 2004). Lack of trust in leadership is a matter of great concern and consequence—after all, if we are going to follow someone we certainly do need to be able to trust them to be acting in the very best interest of us all as a group, rather than in their own self-interest. The notion that trust in leaders allows them a sufficiently wide latitude of behaviors for innovative decisions and actions goes back a long way, for example, to the philosopher John Locke (1690/1988). More recent social psychological research on trust and leadership takes a group-oriented perspective that focuses on justice and fairness perceptions (can we trust our leader to be just and fair?) and is generally or more specifically informed by the social identity perspective (e.g., Hogg, in press; Tyler, 1997, 2003, 2005; Tyler & Blader, 2000; Tyler & Lind, 1992).

Justice and Fairness

One of the key tenets of Tyler's group value model (Lind & Tyler, 1988) and his relational model of authority in groups (Tyler, 1997; Tyler & Lind, 1992) is that fairness and justice perceptions are critical to group life. Because leaders make decisions that have important consequences for followers (e.g., promotions, performance appraisals, and allocation of duties), followers are concerned about how fair the leader is in making these decisions. In judging fairness, followers focus on both distributive justice (how fair are the outcomes of the leader's decisions) and procedural justice (how fair are the procedures that the leader has used to make a decision). Justice and fairness judgments affect reactions to decisions and to the authorities making these decisions and thus influence leadership effectiveness (see De Cremer, 2003; De Cremer & Tyler, in press).

One way in which justice perceptions, particularly procedural justice perceptions, may affect leadership is via social identity processes (e.g., Koper, van Knippenberg, Bouhuijs, Vermunt, & Wilke, 1993; Tyler, 2003; van Knippenberg & Hogg, 2003b). Procedural justice serves a social identity function because fair procedures convey a favorable social evaluation of followers as group members—the respect for group members conveyed by

procedural fairness builds member identification and thus feeds into cooperative and compliant behavior. Research shows that as members identify more strongly with the group they care more strongly that the leader is procedurally fair (e.g., Brockner, Chen, Mannix, Leung, & Skarlicki, 2000), and care less strongly that the leader is distributively fair—this asymmetry arises because with increasing identification, instrumental outcome-oriented considerations (distributive justice) become less important relative to intragroup relational and membership considerations (procedural justice) (e.g., Vermunt, van Knippenberg, van Knippenberg, & Blaauw, 2001).

Social Dilemmas

The fact that justice, particularly procedural justice, may facilitate effective leadership because it builds trust and strengthens group identification raises the possibility that leadership may be a way to resolve social dilemmas. Social dilemmas are essentially a crisis of trust—people behave selfishly because they do not trust others to sacrifice some immediate self-interest for the longer-term greater good of the collective (e.g., Dawes & Messick, 2000; Hardin, 1968; Liebrand, Messick, & Wilke, 1992).

Social dilemmas are notoriously difficult to resolve (Kerr & Park, 2001). However, they are not impossible to resolve if one can address the trust issue. One relatively successful solution is to build mutual trust among people by causing them to identify strongly as a group—people tend to trust ingroup members more (e.g., Brewer, 1981; Hogg, in press; Yamagishi & Kiyonari, 2000), and under these circumstances they are more likely to sacrifice self-interest for the good of all (e.g., Brewer & Schneider, 1990; De Cremer & van Vugt, 1999). One downside to this is that mutual ingroup trust and cooperation are typically balanced by intergroup distrust and competition—the interpersonal dilemma may be transformed into a destructive intergroup dilemma (e.g., Kramer & Jost, 2002; Kramer & Messick, 1998).

Nevertheless, building ingroup trust is one way to resolve a social dilemma. Leadership plays an often critical role in this process precisely because a leader can transform selfish individual goals into shared group goals by building a sense of common identity, shared fate, interindividual trust, and custodianship of the collective good (De Cremer, 2000, 2002; De Cremer & van Knippenberg, 2003; De Cremer & van Vugt, 2002; van Vugt & De Cremer, 1999).

GENDER AND LEADERSHIP

Throughout the world, leadership roles are dominated by men. If one restricts oneself to liberal democracies like the United States where more progressive gender attitudes have developed over the past 40 years, it is still the case that although women are now relatively well represented in middle management, they are still significantly underrepresented in senior management and "elite" leadership positions—there is a "glass ceiling" (e.g., Cejka & Eagly, 1999; Eagly & Karau, 1991; Eagly et al., 1992, 1995).

Perhaps this is because men and women differ in ways that suit men better for leadership. Research does not uphold this idea. Although women and men tend to adopt different leadership styles, women are usually rated as just as effective leaders as men—and in general they are perceived to be slightly more transformational, participative, and engaged in contingent reward behaviors (Eagly, Johannesen-Schmidt, van Engen, & Vinkenburg, 2002).

If women and men are equally capable of being effective leaders, why is there a gender gap in leadership? One explanation is in terms of role incongruity theory (Eagly, 2003; Eagly & Karau, 2002; Heilman, 1983), which argues that because there is greater overlap between general leader schemas and agentic male stereotypes than between leader schemas and communal female stereotypes, people tend to have more favorable perceptions of male leaders than of female leaders. These leadership perceptions facilitate or impede effective leadership. One implication of role incongruity theory is that the evaluation of male and female leaders will change if the leadership schema changes or if people's gender stereotypes change. For example, research has shown that men leaders are evaluated more favorably than women leaders when the role is defined in more masculine terms, and vice versa when the role is defined in less masculine terms (Eagly et al., 1995).

Related to role incongruity theory is a social identity analysis (see above) that argues that in high-salience groups with which members identify, male or female leaders are perceived to be and actually are effective if the group's norms are consistent with the members' gender stereotypes. So, people with traditional gender stereotypes will endorse a male not a female leader of a group with instrumental norms and a female not a male leader of a group with more expressive norms, whereas among people with less traditional gender stereotypes this effect may not be so pronounced and may be reversed (Hogg et al., 2006).

Bowles and McGinn (2005) attribute the gender leadership gap to a tendency for women to claim authority less effectively than men (men claim and hold a greater number of leadership positions than do women), not to any difference in how effective men or women leaders are once they achieve authority. Bowles and McGinn propose four main barriers to women claiming authority. The first is effectively role incongruity, as discussed previously. The second is lack of critical management experience. The third is familial responsibility that restricts a woman's ability to find the time commitment required by leadership positions.

The fourth obstacle is lack of motivation—women are simply not as "hungry" for leadership positions as are men. They shy away from self-promotion and take on less visible background roles with informal titles like "facilitator" or "coordinator." Although the link has not been made explicit, it is very possible that the underlying reason for women's reticence to claim authority is stereotype

threat (Steele & Aronson, 1995; Steele, Spencer, & Aronson, 2002)—women fear that stereotypes about women and leadership will be confirmed, and so they experience reduced motivation to take on leadership. In addition, self-promotion and leadership claiming are non-female-stereotypical behaviors that can be interpreted as "pushy" and can attract negative reactions from most group members (Rudman, 1998; Rudman & Glick, 1999).

SUMMARY AND CONCLUSIONS

This review of leadership has covered a lot of ground. I have tried to capture the breadth of the topic by including the main perspectives and approaches, but this has entailed some sacrifice of detail. I have tried to place more emphasis on social psychological research, in particular recent work, than on management and organizational science work.

It is important, however, to recognize that there is substantial overlap between social psychological research and research in management and organizational settings—this is because leadership is a topic that sits right on the boundary and people who do leadership research, of whatever discipline, are often social psychologists. One of the exciting and timely new opportunities for leadership research is a much closer and more concrete working relationship between social psychology and management/organizational science, and also political science and sociology. This integration will benefit social psychology because leadership plays out in organizations and society, and management/organizational science, political science, and sociology will benefit from the significant advances made in social psychology over the past 30 years in our understanding of social cognition and group processes.

Basic Principles

From a social psychological point of view the take-away message about leadership is that leadership is a group process and that effective leadership is the product of a complex interaction between, on the one hand, what a leader does and how this is perceived by the followers and, on the other hand, the expectations of and constraints imposed on leader behavior by societal norms, group norms, and group tasks. Leadership plays out through interactions, which are largely symbolic, between leaders and followers that are configured by the wider group that leader and followers belong to. Leadership has an important identity dimension. Effective leaders need to be innovative and transformational—they need to be able to influence group norms and persuade members to internalize those norms as their own guides to action. They need to be able to make members sacrifice self-interest for collective interest. They need to make members identify with the group. To accomplish all this, leaders need to be persuasive and perhaps viewed as charismatic and visionary. They need to be trusted to be acting in the interest of the group, not in their own self-interest.

This general message can be reframed and expanded into a set of basic principles about leadership—things that we know fairly certainly about leadership.

1. Leadership is a process of influence that does not require the exercise of coercive power over others—coercion may negate true leadership and merely produce compliance and obedience.
2. Leadership is predominantly a group process in which one person transforms other members of the group so that they adopt a vision (often a new vision) and are galvanized into pursuing the vision on behalf of the group—leadership is not simply managing a group's activities. Transformational leadership is facilitated by charisma, consideration of followers, and inspiring followers.
3. Leadership involves transactions between leader and followers—leaders do something for the group and the group in return does something for the leader to allow the leader to lead effectively.
4. Leadership often has an important identity dimension—followers look to their leaders to mould, transform, and express who they are, their identity. Being perceived to be "one of us" can often facilitate leadership.
5. Trust plays an important role in leadership—leaders have greater scope to be innovative if the group trusts them.
6. Effective and good leadership are not the same thing—effective leaders successfully influence the group to adopt and achieve (new) goals, whereas good leaders pursue goals that we value, use means that we approve of, and have qualities that we applaud.
7. Although some broad personality attributes tend to be associated with effective leadership (e.g., extraversion/surgency, intellect/openness to experience, and conscientiousness), personality alone is rarely sufficient.
8. There is a general distinction between task-focused (structuring) and person/relationship-focused (consideration) leadership style—their relative effectiveness and the effectiveness of other leadership styles depends on context (e.g., the nature of the group, the nature of the task).
9. Leadership effectiveness can be improved if the leaders' attributes and behavior are perceived to fit general or task-specific schemas we have of effective leadership, or the norms/prototype of a group membership/identity that we share with the leader.

Agenda for Leadership Research in Social Psychology

In terms of a concrete research agenda for social psychology, one thing that is badly needed, and was foreshadowed earlier, is a social psychological specification of what goes on in the social cognitive black box of transformational leadership—what happens in the head of a follower to transform leader behavior into follower thought and behavior? We cited research by Shamir and colleagues (1993) and Dvir and colleagues (2002) that quite

clearly suggests that processes of follower identification with the group may play a key role. The social identity theory of leadership may help flesh this out.

Related to this is a need to explore the notion of charisma more fully (cf. Haslam & Platow, 2001). Is it an enduring personality attribute that is brought to a leadership situation? Is it perceptually constructed for an effective leader by group members? Or is it a mixture of the two, and if so how are the two related and in what proportions? Social psychologists are often uncomfortable with explanations of group processes that rely too heavily on enduring personality attributes.

Leadership has a critical identity function that needs to be more fully explored. Not only are leaders often more effective if they are perceived to embody the group's identity, but in many leadership contexts followers actually look to their leaders to define and express and to consolidate or transform who they themselves are—their identity.

Trust also plays a key role in effective leadership. There is a strong strand of social psychology research on trust and leadership, mainly by Taylor and De Cremer and their colleagues. However, this is an area that would probably benefit from more extensive general social psychological research on trust, to link the general notion of trust to risk management, group life, identity, and leadership (Hogg, in press).

Leaders lead groups in the context of, or against, other groups. Intergroup relations impact leadership, and leadership impacts intergroup relations. This intergroup aspect of leadership is a core aspect of Eagly and colleagues' research on gender and leadership—intergroup perceptions, in the guise of gender stereotypes, impact leadership expectations and thus leadership effectiveness. This analysis could be expanded and integrated with a broad social identity analysis to focus more on intergroup behavior, as opposed to perceptions, and also widened to deal with other intergroup situations aside from gender relations. This wider intergroup analysis might help us understand intergroup leadership in the political sphere. It might also help us identify common reasons why many minorities may find it hard to lead—for example, as hinted previously, minorities may not take opportunities to lead because of stereotype threat processes (cf. Steele & Aronson, 1995).

Finally, by focusing on leadership as a basic influence process that involves people's identities as group members, social psychology may be able to speak to a wider conceptualization of leadership that moves beyond business leadership to address public leadership, nonprofit leadership, and so forth. This wider perspective may also make it important to move beyond a discussion of leadership effectiveness to incorporate a fuller causal analysis of good and bad leadership.

ACKNOWLEDGMENTS

REFERENCES

Abrams, D., & Hogg, M. A. (1998). Prospects for research in group processes and intergroup relations. *Group Processes and Intergroup Relations*, *1*, 7–20.

Abrams, D., & Hogg, M. A. (2004). Metatheory: Lessons from social identity research. *Personality and Social Psychology Review*, *8*, 98–106.

Allport, G. W. (1968). The historical background of modern social psychology. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 1, pp. 3–56). Reading, MA: Addison-Wesley.

Antonakis, J., & House, R. J. (2003). An analysis of the full-range leadership theory: The way forward. In B. J. Avolio & F. J. Yammarino (Eds.), *Transformational and charismatic leadership: The road ahead* (pp. 3–33). New York: Elsevier.

Avolio, B. J. (1999). *Full leadership development: Building the vital forces in organizations*. Thousand Oaks, CA: Sage.

Avolio, B. J., & Bass, B. M. (1987). Transformational leadership, charisma and beyond. In J. G. Hunt, B. R. Balaga, H. P. Dachler, & C. A. Schriesheim (Eds.), *Emerging leadership vistas* (pp. 29–50). Elmsford, NY: Pergamon Press.

Avolio, B. J., & Bass, B. M. (1991). *The full-range leadership development programs: Basic and advanced manuals*. Binghamton, NY: Bass, Avolio & Associates.

Avolio, B. J., & Gibbons, T. C. (1988). Developing transformational leaders: A life span approach. In J. A. Conger & R. N. Kanungo (Eds.), *Charismatic leadership: The elusive factor in organizational effectiveness* (pp. 276–308). San Francisco: Jossey-Bass.

Avolio, B. J., & Yammarino, F. J. (Eds.). (2003). *Transformational and charismatic leadership: The road ahead*. New York: Elsevier.

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Reading, MA: Addison-Wesley.

Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York: Free Press.

Bass, B. M. (1990). *Bass and Stogdill's handbook of leadership* (3rd ed.). New York: Free Press.

Bass, B. M., & Avolio, B. J. (1990). *Transformational leadership development: Manual for the Multifactor Leadership Questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Bass, B. M., & Avolio, B. J. (Eds.). (1994). *Improving organizational effectiveness through transformational leadership*. Thousand Oaks, CA: Sage.

Bennis, W. (1989), *On becoming a leader*. Reading, MA: Addison-Wesley.

Berger, J., Fisek, M. H., Norman, R. Z., & Zelditch, M., Jr. (1977). *Status characteristics and social interaction*. New York: Elsevier.

Berger, J., Wagner, D., & Zelditch, M., Jr. (1985). Expectation states theory: Review and assessment. In J. Berger & M. Zelditch, Jr. (Eds.), *Status, rewards and influence* (pp. 1–72). San Francisco: Jossey-Bass.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 193–281). New York: McGraw-Hill.

Billig, M. (1976). *Social psychology and intergroup relations*. London: Academic Press.

Bowles, H. R., & McGinn, K. L. (2005). Claiming authority: Negotiating challenges for women leaders. In D. M. Messick & R. M. Kramer (Eds.), *The psychology of leadership: New perspectives and research* (pp. 191–208). Mahwah, NJ: Erlbaum.

Boyle, M., & Tkaczyk, C. (2004). When will they stop? *Fortune*, *149*, 123–126.

Brewer, M. B. (1981). Ethnocentrism and its role in interpersonal trust. In M. B. Brewer & B. Collins (Eds.), *Scientific inquiry and the social sciences* (pp. 345–360). San Francisco: Jossey-Bass.

Brewer, M. B., & Schneider, S. (1990). Social identity and social dilemmas: A double-edged sword. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 169–184). London: Harvester Wheatsheaf.

Brockner, J., Chen, Y.-R., Mannix, E. A., Leung, K., & Skarlicki, D. P. (2000). Culture and procedural fairness: When the effects of what you do depend on how you do it. *Administrative Science Quarterly*, *45*, 1238–159.

Bryman, A. (1992). *Charisma and leadership*. London: Sage.

Burns, J. M. (1978). *Leadership*. New York: Harper & Row.

Burns, J. M. (2003). *Transforming leadership: The new pursuit of happiness*. New York: Atlantic Monthly Press.

Carlyle, T. (1841). *On heroes, hero-worship, and the heroic*. London: Fraser.

Carter, L. F., & Nixon, M. (1949). An investigation of the relationship between four criteria of leadership ability for three different tasks. *Journal of Psychology*, *27*, 245–261.

Cartwright, D., & Zander, A. (Eds.). (1968). *Group dynamics: Research and theory* (3rd ed.). London: Tavistock.

Cawthon, D., & Clark, B. (2002). *Philosophical foundations of leadership*. Somerset, NJ: Transaction.

Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, *25*, 413–423.

Chemers, M. M. (1997). *An integrative theory of leadership*. Mahwah, NJ: Erlbaum.

Chemers, M. M. (2001). Leadership effectiveness: An integrative review. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 376–399). Oxford, UK: Blackwell.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 151–192). New York: McGraw-Hill.

Conger, J. A., & Kanungo, R. N. (1998). *Charismatic leadership in organizations*. Thousand Oaks, CA: Sage.

Counselman, E. (1991). Leadership in a long-term leaderless women's group. *Small Group Research*, *22*, 240–257.

Dansereau, F., Graen, G., & Haga, W. J. (1975). A vertical dyad linkage approach to leadership within formal organizations: A longitudinal investigation of the role making process. *Organizational Behavior and Human Performance*, *13*, 46–78.

Dawes, R. M., & Messick, D. M. (2000). Social dilemmas. *International Journal of Psychology*, *35*, 111–116

De Cremer, D. (2000). Leadership selection in social dilemmas– Not all prefer it: The moderating effect of social value orientation. *Group Dynamics*, *4*, 330–337.

De Cremer, D. (2002). Charismatic leadership and cooperation in social dilemmas: A matter of transforming motives? *Journal of Applied Social Psychology*, *32*, 997–1016.

De Cremer, D. (2003). A relational perspective on leadership and cooperation: Why it matters to care and be fair. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 109–122). London: Sage.

De Cremer, D., & Tyler, T. R. (in press). Managing group behavior: The interplay between procedural fairness, sense of self, and cooperative behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology*. San Diego CA: Academic Press.

De Cremer, D., & van Knippenberg, D. (2003). Cooperation with leaders in social dilemmas: On the effects of procedural fairness and outcome favorability in structural cooperation. *Organizational Behavior and Human Decision Processes*, *91*, 1–11.

De Cremer, D., & van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology*, *29*, 871–893.

De Cremer, D., & Van Vugt, M. (2002). Intergroup and intragroup aspects of leadership in social dilemmas: A relational model of cooperation. *Journal of Experimental Social Psychology*, *38*, 126–136.

Devine, P. G., Hamilton, D. L., & Ostrom, T. M. (Eds.). (1994). *Social cognition: Impact on social psychology*. San Diego, CA: Academic Press.

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology*, *87*, 611–628.

Downton, J. V. (1973). *Rebel leadership*. New York: Free Press.

Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. (2002). Impact of transformational leadership training on follower development and performance: A field experiment. *Academy of Management Journal*, *45*, 735–744.

Eagly, A. H. (2003). Few women at the top: How role incongruity produces prejudice and the glass ceiling. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 79–93). London: Sage.

Eagly, A. H., Johannesen-Schmidt, M., van Engen, M. L., & Vinkenburg, C. (2002). Transformational, transactional, and laissez-faire styles: A meta-analysis comparing men and women. *Psychological Bulletin*, *129*, 569–591.

Eagly, A. H., & Karau, S. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, *60*, 685–710.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*, 573–598.

Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, *117*, 125–145.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*, 3–22.

Ellemers, N., de Gilder, D., & Haslam, S. A. (2004). Motivating individuals and groups at work: A social identity perspective on leadership and group performance. *Academy of Management Review*, *29*, 459–478.

Fiedler, F. E. (1964). A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 149–190). New York: Academic Press.

Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw-Hill.

Field, R. H. G., & House, R. J. (1990). A test of the Vroom-Yetton model using manager and subordinate reports. *Journal of Applied Psychology*, *75*, 362–366.

Fiol, C. M. (2002). Capitalizing on paradox: The role of language in transforming organizational identities. *Organization Science*, *13*, 653–666.

Fiske, S. T., & Dépret, E. (1996). Control, interdependence and power: Understanding social cognition in its social context. *European Review of Social Psychology*, *7*, 31–61.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fleishman, E. A. (1973). Twenty years of consideration and structure. In E. A. Fleishman & J. F. Hunt (Eds.), *Current developments in the study of leadership*. Carbondale: South Illinois University Press.

Galton, F. (1892). *Heredity genius: An inquiry into its laws and consequences*. London: Macmillan.

Gardner, M. J., Paulsen, N., Gallois, C., Callan, V. J., & Monaghan, P. (2001). Communication in organizations: An intergroup perspective. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 561–584). Chichester, UK: Wiley.

Gergen, D. (2000). *Eyewitness to power: The essence of leadership*. New York: Simon & Schuster.

Gerstner, C. R., & Day, D. V. (1997). Meta-analytic review of leader–member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, *82*, 827–844.

Gilbert, D. T., Fiske, S. T., & Lindzey, G. (Eds.). (1998). *The handbook of social psychology* (4th ed.). New York: McGraw-Hill.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Goethals, G. R., Sorenson, G. J., & Burns, J. M. (Eds.). (2004). *Encyclopedia of leadership*. Thousand Oaks, CA: Sage.

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader–member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain approach. *Leadership Quarterly*, *6*, 219–247.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Haslam, N., Rothschild, L., & Ernst, D. (1998). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*, 113–127.

Haslam, S. A., & Platow, M. J. (2001). Your wish is our command: The role of shared social identity in translating a leader's vision into followers' action. In M. A. Hogg & D. J. Terry (Eds.), *Social identity processes in organizational contexts* (pp. 213–228). Philadelphia: Psychology Press.

Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. *Research in Organizational Behavior*, *5*, 269–298.

Hogg, M. A. (1993). Group cohesiveness: A critical review and some new directions. *European Review of Social Psychology*, *4*, 85–111.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, *5*, 184–200.

Hogg, M. A. (2003). Social identity. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 462–479). New York: Guilford Press.

Hogg, M. A. (2005). Social identity and leadership. In D. M. Messick & R. M. Kramer (Eds.), *The psychology of leadership: New perspectives and research* (pp. 53–80). Mahwah, NJ: Erlbaum.

Hogg, M. A. (2006). Social identity theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (pp. 111–136). Palo Alto, CA: Stanford University Press.

Hogg, M. A. (in press). Social identity and the group context of trust: Managing risk and building trust through belonging. In M. Siegrist, T. C. Earle, & H. Gutscher (Eds.), *Trust in cooperative risk management: Uncertainty and scepticism in the public mind*. London: Earthscan.

Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.

Hogg, M. A., Fielding, K. S., Johnson, D., Masser, B., Russell, E., & Svensson, A. (2006). Demographic category membership and leadership in small groups: A social identity analysis. *Leadership Quarterly*, *17*, 335–350.

Hogg, M. A., & Martin, R. (2003). Social identity analysis of leader–member relations: Reconciling self-categorization and leader–member exchange theories of leadership. In S. A. Haslam, D. van Knippenberg, M. J. Platow, & N. Ellemers (Eds.), *Social identity at work: Developing theory for organizational practice* (pp. 139–154). New York: Psychology Press.

Hogg, M. A., Martin, R., Epitropaki, O., Mankad, A., Svensson, A., & Weeden, K. (2005). Effective leadership in salient groups: Revisiting leader-member exchange theory from the perspective of the social identity theory of leadership. *Personality and Social Psychology Bulletin*, *31*, 991–1004.

Hogg, M. A., Martin, R., & Weeden, K. (2004). Leader–member relations and social identity. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 18–33). London: Sage.

Hogg, M. A., & Tindale, R. S. (2005). Social identity, influence, and communication in small groups. In J. Harwood & H. Giles (Eds.), *Intergroup communication: Multiple perspectives* (pp. 141–164). New York: Peter Lang.

Hogg, M. A., & van Knippenberg, D. (2003). Social identity and leadership processes in groups. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 35, pp. 1–52). San Diego, CA: Academic Press.

Hollander, E. P. (1958). Conformity, status, and idiosyncracy credit. *Psychological Review*, *65*, 117–127.

Hollander, E. P. (1985). Leadership and power. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 485–537). New York: Random House.

Hollander, E. P., & Julian, J. W. (1969). Contemporary trends in the analysis of leadership processes. *Psychological Bulletin*, *71*, 387–391.

Hollander, E. P., & Julian, J. W. (1970). Studies in leader legitimacy, influence, and innovation. In L. Berkowitz (Ed.), *Advances in ex-*

perimental social psychology (Vol. 5, pp. 34–69). New York: Academic Press.

House, R. J. (1971). A path-goal theory of leadership effectiveness. *Administrative Science Quarterly*, *16*, 321–338.

House, R. J. (1977). A 1976 theory of charismatic leadership. In J. G. Hunt & L. Larson (Eds.), *Leadership: The cutting edge* (pp. 189–207). Carbondale: Southern Illinois University Press.

House, R. J. (1996). Path-goal theory of leadership: Lessons, legacy, and a reformulated theory. *Leadership Quarterly*, *7*, 323–352.

House, R. J., & Mitchell, T. R. (1974). Path-goal theory of leadership. *Journal of Contemporary Business*, *3*, 81–98.

House, R. J., Spangler, W. D., & Woycke, J. (1991). Personality and charisma in the US presidency: A psychological theory of leader effectiveness. *Administrative Science Quarterly*, *36*, 364–396.

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology*, *85*, 751–765.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*, 765–780.

Kellerman, B. (1999). *Re-inventing leadership*. Albany: State University of New York Press.

Kellerman, B. (2004). *Bad leadership: What it is, how it happens, why it matters*. Cambridge, MA: Harvard Business School Press.

Kennedy, J. (1982). Middle LPC leaders and the contingency model of leader effectiveness. *Organizational Behavior and Human Performance*, *30*, 1–14.

Kerr, N. L., & Park, E. S. (2001). Group performance in collaborative and social dilemma tasks: Progress and prospects. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 107–138). Oxford, UK: Blackwell.

Koper, G., van Knippenberg, D., Bouhuijs, F., Vermunt, R., & Wilke, H. (1993). Procedural fairness and self-esteem. *European Journal of Social Psychology*, *23*, 313–325.

Kramer, R. M., & Jost, J. T. (2002). Close encounters of the suspicious kind: Outgroup paranoia in hierarchical trust dilemmas. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 173–189). New York: Psychology Press.

Kramer, R. M., & Messick, D. M. (1998). Getting by with a little help from our enemies: Collective paranoia and its role in intergroup relations. In C. Sedikides & J. Scholpler (Eds.), *Intergroup cognition and intergroup behavior* (pp. 233–255). Mahwah, NJ: Erlbaum.

Levine, J. M., & Moreland, R. L. (1994). Group socialization: Theory and research. *European Review of Social Psychology*, *5*, 305–36.

Levine, J. M., & Moreland, R. L. (1995). Group processes. In A. Tesser (Ed.), *Advanced social psychology* (pp. 419–465). New York: McGraw-Hill.

Liden, R. C., Sparrowe, R. T., & Wayne, S. J. (1997). Leader–member exchange theory: The past and potential for the future. *Research in Personnel and Human Resources Management*, *15*, 47–119.

Liebrand, W., Messick, D., & Wilke, H. (Eds.). (1992). *A social psychological approach to social dilemmas*. New York: Pergamon Press.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum Press.

Lindholm, C. (1990). *Charisma*. Oxford, UK: Blackwell.

Lippitt, R., & White, R. (1943). The "social climate" of children's groups. In R. G. Barker, J. Kounin, & H. Wright (Eds.), *Child behavior and development* (pp. 485–508). New York: McGraw-Hill.

Locke, J. (1988). *Two treatises of government*. Cambridge, UK: Cambridge University Press. (Original work published 1690)

Lord, R. G., & Brown, D. J. (2004). *Leadership processes and follower self-identity*. Mahwah, NJ: Erlbaum.

Lord, R. G., Brown, D. J., & Harvey, J. L. (2001). System constraints on leadership perceptions, behavior and influence: An example of connectionist level processes. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 283–310). Oxford, UK: Blackwell.

Lord, R. G., Brown, D. J., Harvey, J. L., & Hall, R. J. (2001). Contex-

tual constraints on prototype generation and their multilevel consequences for leadership perceptions. *Leadership Quarterly*, *12*, 311–338.

Lord, R. G., Foti, R. J., & DeVader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. *Organizational Behavior and Human Performance*, *34*, 343–378.

Lord, R. G., Foti, R. J., & Phillips, J. S. (1982). A theory of leadership categorization. In J. G. Hunt, U. Sekaran, & C. Schriesheim (Eds.), *Leadership: Beyond establishment views* (pp. 104–121). Carbondale: Southern Illinois University Press.

Lord, R., & Hall, R. (2003). Identity, leadership categorization, and leadership schema. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 48–64). London: Sage.

Lord, R. G., & Maher, K. J. (1991). *Leadership and information processing: Linking perceptions and performance*. Boston: Unwin Hyman.

Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformational and transactional leadership: A meta-analytic review. *Leadership Quarterly*, 7, 385–425.

Martell, R. F., Parker, C., Emrich, C. G., & Swerdlin Crawford, M. (1998). Sex stereotyping in the executive suite: "Much ado about something." *Journal of Social Behavior and Personality*, *13*, 127–138.

McGrath, J. E. (1997). Small group research, that once and future field: An interpretation of the past with an eye to the future. *Group Dynamics: Theory, Research, and Practice*, *1*, 7–27.

Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. (1985). The romance of leadership. *Administrative Science Quarterly*, *30*, 78–102.

Meindl, J. R., & Lerner, M. (1983). The heroic motive: Some experimental demonstrations. *Journal of Experimental Social Psychology*, *19*, 1–20.

Merei, F. (1949). Group leadership and institutionalization. *Human Relations*, *2*, 23–39.

Messick, D. M. (2005). On the psychological exchange between leaders and followers. In D. M. Messick & R. M. Kramer (Eds.), *The psychology of leadership: New perspectives and research* (pp. 81–96). Mahwah, NJ: Erlbaum.

Messick, D. M., & Kramer, R. M. (Eds.). (2005). *The psychology of leadership: New perspectives and research*. Mahwah, NJ: Erlbaum.

Misumi, J., & Peterson, M. F. (1985). The performance–maintenance (P-M) theory of leadership: Review of a Japanese research program. *Administrative Science Quarterly*, *30*, 198–223.

Moreland, R. L., Hogg, M. A., & Hains, S. C. (1994). Back to the future: Social psychological research on groups. *Journal of Experimental Social Psychology*, *30*, 527–555.

Moscovici, S. (1976). *Social influence and social change*. London: Academic Press.

Mowday, R. T., & Sutton, R. I. (1993). Organizational behavior: Linking individuals and groups to organizational contexts. *Annual Review of Psychology*, *44*, 195–229.

Northhouse, P. (2004). *Leadership: Theory and practice* (3rd ed.). Thousand Oaks, CA: Sage.

O'Connor, J., Mumford, M. D., Clifton, T. C., Gessner, T. L., & Connelly, M. S. (1995). Charismatic leaders and destructiveness: A historiometric study. *Leadership Quarterly*, *6*, 529–558.

Peters, L. H., Hartke, D. D., & Pohlmann, J. T. (1985). Fiedler's contingency theory of leadership: An application of the meta-analytic procedure of Schmidt and Hunter. *Psychological Bulletin*, *97*, 274–285.

Platow, M. J., Hoar, S., Reid, S. A., Harley, K., & Morrison, D. (1997). Endorsement of distributively fair and unfair leaders in interpersonal and intergroup situations. *European Journal of Social Psychology*, *27*, 465–94.

Platow, M. J., Reid, S. A., & Andrew, S. (1998). Leadership endorsement: The role of distributive and procedural behavior in interpersonal and intergroup contexts. *Group Processes and Intergroup Relations*, *1*, 35–47.

Platow, M. J., & van Knippenberg, D. (2001). A social identity analysis of leadership endorsement: The effects of leader ingroup prototypicality and distributive intergroup fairness. *Personality and Social Psychology Bulletin*, *27*, 1508–1519.

Raven, B. H. (1993). The bases of power: Origins and recent developments. *Journal of Social Issues*, *49*, 227–251.

Reicher, S. D., & Hopkins, N. (1996). Self-category constructions in political rhetoric: An analysis of Thatcher's and Kinnock's speeches concerning the British miners' strike (1984–5). *European Journal of Social Psychology*, *26*, 353–371.

Reicher, S. D., & Hopkins, N. (2001). *Self and nation*. London: Sage.

Reicher, S., & Hopkins, N. (2003). On the science of the art of leadership. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 197–209). London: Sage.

Reid, S. A., & Ng, S. H. (2000). Conversation as a resource for influence: Evidence for prototypical arguments and social identification processes. *European Journal of Social Psychology*, *30*, 83–100.

Ridgeway, C. L. (2001). Social status and group structure. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 352–375). Oxford, UK: Blackwell.

Ridgeway, C. L. (2003). Status characteristics and leadership. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 65–78). London: Sage.

Riggio, R. E., & Carney, D. R. (2003). *Social Skills Inventory manual* (2nd ed.). Redwood City, CA: MindGarden.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 174–220). New York: Academic Press.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, *74*, 629–645.

Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, *75*, 1004–1010.

Sanna, L. J., & Parks, C. D. (1997). Group research trends in social and organizational psychology: Whatever happened to intragroup research? *Psychological Science*, *8*, 261–267.

Scandura, T. A. (1999). Rethinking leader–member exchange: An organizational justice perspective. *Leadership Quarterly*, *10*, 25–40.

Schriesheim, C. A., Castro, S. L., & Cogliser, C. C. (1999). Leader–member exchange (LMX) research: A comprehensive review of theory, measurement, and data-analytic practices. *Leadership Quarterly*, *10*, 63–113.

Schriesheim, C. A., & Neider, L. L. (1996). Path-goal leadership theory: The long and winding road. *Leadership Quarterly*, 7, 317–321.

Schriesheim, C. A., Tepper, B. J., & Tetrault, L. A. (1994). Least preferred co-worker score, situational control, and leadership effectiveness: A meta-analysis of contingency model performance predictions. *Journal of Applied Psychology*, *79*, 561–573.

Shamir, B., House, R., & Arthur, M. (1993). The motivational effects of charismatic leadership: A self-concept based theory. *Organization Science*, *4*(3), 1–17.

Shartle, C. L. (1951). Studies in naval leadership. In H. Guetzkow (Ed.), *Groups, leadership, and men* (pp. 119–133). Pittsburgh, PA: Carnegie Press.

Shaw, M. E. (1981). *Group dynamics: The psychology of small group behavior* (2nd ed.). New York: McGraw-Hill.

Sherif, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation*. Boston: Houghton-Mifflin

Sherif, M., Harvey, O. J., White, B. J., Hood, W., & Sherif, C. (1961). *Intergroup conflict and cooperation: The robbers' cave experiment*. Norman: University of Oklahoma, Institute of Intergroup Relations.

Shore, T. H. (1992). Subtle gender bias in the assessment of managerial potential. *Sex Roles*, *27*, 499–515.

Simonton, D. K. (1980). Land battles, generals and armies: Individual and situational determinants of victory and casualties. *Journal of Personality and Social Psychology*, *38*, 110–119.

Simonton, D. K. (1994). *Greatness: Who makes history and why.* New York: Guilford Press.

Slater, P. E. (1955). Role differentiation in small groups. *American Sociological Review*, *20*, 300–310.

Smith, P. B., Misumi, J., Tayeb, M., Peterson, M., & Bond, M. (1989). On the generality of leadership style measures across cultures. *Journal of Occupational Psychology*, *62*, 97–109.

Sorrentino, R. M., & Field, N. (1986). Emergent leadership over time: The functional value of positive motivation. *Journal of Personality and Social Psychology*, *50*, 1091–1099.

Sparrowe, R. T., & Liden, R. C. (1997). Process and structure in leader-member exchange. *Academy of Management Review*, *22*, 522–552.

Steele, C. M., & Aronson, J. (1995). Stereotype vulnerability and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.

Steiner, I. D. (1974). Whatever happened to the group in social psychology? *Journal of Experimental Social Psychology*, *10*, 94–108.

Stogdill, R. M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology*, *25*, 35–71.

Stogdill, R. (1974). *Handbook of leadership.* New York: Free Press.

Strube, M. J., & Garcia, J. E. (1981). A meta-analytic investigation of Fiedler's contingency model of leadership effectiveness. *Psychological Bulletin*, *90*, 307–321.

Tajfel, H. (Ed.). (1984). *The social dimension: European developments in social psychology*. Cambridge, UK: Cambridge University Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago: Nelson-Hall.

Taylor, S. E. (1998). The social being in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 58–95). New York: McGraw-Hill.

Tindale, R. S., & Anderson, E. M. (1998). Small group research and applied social psychology: An introduction. In R. S. Tindale, L. Heath, J. Edwards, E. J. Posavac, F. B. Bryant, Y. Suarez-Balcazar, et al. (Eds.), *Social psychological applications to social issues: Theory and research on small groups* (Vol. 4, pp. 1–8). New York: Plenum Press.

Tolstoy, L. (1869). *War and peace*. Harmondsworth, UK: Penguin.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, *63*, 384–99.

Turner, J. C. (1991). *Social influence*. Buckingham, UK: Open University Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford, UK: Blackwell.

Tyler, T. R. (1997). The psychology of legitimacy: A relational perspective on voluntary deference to authorities. *Personality and Social Psychology Review*, *1*, 323–345.

Tyler, T. R. (2003). Justice, identity, and leadership. In D. van Knippenberg & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 94–108). London: Sage.

Tyler, T. R. (2005). Process-based leadership: How do leaders lead? In D. M. Messick & R. M. Kramer (Eds.), *The psychology of leadership: New perspectives and research* (pp. 163–189). Mahwah, NJ: Erlbaum.

Tyler, T. R., & Blader, S. L. (2000). *Cooperation in groups*. Philadelphia: Psychology Press.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–191). New York: Academic Press.

van Knippenberg, B., & van Knippenberg, D. (2005). Leader self-sacrifice and leadership effectiveness: The moderating role of leader prototypicality. *Journal of Applied Psychology*, *90*, 25–37.

van Knippenberg, D., & Hogg, M. A. (Eds.). (2003a). *Leadership and power: Identity processes in groups and organizations.* London: Sage.

van Knippenberg, D., & Hogg, M. A. (2003b). A social identity model of leadership in organizations. In R. M. Kramer & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 25, pp. 243–295). Greenwich, CT: JAI Press.

van Knippenberg, D., van Knippenberg, B., De Cremer, D., & Hogg, M. A. (2004). Leadership, self, and identity: A review and research agenda. *Leadership Quarterly*, *15*, 825–856.

van Vugt, M., & De Cremer, D. (1999). Leadership in social dilemmas: The effects of group identification on collective actions to provide public goods. *Journal of Personality and Social Psychology*, *76*, 587–599.

Vermunt, R., van Knippenberg, D., van Knippenberg, B., & Blaauw, E. (2001). Self-esteem and outcome fairness: Differential importance of procedural and outcome considerations. *Journal of Applied Psychology*, *86*, 621–628.

Vroom, V. H., & Jago, A. G. (1988). *The new leadership*. Englewood Cliffs, NJ: Prentice Hall.

Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision-making*. Pittsburgh, PA: University of Pittsburgh Press.

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity theory and research*. Boston: Allyn & Bacon.

Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, *63*, 116–132.

Yukl, G. (2002). *Leadership in organizations* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Dynamical Social Psychology

Finding Order in the Flow of Human Experience

ROBIN R. VALLACHER
ANDRZEJ NOWAK

Intuition is a double-edged sword for a social psychologist. The advantages are obvious. Particularly if the psychologist is a keen observer of the human condition and is sensitive to the nuances of his or her own experience, intuition can provide a head start on theory construction that is simply not available to a scientist attempting to identify the principles of, say, photosynthesis, radioactivity, or quasar formation. The psychologist's privileged access to his or her subject matter—the experience of humans—provides an abundance of raw data for identifying lawful regularities concerning emotion, motivation, social influence, and all manner of interpersonal behavior. Therein lies the other edge of the sword. There are no limits to the amount and diversity of insights forthcoming from introspection, observation, and empathy. The insights emanating on a continual basis from our private and social experiences can generate a dizzying amount of relevant data, expose an enormous range of important phenomena, and promote a bewildering array of explanations with unclear points of confluence and conflict. It is not surprising, then, that social psychology has developed into a discipline that is rich in insight but devoid of theoretical consensus, let alone explanatory coherence.

One could argue that the fragmented nature of contemporary social psychology simply mirrors the complexity of human experience. The intricacies of information processing, after all, seem to belong to a different theoretical realm than the passion and dependency associated with close relationships. With a topical landscape that runs the gamut from vascular processes to social movements, perhaps it is entirely reasonable that the explanatory landscape of social psychology should be correspondingly diverse. Indeed, this idea itself is represented in the intuitions of laypeople and psychologists alike.

Recent years, however, have witnessed the emergence of a new perspective on the relationship between surface and explanatory complexity in many areas of science (cf. Haken, 1978; Holland, 1995; Johnson, 2001; Schuster, 1984; Strogatz, 2003; Weisbuch, 1992; Wolfram, 2002). This perspective trades on the idea that any phenomenon can be conceptualized as a system of interacting elements, the dynamics of which give rise to new properties and processes at an aggregate level. In analyzing phenomena as nonlinear dynamical systems, the task of the scientist is to identify the rules of interaction among system elements and to investigate how these rules promote the emergence of macrolevel behavior. In fields as distinct as population ecology and organic chemistry, the dynamical systems perspective has revealed that highly complex behavior at the macrolevel can emerge from very simple rules of influence among the system's elements. The rules linking the micro and macro features of a system, moreover, have been found to be remarkably similar for otherwise highly distinct phenomena.

Our aim in this chapter is to outline the dynamical perspective and demonstrate its relevance to social psychology. Although in its infancy, this approach has already proven useful in highlighting many distinct phenomena

and providing the discipline with a new set of research tools. Social psychology covers a lot of ground, however, and there are many topics and issues that have yet to be addressed from an explicitly dynamical perspective. In addition to noting the accomplishments to date, then, we suggest how various phenomena of current interest might be reframed in dynamical terms and investigated with methods adapted from the study of nonlinear systems in other areas of science. In a concluding section, we describe dynamical minimalism, an approach to theory construction and research that mirrors the dynamical properties of human experience. Because dynamical principles characterize complex systems generally, the dynamical approach to investigating personal and interpersonal experience holds potential for bridging social psychology with other areas of science.

THE DYNAMICAL PERSPECTIVE

Over the past decade, theorists and researchers in the social sciences have come to recognize the relevance of the dynamical approach to their respective areas of concern. This recognition has fueled ambitious attempts to adapt the concepts and methods associated with nonlinear dynamical systems to various topics in social psychology, including attitudes (e.g., Eiser, 1994; Latané & Nowak, 1994; Shultz & Lepper, 1996; Simon & Holyoak, 2002), social judgment (e.g., Kunda & Thagard, 1996; Queller, 2002; Read & Miller, 1998; Smith, 1996; Vallacher, Nowak, & Kaufman, 1994), decision making (Kaplowitz & Fink, 1992; Richards, 1990; Townsend & Busemeyer, 1995), self-regulation (e.g., Carver & Scheier, 1999; Hsee, Abelson, & Salovey, 1991; Vallacher & Nowak, 1999), self-concept (e.g., Nowak, Vallacher, Tesser, & Borkowski, 2000; Vallacher, Nowak, Froehlich, & Rockloff, 2002), action (e.g., Newtson, 1994; Thelen & Smith, 1994; Turvey & Carello, 1995; Vallacher, Nowak, Markus, & Strauss, 1998), social interaction (e.g., Baron, Amazeen, & Beek, 1994; Buder, 1991; Newtson, 1994; Nowak & Vallacher, 1998b), social influence (e.g., Nowak, Szamrej, & Latané, 1990; Vallacher, Nowak, & Miller, 2002), group dynamics (e.g., Arrow, McGrath, & Berdahl, 2000), close relations (e.g., Gottman, Murray, Swanson, & Tyson, 2004; Kenrick, Li, & Butner, 2003; Tesser & Achee, 1994), organizational behavior (e.g., Guastello, 1995), and social systems (e.g., Goldstein, 1996; Nowak & Vallacher, 2001). This general approach has also been used to investigate and model personality development (e.g., Fischer & Bidell, 1997; Lewis, 1997; Nowak, Vallacher, & Zochowski, 2002) and personality dynamics (e.g., Carver & Scheier, 2002; Cervone, 2004; Mischel & Shoda, 1995; Read & Miller, 2002; Shoda, LeeTiernan, & Mischel, 2002). The various adaptations differ in crucial respects, but they share the assumption that personal and interpersonal phenomena can be profitably viewed as sets of interconnected elements that evolve due to the mutual influences among the elements (cf. Barton, 1994; Weidlich, 1991). Research within the dynamical perspective describes the connections among a system's elements and investigates the system-level properties and behaviors that these connections promote.

Below we highlight the primary features of nonlinear dynamical systems that are especially relevant for the subject matter of social psychology. We focus first on the tendency for systems to display internally generated dynamics and on the temporal patterns to which such dynamics conform. Building on this foundation, we develop the implications of nonlinear relations for the nature of complexity and the emergence of global properties and higher-order processes in human functioning. We develop these ideas in the context of specific personal or interpersonal processes of interest to social psychology. Our aim is not to be exhaustive but, rather, to illustrate the heuristic and integrative potential of the dynamical perspective.[1]

Intrinsic Dynamics

Dynamical systems are not new to science. This perspective, in fact, has been employed to investigate phenomena since the rise of modern science. Perhaps the most fundamental characteristic of a dynamical system is that it changes and evolves in time. Consider the proverbial falling apple that is said to have knocked insight regarding gravitation into Isaac Newton. The apple begins moving, continues moving for some time, eventually hits the ground (after bonking the hapless Newton), and finally stops moving. The state of the apple at any moment in time (i.e., its position, velocity, and angular momentum) determines its state in the next moment in time. The apple's motion can be described and modeled using differential or difference equations. The resultant model is deterministic in the sense that the future of the apple can be fully predicted from full knowledge of the apple's current state.

Clearly humans are not apples, nor should reducing complex human experience to differential equations be the goal of social psychology. Nonetheless, the experience gained in the investigation of dynamics in other systems in nature may provide important insights concerning the nature of processes occurring in people and social groups. People's thoughts, emotions, and actions can clearly evolve and change in the absence of external influence. Such patterns of "intrinsic dynamics," however, are rarely the focus of social psychological theory and research (cf. McGrath & Kelley, 1986). Research instead largely concentrates on trying to predict the values of some outcome variables (operationalized as dependent measures) from the knowledge of other variables (independent variables). External causation clearly characterizes social processes, of course, and this approach has generated enormous insight into personal and interpersonal phenomena. External forces, however, do not act on an empty or passive system. Rather, such factors interact with the intrinsic dynamics associated with the process in question. The concepts, principles, and methods developed within nonlinear dynamical systems thus may enhance our understanding of social psychological phenomena and how otherwise distinct phenomena are related in terms of common properties.

Personal Dynamics

The importance of intrinsic dynamics is apparent at different levels of social reality, from personal processes to interpersonal processes and even societal phenomena. At the individual level, the spontaneous activation and turnover of cognitive and affective elements in the stream of thought (James, 1890) may be more indicative of a person's mental makeup than the summary aspects of thought (e.g., overall attitude, final decision, and global self-evaluation) that are more commonly the focus of investigation (cf. Vallacher & Nowak, 1994b). Attention to the internal workings of the cognitive–affective system is apparent in several contemporary lines of theory and research. Theory and research on thought-induced attitude polarization, for example, has shown that when people simply think about an attitude object (e.g., another person) in the absence of external influence or new information, their evaluation of the object tends to be become more extreme over time (e.g., Tesser, 1978).

Research on the time course of social judgment has shown that internally generated thoughts and feelings about someone can reflect rich and elaborate patterns of change and that these patterns convey important information about the person's judgment. A seemingly neutral judgment of someone, for instance, can have quite different meanings and implications, depending on the intrinsic dynamics of the judgment process (Vallacher et al, 1994). If neutrality reflects little variation in evaluation occurring on a relatively slow time scale, the summary judgment might indicate detachment or a true neutral sentiment. But if neutrality reflects oscillation between highly positive and highly negative judgments, and this variation occurs on a rapid time scale, the summary judgment is likely to reflect heightened involvement and ambivalence rather than neutrality per se.

Similarly, knowing only a person's level of self-esteem may provide an unclear, potentially misleading indication of how the person processes self-relevant information, responds to social feedback, interacts with others, or forms social relationships (e.g., Baumgardner, 1990; Campbell et al., 1996; Kernis, 1993; Nowak et al., 2000; Pelham, 1991; Showers, 1992; Vallacher, 1978, 1980; Vallacher & Nowak, 2000). Insight into such issues is enhanced by tracking the moment-to-moment changes in self-evaluation in the stream of self-reflective thought (Vallacher et al., 2002). A person may express a highly positive assessment of him- or herself on a self-report measure, but the manifestation of the person's high self-esteem depends on the dynamic properties of his or her self-reflection. If the person's self-evaluative thought is relatively quiescent, with modest changes in self-evaluation occurring on a slow time scale, his or her level of self-regard is likely to be associated with high self-concept certainty, the maintenance of particular self-views despite contradictory incoming information, and a nondefensive orientation toward other people. But if the person's high self-esteem masks underlying volatility in self-reflection—the spontaneous expression of highly pos-

itive and highly negative self-relevant thoughts on a relatively fast time scale—he or she is likely to display lower self-concept certainty and heightened vulnerability to social feedback that challenges his or her positive self-image.

Contemporary models of personality also incorporate the potential for intrinsic dynamics into their conceptualizations and research strategies (e.g., Carver & Scheier, 1999; Cervone, 2004; Lewis, 1997; Mischel & Shoda, 1995; Read & Miller, 2002). These models differ in a number of important respects, of course, but they share an appreciation for the role of internal mechanisms in giving meaning to, and modulating the effects of, environmental cues and demands. In this perspective, the stability of personality does not refer to stable traits or to the control provided by a stable environment but, rather, to stable patterns of thought, emotion, and action in the context of goals, opportunities, pressures, and other situational parameters. Such models go beyond the recognition of personality by situation interactions to focus on the structures within a person that interpret, synthesize, and utilize incoming information.

Interpersonal Dynamics

Interpersonal behavior involves the temporal coordination of behavior at different levels, from motor movements and utterances to high-level action categories reflecting momentary goals or long-range plans. Actions as basic as talking or passing in a hallway require coordination in time (e.g., turn taking) to prevent individuals from stumbling over one another. As discussed previously, a person can be viewed as a separate system capable of displaying rich dynamics. Social coordination, then, involves the synchronization of partners' respective dynamics to produce a higher-order system with its own dynamic properties.

Several lines of research have explored the coordination of dynamics in social interaction with respect to speech (e.g., Condon & Ogston, 1967; Dittman & Llewellyn, 1969) and motor movement (e.g., Beek & Hopkins, 1992; Kelso, 1995; Newtson, 1994; Turvey, 1990). Two individuals, for example, might simply be asked to swing their legs. One person swings his or her legs in time to a metronome and the other person tries to match those movements. This simple paradigm reveals interesting coordination phenomena. First, synchronization may be in phase, with people swinging their legs in unison, or in antiphase, with people swinging their legs with the same frequency but in the opposite direction. Second, hysteresis is commonly observed. When participants are instructed to synchronize out of phase, they are able to do so only up to a certain frequency of movement, at which they switch to in-phase synchronization. When the tempo decreases, at some value they are able to coordinate out of phase again, but this tempo is significantly lower than the point at which they originally started to synchronize in phase. The appearance of hysteresis shows that movement coordination can be analyzed as a nonlinear dynamical system (Kelso, 1995). Yet more complex modes of coordination have been cap-

tured in this line of research (cf. Baron et al., 1994; Turvey, 1990).

Social relations go beyond the coordination of speech and motor movements to include the temporal coordination of higher-order actions and internal states (moods, judgments, etc.). The quality of a social relationship is reflected in the ability of partners to synchronize in this fashion (e.g., Baron et al., 1994; McGrath & Kelly, 1986; Nowak et al., 2002; Tickle-Degnen & Rosenthal, 1987). In everyday parlance, people who like one another are said to "be in synch" or "on the same wavelength." Dynamically speaking, the ebb and flow of sentiment, information exchange, and action may be more informative about the nature of a relationship than are the average sentiment, the amount of information exchanged, or the summary action tendencies.

In a close relationship, for example, the partners' mutual affect may alternate between intense passion and equally intense anger (e.g., Gottman, Swanson, & Swanson, 2002). If one assessed partners' feelings at a point in time dictated by measurement convenience or the needs of the experimenter, one would conclude that they experienced either positive or negative feelings, depending on which phase the partners were experiencing at that time. If, instead, one were to collapse partners' expressed feelings over time to obtain a summary measure (with temporal variance treated as noise), one would be left with the curious conclusion that the predominant affective tone was mutual neutrality—a state that is never experienced. What is important is the pattern of feelings over time, which may have a life of its own, independent of factors outside the dyadic system that are capable of promoting some feelings rather than others.

Societal Dynamics

At a societal level, tracking the temporal trajectory associated with the emergence of norms and public opinion may provide greater insight into the society's future makeup and likely response to external threat than simply knowing what the norms and public sentiments are (cf. Nowak et al., 1990). When norms and opinions develop slowly and incrementally, for example, society is likely to display resistance to external threats or even to new information that might promote better economic conditions. Societal change in political and economic ideology, however, can also occur in a rapid, nonlinear manner (e.g., Nowak & Vallacher, 2001; Nowak, Vallacher, Kus, & Urbaniak, 2005), with a trajectory that resembles phase transitions in physical systems (Lewenstein, Nowak, & Latané, 1993). Such nonlinear societal transitions can make society vulnerable to subsequent rebounds of the earlier ideologies and highly responsive to threats and new information, and they can promote a period of sustained oscillation between conflicting worldviews (Nowak & Vallacher, 2001).

Attractor Dynamics

The proclivity for intrinsic dynamics cannot be denied, but neither can the tendency for psychological systems to demonstrate stability and resistance to change. An enormous amount of diverse information relevant to self-understanding, social judgment, goal-directed action, and interpersonal relations is encountered on a daily basis, yet people manage to maintain relatively stable platforms for thought and action in each of these domains. Clearly, then, the flow of thoughts, feelings, and behavior is not random but, rather, converges on specific states or patterns of change between states. In reaction to negative life experiences, for example, a person is likely to experience sadness and perhaps even depression, but over time his or her self-regulatory mechanisms are likely to reinstate positive moods and a sense of personal control (e.g., Johnson & Nowak, 2002). Similarly, a person with high self-esteem may experience negative self-relevant thoughts, but these are likely to be unstable, giving way over time to a flow of self-evaluative thinking that converges on a positive state (e.g., Vallacher, Nowak, Froehlich, & Rockloff, 2002).

The constraints on psychological process can be couched in terms of *attractor dynamics*. An attractor is a state or a reliable pattern of changes (e.g., oscillation between two states) toward which a dynamical system evolves over time, and to which the system returns after it has been perturbed. In a system governed by attractor dynamics, a relatively wide range of starting points (initial states) will eventually converge on a much smaller set of states or on a pattern of change between states. In effect, an attractor "attracts" the system's dynamics, so that despite differences at the outset in one's thoughts, feelings, or behaviors, the process unfolds in the direction of the attractor. Attempting to move the system out of its attractor, moreover, promotes forces that reinstate the system at its attractor. Across different areas of science, research has identified three types of attractors (cf. Eckmann & Ruelle, 1985; Nowak & Lewenstein, 1994; Schuster, 1984): fixed-point attractors, periodic (including multiperiodic) attractors, and deterministic chaos (intrinsic dynamics characterized by very irregular, seemingly random temporal evolution).

Fixed-Point Attractors

A *fixed-point attractor* describes the case in which the state of the system converges to a stable value. In psychological systems, this tendency is similar to the notion of equilibrium or homeostasis (cf. Cannon, 1932; Miller, 1944). In some cases, the attractor may correspond a desired end state or goal (Vallacher & Nowak, 1997) that functions as a standard of self-regulation (Carver & Scheier, 1999). In the face of obstacles to goal attainment, a person's thought and behavior are configured (and reconfigured if necessary) in order to overcome the obstacles and ensure achievement of the goal or maintenance of the regulatory standard. Thus, people are described as motivated to bring about and maintain various psychological states, whether cognitive (e.g., a belief), affective (e.g., a judgment or an attitude), or behavioral (e.g., an action tendency or desire) in nature.

Attractors are not limited to goals, however, nor do they necessarily refer to intentions or desired states. A

person may display patterns of behavior that repetitively converge on hostility across his or her social relations, for example, even when he or she attempts to avoid behaving in this manner. In an intergroup context, meanwhile, warring factions embroiled in conflict may exhibit conciliatory gestures when prompted to do so but revert to a pattern of antagonistic thought and behavior if left to their own devices. In like manner, a person with low self-esteem may initially embrace positive feedback from someone, but over time he or she is likely to discount such feedback, displaying a pattern of self-reflective thought that converges on a negative state (Swann, Hixon, Stein-Seroussi, & Gilbert, 1990). A system governed by fixed-point attractor dynamics, in other words, will consistently evolve to a particular state, whether or not this state is hedonically pleasant, and will return to this state even when perturbed by outside influences that might promote a more pleasant state.

A system may have multiple stable states, each corresponding to an attractor for the system's behavior. The particular attractor that is reached depends on the initial states or starting values of the system's evolution. The set of initial states leading to each attractor represents the *basin of attraction* for that attractor. In a system with multiple fixed-point attractors, then, the system can display correspondingly different equilibrium tendencies, each associated with a distinct basin of attraction. Within a given basin, even quite different initial states will promote a trajectory that eventually converges on the same stable value. But by the same token, even a slight deviation in initial conditions has the potential to promote a dramatic change in the system's trajectory if the deviation represents a state that falls just outside the original basin of attraction and within a basin for a different attractor. In a high-threat situation, for example, there are two dominant responses, corresponding to fight versus flight. Minimal differences in the circumstances associated with the threat will thus lead to dramatically different behaviors, with virtually no option for a response that integrates the two tendencies.

The essence of the attractor concept and its relevance for personal and social processes can be illustrated with a simple metaphor. Figure 32.1 portrays a ball on a hilly landscape. The ball represents the current state of the system and the two valleys (A and B) represent different fixed-point attractors for the system. The ball will roll down the hill and come to rest at the bottom of a valley. Each attractor has its own basin of attraction—that is, a set of states that will evolve toward the attractor. The ba-

sin of attraction for attractor A is somewhat wider than the basin for attractor B. This means that a wider variety of states will evolve toward attractor A than toward attractor B. Attractors can also vary in their respective strength, which is depicted as the relative depth of the two valleys. Attractor B, then, is stronger than attractor A. This means that when a system is at attractor B, it is more difficult for it to be dislodged by external influence.

The potential for multiple fixed-point attractors captures the idea that people may have different (sometimes conflicting) goals, self-views, and behavior patterns. A person may have more than one self-regulatory standard, for example, each providing for behavioral direction and self-control under different conditions. His or her actions may be in service of an achievement standard under a range of conditions that promote this tendency but in service of affiliation standards under a different set of conditions. In similar fashion, a person may have multiple self-views (e.g., Markus & Nurius, 1986), each representing a stable way of thinking about him- or herself that comes to the fore when a specific set of self-relevant information is primed, made salient by virtue of context or role expectations, or otherwise brought to consciousness. Apparent inconsistency or conflict in personality, too, can be viewed as the existence of multiple attractors associated with different basins of attraction for thought and action (cf. Nowak et al., 2002). One set of conditions and initial states might promote a trajectory that evolves toward dominance and competition, but another might promote instead warmth and compassion.

Because an attractor's strength may be independent of the size of its basin of attraction, one can envision different combinations of these properties and consider their implications for psychological processes.[2] Attractors A and B in Figure 32.1 present contrasting combinations. Attractor A is relatively weak but has a wide basin of attraction. In this case, a small force may be sufficient to change the state of the system (i.e., move the ball up the gradual slope), but even if these changes are relatively large, the system will still have a tendency to return to the attractor (i.e., it will roll back into the valley). Attractor B is stronger but has a relatively narrow basin of attraction. In this case, considerable influence is necessary to have even a slight impact on the system (i.e., move the ball up the steep slope), but once such an effect is achieved, the system will lose its ability to return to the attractor (i.e., it will escape the valley).

To illustrate this difference, consider two people, A and B, who express equally flattering self-appraisals but whose attractors for positive self-evaluation differ in their respective strength and basin of attraction. Person A has a relatively weak attractor but a relatively wide basin of attraction (corresponding to attractor A in Figure 32.1). This person may entertain a broad range of evaluatively discrepant thoughts (e.g., reminders of past deeds, negative social feedback), but over time his or her trajectory of self-evaluation is nonetheless likely to converge on the positive attractor. Person B has a stronger attractor for positive self-evaluation but a narrower basin of attraction (corresponding to attractor B in Figure 32.1). Unlike per-
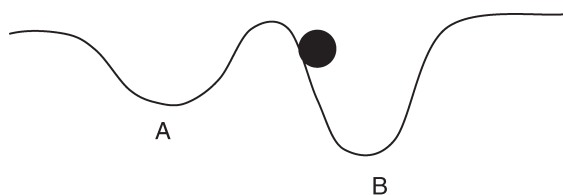
FIGURE 32.1. A dynamical system with two attractors (A and B).

son A, he or she may actively resist even mildly negative thoughts and immediately discount unflattering self-relevant information. If such thoughts and elements of information fail to be rejected, however, they may be sufficient to generate self-doubt and, if a potential negative attractor exists for the person, a trajectory of self-evaluation that converges on a considerably less flattering self-appraisal.

Differences in the properties of fixed-point attractors can be envisioned as well at an interpersonal level of functioning. Imagine a couple that has a strong attractor associated with positive feelings but also a weak attractor associated with negative feelings. The couple may consistently evolve toward positive feelings if the partners begin an interaction within a certain range of affective states (e.g., neutral to very positive) but just as consistently end up feeling negative about one another if they begin an interaction within a different range of affective states (e.g., mildly to highly negative). If the couple has a wider basin of attraction for positive feelings than for negative feelings, a broader range of initial states is likely to produce a communication trajectory that results in an exchange of warm sentiments. If the couple routinely starts out with negative feelings, though, the relatively narrow negative attractor may nonetheless dictate the predominant trajectory for feelings expressed in the couple's interactions. On the other hand, if the couple has a wider basin for negative feelings, anything short of a highly positive initial state could dissolve into a negatively toned exchange.

Latent Attractors

When a system is at one of its attractors, other attractors for the system's behavior may not be visible to observers, perhaps not even to the actors themselves. Indeed, the existence of these potential states of the system might not even be suspected. Such *latent attractors* are nonetheless important in the long run because they determine which states are possible for the system when conditions change. The concept of latent attractor thus goes beyond the traditional notion of equilibrium by specifying possibilities for a system that have yet to be observed or experienced. Important changes in a system might not be reflected in the system's observable state but, rather, in the creation or destruction of a latent attractor corresponding to a potential state that is currently invisible to all concerned.

The possibility of latent attractors has recently been explored in the context of social relations characterized by seemingly intractable conflict (Coleman, Vallacher, Nowak, & Bui-Wrzosinska, in press; Nowak, Vallacher, Bui-Wrzosinska, & Coleman, in press). In intergroup relations, for example, factors such as objectification, dehumanization, and stereotyping of outgroup members have been shown to be preconditions for the development of intractable conflict (Coleman, 2003; Deutsch, 1973), but their immediate impact may not be apparent. Rather, these factors may gradually create a latent attractor to which the system can abruptly switch in response to seemingly minor provocations. By the same token,

though, seemingly fruitless efforts at conflict resolution may have the effect of creating a latent positive attractor for intergroup relations, thereby establishing a potential relationship to which the groups can switch if other conditions permit. The existence of a latent positive attractor can promote a rapid deescalation of conflict, even among groups with a long history of seemingly intractable conflict.[3]

The concept of latent attractor resonates well with the concept of implicit attitude (cf. Greenwald & Banaji, 1995; Greenwald et al., 2002; Greenwald, McGhee, & Schwarz, 1998). In the course of socialization (through peer and parental influence, mass media, idiosyncratic experiences, etc.), some thoughts, beliefs, and attitudes may become integrated in a relatively coherent manner and thus constitute an attractor for thoughts and feelings concerning a person or group (or oneself). Such an attractor may become inaccessible, however, by virtue of control mechanisms that reject or suppress the cognitive and affective elements comprising it. In contemporary society, for example, racist attitudes are widely condemned and strongly rejected, yet they are potentially accessible in a prepackaged manner as an integrated belief system (cf. Devine, 1989).

Because access to this set of thoughts may be constantly blocked by self-regulation (e.g., controlled cognitive processes), this attractor for mental dynamics is unlikely to be experienced. Under conditions that disrupt self-regulation, however, the latent attractor may become manifest and shape the person's trajectory of thoughts, feelings, and actions. This perspective is consistent with research showing that such factors as stress, cognitive load, deindividuation, and heightened arousal may facilitate the expression of attitudes and emotions that are normally suppressed (cf. Macrae, Bodenhausen, Milne, & Jetten, 1994; Sherman, Stroessner, Loftus, & Deguzman, 1997; Wegner, 1994; Zimbardo, 1970).

Repellors

Social psychological dynamics can be described in terms of not only seeking and maintaining specific states (i.e., fixed-point attractors) but also avoiding or escaping various states. In classic and contemporary psychological theories, avoidance tendencies are accorded equal theoretical status to approach tendencies (e.g., Higgins, 2000; Lewin, 1936). Sometimes, in fact, stronger predictions can be made about actions people are likely to avoid than about actions they are likely to perform. It is difficult to predict who will marry whom in a society, for example, but it is often clear who will not become marriage partners (e.g., siblings).

From a dynamical perspective, states in which a system cannot stabilize and from which the system escapes are termed "repellors." It is possible to construct dynamical models in which specific states and their associated "basins of repulsion" can be specified (Nowak, Lewenstein, & Tarkowski, 1994). Although these models have not been developed in the context of social psychology, one might speculate that repellor dynamics are relevant to any aspect of human experience involving avoidance or repres-

sion. Negative standards of self-regulation, for example, may be represented as repellors (Carver & Scheier, 1999, 2002). This is not to suggest that repellors are synonymous with conscious efforts to avoid a state. In a multistable system, for instance, any two adjacent fixed-point attractors are separated by a repellor. Movement out of the basin of attraction for one attractor is avoided, but the person's subjective experience revolves around the state defining the attractor.

Periodic Attractors

Some systems display sustained rhythmic behavior rather than convergence on a stable value over time. This temporal pattern is referred to as a periodic or limit-cycle attractor. Periodic attractors capture the essence of many biological phenomena (cf. Glass & Mackey, 1988), such as circadian rhythms and menstrual cycles. There is reason to think that cyclicity is associated with psychological phenomena as well (Gottman, 1979). Moods, for example, have been shown to demonstrate periodicity, often corresponding to a weekly cycle (e.g., Brown & Moskowitz, 1998; Larsen, 1987; Larsen & Kasimatis, 1990). Individuals have been found to differ in this tendency, with some people reliably demonstrating greater temporal variation in mood than others (Penner, Shiffman, Paty, & Pritzche, 1994). Research into the intrinsic dynamics of social judgment (Vallacher et al., 1994) and self-evaluation (Vallacher, Nowak, Froehlich, & Rockloff, 2002), meanwhile, has found that the stream of thought often oscillates between positive and negative assessments, sometimes in accordance with remarkably fast time scales. Periodic structure also characterizes human action (Newtson, 1994) and is apparent as well in social interaction (Beek & Hopkins, 1992; Gottman, 1979; Nezlek, 1993).[4]

A periodic attractor may be difficult to distinguish from the existence of multiple fixed-point attractors. In both cases, the system has a tendency to move between different states over time. The distinction between these two forms of attractor dynamics centers on the regularity of the movement and the role of external factors in producing the movement between states. A periodic attractor is defined in terms of a repetitive temporal pattern, such that the values of the dynamical variable repeat after a time T, $x_i(t) = (t + T)$, where T is the period of motion. Thus, even in the absence of noise or external influence, the state of the system undergoes constant change. For a pattern of change to qualify as an attractor for a system, then, it must represent a pattern on which the system converges, and to which it returns after small perturbations. In a daily cycle of activity, for example, a departure (e.g., a sleepless night) might temporarily disrupt the pattern (e.g., oversleeping the next few days), but eventually the pattern will be restored.

In contrast, a system characterized by fixed-point attractors exhibits a tendency to stabilize on a particular state or set of states. Because such attractors capture all trajectories within their respective basins, a movement from one stable state to another necessarily reflects a disturbance, noise, or external influence operating on the system. A person with self-regulatory standards for both assertiveness and conciliation, for example, will display one of these tendencies as long as the context surrounding the person is within the basin of attraction for that tendency. If the two attractors differ in the size of their respective basins, and if contexts are avoided that attract the person's mental, emotional, or behavioral state toward the smaller basin, the person may operate for extended periods of time in accordance with the stronger attractor. Similarly, an intimate couple may have fixed-point attractors for both positive and negative affective states, but whether they display periodic movement between them will depend on the starting conditions associated with their interactions. Even if the couple demonstrates fairly regular change between positive and negative states, each of these states provides at least temporary stability. In periodic evolution, on the other hand, stability is not afforded by any particular state but, rather, by the pattern of changes between states.

The distinctiveness of fixed-point attractor dynamics was observed in a study exploring the temporal trajectories of affective states on the part of bipolar depressive individuals (Johnson & Nowak, 2002). Time-series analysis of mood and other symptoms revealed that patients whose temporal dynamics were not characterized by fixed-point attractor tendencies were at highest risk for suicide and were hospitalized more often for their depression. Interestingly, these risks were equally low for individuals whose moods oscillated around a single attractor, even one corresponding to a depressed state, and those whose moods switched between two distinct attractors representing a normal state and a depressed state. These results may be interpreted in light of the connection between attractor dynamics and self-regulatory tendencies. Self-regulation implies approach and stabilization regarding some states and avoidance and destabilization of other states. The stable states are manifest as fixed-point attractors for a person's dynamics. From this perspective, the absence of fixed-point attractors for one's affective state signals a breakdown in the capacity for self-regulation.

Deterministic Chaos

Perhaps the most popular insight concerning nonlinear dynamical systems involves *deterministic chaos*. Indeed, many researchers—especially those outside mathematics and physics—often discuss the primary insights from the work on nonlinear dynamics as chaos theory (cf. Gleick, 1987). It is true that only nonlinear systems can display chaos, but as we have seen, such systems are capable of displaying far less spectacular temporal patterns (e.g., fixed-point attractors) as well.

The fascination with chaos reflects the fact that although the behavior of a chaotic system is fully determined, it may be totally unpredictable in practice. This seeming paradox reflects the system's *sensitivity to initial conditions*. In modeling weather patterns, for instance, Lorenz (1963) found that even the most trivial changes in initial conditions, such as rounding the initial humidity and ambient temperature to the third instead of the sec-

ond decimal point, eventually led to entirely different weather patterns. In a chaotic system, anything short of infinite precision in the knowledge of a system at one point in time can undermine knowledge of the system's future states. This occurs because all initial inaccuracies are amplified by the system's intrinsic dynamics, so that the inaccuracies grow exponentially over time. Exponential growth, in turn, ensures that after some finite (often quite short) time, the size of the error will exceed the possible range of states of the system's behavior.

Of course, one can never specify completely the initial conditions for any real-world system, because there is always some error of measurement. It is a matter of time, then, before any initial imprecision grows to values that make prediction impossible. Beyond measurement error, slight and momentary perturbations of a system's dynamics can cause arbitrarily large effects after some time. In the popular literature, this aspect of chaotic systems is known as the "butterfly effect." Lorenz coined this phrase to capture the illusive nature of weather forecasting. He suggested that the flapping of a butterfly's wings in one location—say, Brazil—could generate a cascade of events culminating in a tornado in a distant location such as Texas.

Chaos is always a possibility in nonlinear dynamical systems and, in fact, has been demonstrated in many biological and physical phenomena. It is conceivable, then, that human thought and behavior may often follow a chaotic trajectory. Indeed, social psychology is replete with nonlinear phenomena, such as threshold functions, inverted-U relations, and complex interactions among causal variables (Nowak & Vallacher, 1998a; Vallacher & Nowak, 1997). Despite this potential, however, unequivocal evidence for deterministic chaos in human thought and behavior remains to be documented.

To be sure, human behavior often appears to be unpredictable. But the source of unpredictability in chaotic systems is different from that in random systems. When unpredictability reflects randomness, either no deterministic rules exist for dictating the system's behavior or so many factors influence the system that it is impossible to account for all of them except by treating their combined effects as random influences or noise. In a chaotic system, in contrast, unpredictability results from the nonlinear interactions among a small number of variables. Indeed, as few as three elements can interact in sufficiently nuanced fashion to produce a chaotic—hence, unpredictable—trajectory for the system as a whole (Poincaré, 1908/1952). Deterministic chaos may well play a role in human affairs, but it is also the case that human behavior is often the nexus of innumerable influences, so randomness must be seriously considered when a particular phenomenon proves difficult to forecast.

Nonlinearity, Complexity, and Emergence

Intuition leads us to believe that behavior conforms to two basic rules: one concerning cause and effect, the other concerning complexity. The intuition about causation is that there should be proportionality between cause and an effect. When we see a large effect, it seems reasonable to attribute the effect to a correspondingly large cause. Minor effects, in turn, are reasonably attributed to smaller, incidental causes. It is hard for people to accept that one insignificant disgruntled young man could command the world's attention, perhaps even change the course of history, by assassinating John F. Kennedy in 1963. And who could imagine that a slight peculiarity in the connection between names and punch holes on a voting card would determine the outcome of the U.S. presidential election in 2000—in essence, a "butterfly ballot effect."

The complexity intuition concerns the relation between surface and explanatory complexity. When something behaves in a complex or intricate fashion, we naturally assume that the behavior is a manifestation of complex processes involving multiple factors. We assume that simple thoughts, for example, represent simple assumptions and a paucity of relevant information, whereas complex or nuanced thoughts stem from multiple, perhaps conflicting, assumptions and an abundance of relevant information.

Prior to the advent of the mathematical theory of nonlinear dynamical systems, scientists tended to share these assumptions about the workings of reality. Generally speaking, it was assumed that causal relations were linear—that there was proportionality between cause and effect—and that the complexity of a system's behavior was a direct reflection of the number of interacting elements and the complexity of their mutual influences. Scientists recognized that many natural phenomena involved nonlinear relations, but they believed that such relations could be approximated with adequate precision by linear equations. The revolution in science during the late 1970s and early 1980s reflected a realization that even simple systems consisting of a few elements often exhibit behavior of enormous complexity when the interactions among elements are nonlinear as opposed to linear.

Nonlinearity means that the effects of changes in one variable are not reflected in a proportional manner in other variables. A variable may increase in magnitude with no change in the magnitude of another variable, for example, until a threshold is reached, beyond which even slight changes in the first variable can promote large changes in the second variable. In a nonlinear system, moreover, system-level behavior cannot be decomposed into separate additive influences. Rather, the relations among variables depend on the values of other variables in the system. Even if one's theoretical concern centers on the relation between two variables, one cannot ignore the influence of myriad other variables relevant to the phenomenon.

Nonlinear systems are capable of far richer behavior than can be generated in linear systems. Even a system consisting of a few elements can display highly complex macrolevel behavior if the elements interact in a nonlinear manner over time. The generation of macrolevel properties and behavior from the internal workings of a nonlinear system is referred to as *emergence* (cf. Holland, 1995). Emergence is reminiscent of pattern formation in Gestalt psychology and is captured by the phrase, "the

whole is greater than the sum of its parts." Emergence is due to self-organization among elements, such that each element adjusts to the current state of other elements. No higher-order agent is required for the emergence of such structures. Rather than being imposed on the system from above or from outside the system altogether, they emerge from the nonlinear relations among the system's elements. The higher-order properties that result from the mutual adjustment among lower-level elements provide coordination for the lower-level elements (cf. Haken, 1978; Kelso, 1995).

Emergence via self-organization has implications for the achievement of higher-order structure at different levels of social reality. At the individual level, specific movements and perceptions become coordinated to produce meaningful action (cf. Kelso, 1995; Newtson, 1994), detailed representations of behavior are integrated to form comprehensive representations of the action (Vallacher et al., 1998), and the cognitive and affective elements populating the stream of thought influence each other to promote the emergence of higher-order judgments (Vallacher et al., 1994) and self-concepts (Nowak et al., 2000). The emergence of system-level properties is apparent as well in interpersonal phenomena. Classic accounts of group and societal dynamics have noted that shared norms and beliefs often develop through the spontaneous coordination of group members' impulses and actions, without the intervention of a higher-level authority to impose rules and standards (Durkheim, 1938; Turner & Killian, 1957).

The self-organizing nature of groups and societies has been verified in recent empirical research and computer simulations on social influence and interdependence (e.g., Axelrod, 1984; Latané & Bourgeois, 1996; Messick & Liebrand, 1995; Nowak et al., 1990; Nowak & Vallacher, 1998b, 2001). This work shows that social interaction with one's neighbors is responsible for the emergence of public opinion, altruistic values, and other group-level properties. Thus, individuals become integrated into higher-order functional units such as dyads and social groups (e.g., Nowak et al., 2002), and social groups become coordinated with respect to larger goals and values that define the social system in which they are embedded (cf. Nowak & Vallacher, 2001).

PERSONAL VERSUS SITUATIONAL CAUSATION

Social psychology is preoccupied with the "person versus situation" issue. On the one hand, a variety of theories and research traditions emphasize internal forces that maintain a semblance of stability in thought, emotion, and action. Thus, people are said to maintain or protect their self-esteem (Tesser, Martin, & Cornell, 1996); verify a particular level of self-regard (Swann, 1990); act in accordance with personal standards of achievement, morality, and other values (e.g., Carver & Scheier, 1999; Higgins, 1996); resist attempts by others to dictate their behavior or change their beliefs (e.g., Brehm & Brehm,

1981; Knowles & Linn, 2004); and express a small handful of personality dispositions (e.g., McCrae & John, 1992). The power of the person is often characterized in positive terms such as self-regulation, self-control, self-efficacy, willpower, and ego strength, although it can provide fodder for a focus on dysfunction as well, as reflected in such pejorative (and sometimes clinical) notions as dogmatism, inflexibility, and obsessive–compulsive disorder.

On the other hand, social psychology emphasizes the power of the situation to shape people's internal states and overt behaviors. In this view, people are highly responsive to social norms, authority figures, the presence of others, social feedback, incentives and threats, social influence strategies, environmental conditions such as temperature and crowding, and a host of other general and localized factors. Such situational factors are commonly operationalized as independent variables in experimental research and are assumed to provide the proximate causes of human thought, affect, and action. Internal states are not denied in this perspective, but their purported role in directing thought and behavior is often marginalized when clear situational forces are at work. Some accounts emphasize person-by-situation interactions, in which the expression of personal characteristics (traits, motives, etc.) is either enhanced or minimized in different situations or takes on different forms as a function of variation in situational factors.

The dynamical perspective reframes the person-versus-situation issue somewhat. The basic idea is that external factors do not promote change directly but, rather, shape thought and behavior by influencing the person's intrinsic dynamics. This scenario is manifest in different ways depending on the magnitude of the situational influences and characteristics of the person's intrinsic dynamics. One is likely to observe the power of the person under some configurations of influence and intrinsic dynamics but the power of the situation under other configurations.

The Power of the Person

When intrinsic dynamics are constrained by an attractor, situational influences may have an effect on people's thought and behavior. However, this effect is likely to be transient in nature because when a system is at its attractor, it tends to counter any force that would move it out of the attractor and will return to the attractor if perturbed. The power of the person, then, represents both the ability to resist change despite situational influences and the ability to return to an unperturbed state once the influence subsides. This suggests that situational influences may be most visible directly after their occurrence but fade over time as personal properties, represented as attractor tendencies, regain their prepotence.

Evidence for this idea is provided by research examining the intrinsic dynamics of self-evaluation (Vallacher, Nowak, Froehlich, & Rockloff, 2002). Among participants with a relatively coherent self-concept—and thus presumably with a strong attractor for self-evaluation—

the flow of self-reflective thought initially showed the effect of a manipulation designed to prime either positive or negative thoughts about the self. Over time, however, the evaluative tone of participants' self-reflection increasingly conformed to their chronic self-evaluation tendencies. Participants with high self-esteem who were asked to recall past events that reflected negatively on themselves, for example, tended to express negative self-aspects at the outset, but these sentiments gave way over the course of self-reflection to positively toned self-relevant thoughts.

In a related vein, research on affective forecasting (e.g., Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000) has shown that people commonly overestimate how enduring a mood will be after experiencing a pleasant or an unpleasant event. Even major life events that would seem to alter one's view of the world—whether winning a lottery or experiencing a debilitating illness—typically have relatively short-term effects. Over time, the person's characteristic worldview and mood tend to overpower the mood-inducing effect of the event.

Research has identified various mechanisms that promote resistance to, and recovery from, situational influences that have potential to impact people's attitudes, beliefs, self-concepts, and response tendencies. Social cognition research, for example, has investigated biases in thought and memory (e.g., selective attention and confirmatory bias) that enable people to resist challenges to their judgments about other people and social groups (cf. Kunda, 1999). The work on self-esteem maintenance (Tesser et al., 1996), meanwhile, illustrates the interchangeability of various means for protecting self-esteem against incoming information (e.g., social feedback) that holds potential for undermining positive self-evaluation. Theory and research on self-verification (cf. Swann, 1990) has shown that self-protective tendencies operate to maintain any level of chronic self-regard, even a largely negative self-view that is presumably not consciously desired.

People often have different responses to the same external forces. Some people seem unaffected by information and social influence that contradicts their beliefs or self-concepts, whereas others may readily succumb to such perturbations. Yet others show an exaggerated and highly defensive response to the same factor. Such individual variation can be cast in terms of the strength and basins of attraction associated with the relevant attractors. The strength of an attractor determines the likelihood of rejecting perturbing influences, while the width of the basin of attraction determines the likelihood of the system returning to the attractor after being perturbed. Thus, a person with a strong attractor is likely to reject contradictory information, and someone with a wide basin of attraction is likely to return to his or her attractor after being influenced to think or act in a different manner. If a person has a weak attractor, on the other hand, relatively small influences may prove sufficient to dislodge a system, and if the basin of attraction is also narrow, the system will not return to its attractor. Under this configuration of attractor properties, the person may be especially vulnerable to inconsistent information or other perturbing influences.

Yet another configuration may be at work for people who react in a disproportionately strong manner to any information that challenges their predominant mode of thought or behavior. Such people may have a strong attractor, which promotes rejection, but a narrow basin of attraction, which reduces the potential for recovery from perturbing influences. This scenario, though clearly speculative pending empirical confirmation, is consistent with theory and research on threatened egotism (Baumeister, Smart, & Boden, 1996). The threatened egoist has inflated but insecure self-esteem and responds defensively, even aggressively, toward others who provide feedback challenging his or her self-appraisal. The same dynamics may underlie traits reflecting closed-mindedness, such as dogmatism and need for closure (cf. Kruglanski & Webster, 1996).

Rigid attempts to maintain a particular point of view while rejecting information and influences that challenge this view may promote a paradoxical effect. Imagine, for instance, a person who wishes to maintain a positive view of a close friend despite exposure to information that implies an unflattering assessment of the friend. By virtue of having a shared valence, the thoughts that are rejected are likely to provide mutual reinforcement for each other and thus hold potential for creating a latent attractor. Assuming the latent negative attractor is unwanted, its potential activation may promote increased effort at maintaining the preferred (positive) attractor. But this effort at mental control is likely to increase the number of rejected thoughts, which can then become organized with respect to the latent attractor. The result is further integration of the latent attractor, which augments its potential for capturing the person's trajectory of thoughts should his or her self-regulatory efforts be disrupted or weakened.

In effect, heightened mental control can initiate a tendency that works in opposition to the control effort. This tendency is particularly likely to become prepotent when cognitive resources are drained or mental control is otherwise undermined. This scenario is speculative at this point, of course, but it follows from the notion of latent attractor and its manifestation is consistent with the well-documented ironic effect of thought suppression (cf. Wegner, 1994).

The Power of the Situation

In a system governed by attractor dynamics, situational influences up to a certain magnitude may have only a transient effect because of the tendency for a system to return to its attractor after perturbation. Beyond a certain threshold (i.e., a magnitude that positions the system outside the attractor's basin of attraction), however, an external factor may have a lasting effect on the person's thought or behavior. In a system with multiple attractors, such an influence will dislodge the system from one attractor and position it in the basin of another attractor. This suggests that even when a strong force is experi-

enced, its manifestation may vary from one person to another, depending on the respective attractors available for each person. Thus, a strong provocation might undermine a person's tendency toward friendliness, but whether it promotes aggression or surrender will depend on which of these alternative behavioral tendencies represents an alternative attractor for the person. The person's attractor landscape, in other words, is critical in determining just what effect a situational force is likely to have.

Although strong influences are typically required to promote a lasting change in a person's thought and behavior, small influences may prove sufficient under some conditions. As noted previously, small external influences are likely to be countered by the system, whereas larger perturbations will be accommodated and even amplified in their effects. This suggests that a series of small effects are likely to have little impact on the long-term dynamics of the system. If these effects are concentrated in time, however, their combined effect may be sufficient to move the system into a basin of attraction for a qualitatively different state. Such a scenario was the emergent outcome of the model developed by Nowak and colleagues (2000) to simulate self-concept structure and process. The same amount of conflicting self-relevant information tended to be resisted if presented in small packets separated by time, but tended to promote noteworthy change in global self-evaluation if presented in a large packet in a short period of time. This result can be understood in terms of the system's ability to restore its state after each small perturbation but its inability to do so when a single massive perturbation is experienced.

Weak situational influences may also have effects on a person if his or her thought and behavior are not constrained by attractor dynamics. Research on action identification is relevant in this regard. This work has established that when people are engaged in unfamiliar, complex, or difficult actions, they are highly responsive to suggestions regarding the effects, meanings, or implications of what they are doing (Vallacher & Wegner, 1985, 1987). This vulnerability to influence reflects the lack of integration associated with unfamiliar and personally difficult actions. Such an action is identified in terms of independent low-level components rather than in terms of an integrated high-level identity that provides a stable and comprehensive interpretation. Isolated in this way, each low-level identity lacks the strength to resist social and situational cues regarding its larger meaning—in effect, it is each thought for itself.

If, however, lower-level action components are integrated into higher-level, meaningful structures, then each element receives support from related elements and the combined resistance of these integrated elements can effectively counter outside influence—in effect, one for all and all for one. This general scenario has been demonstrated in computer simulations of self-concept dynamics (Nowak et al., 2000). Contradictory information presented to clusters of self-relevant information that were poorly organized tended to disrupt the clusters, but the same information presented to well-integrated clusters tended to be rejected.

This reasoning suggests that some people may be more vulnerable to situational influences than others. Individuals differ in their characteristic level of action identification (Vallacher & Wegner, 1989), for example, and this difference has been shown to shape people's response to social influence. Thus, people who think about their behavior in low-level (fragmented) terms tend to embrace bogus feedback about the personality characteristics purportedly revealed in a recent action on their part (Wegner, Vallacher, Kiersted, & Dizadji, 1986). From a dynamical perspective, people who identify their action in low-level terms lack stable attractors for thinking about and regulating their behavior (Vallacher et al., 1998). Lacking such attractors, low-level agents are highly vulnerable to external influence.

People also differ with respect to dimensions of self-concept that are relevant to attractor dynamics. Considerable research in recent years has focused in particular on self-concept certainty (e.g., Baumgardner, 1990; Swann & Ely, 1984; Vallacher, 1978; Vallacher, Nowack, Froehlich, & Rockloff, 2002), self-concept clarity (e.g., Campbell et al., 1996), and self-concept stability (e.g., Kernis, 1993). People with an uncertain, unclear, or unstable sense of self have been shown to be highly responsive to social feedback, social comparison, and social influence (cf. Story, 2004). Two possible attractor configurations could be at work here. First, such people may have multiple and conflicting attractors with narrow basins of attraction. Even a slight change in the initial state induced by a social or situational factor might be sufficient to dislodge the system from one attractor and position it in the basin of another attractor. Alternatively, the uncertain person may have a paucity of fixed-point attractors for stabilizing his or her thought and behavior on specific states. In this view, uncertain people are not only responsive to influence but actively seek it out in an attempt to impose structure and clarity on their mental, affective, and behavioral processes (e.g., Vallacher & Nowak, 2000).

PSYCHOLOGICAL CHANGE

In the preceding discussion, the focus was how people change with respect to their location in an existing attractor landscape. Thus, situational forces were described as moving a person within an attractor or, under some conditions, dislodging the person from one basin of attraction and positioning him or her within the basin of another attractor. There is a more fundamental way, however, in which psychological change may occur. Rather than moving a person (or group) within an existing attractor landscape, change can instead involve reconfiguration of the attractor landscape itself. Change reflecting movement between existing attractors is reversible, at least in principle. Thus, if a person moves from attractor A to attractor B because of a specific influence, then an influence in the opposite direction could

reinstate the prepotence of attractor A. A change in the attractor landscape, however, may have lasting and irreversible consequences because it changes the set of possible stable states for the system.

The variables that control the qualitative features of a system—such as the number, location, and type of attractors—are *control parameters*. When developing a theoretical model, then, the central concern is determining which causal factors function as control parameters and have a qualitative effect on the system's dynamics and which factors have a quantitative effect, changing only the momentary state of the system. In modeling psychological change, the system's control parameters clearly are more important and of greater interest than are variables controlling the momentary state of the system.

Changing the attractor landscape of a system can take two different forms. In a system with fixed-point attractors, first of all, change may involve altering the number, position, and shape of the attractors. This type of change is described in topological terms by catastrophe theory (cf. Guastello, 1995; Thom, 1975). A more comprehensive account that incorporates periodic and chaotic attractors is described in dynamical terms by bifurcation theory (cf. Nowak & Lewenstein, 1994; Ruelle, 1989). The second form involves altering the entire pattern of intrinsic dynamics in a system and transforming the types of attractors available for constraining system dynamics. This type of change is often illustrated with the logistic equation, a simple model that captures important features of dynamical systems (cf. Feigenbaum, 1978; Schuster, 1984).

Changing Fixed-Point Attractors

The cusp catastrophe (Thom, 1975) exemplifies change in fixed-point attractors. In this model, two variables—a *splitting factor* and a *normal factor*—control the attractor landscape of the system. The splitting factor, functioning as the control parameter, decides whether the system reacts in a linear or nonlinear way to changes in the normal factor. For low values of the splitting factor, the system is characterized by a single attractor, the position of which changes in a monotonic fashion with changes in the normal factor. For higher values, the single attractor splits into two attractors whose values diverge with increases in values of the splitting factor.

A model of close relations proposed by Tesser and Achee (1994), recast and tested in dynamical terms by Kozlowska, Nowak, and Kus (cited in Nowak & Vallacher, 1998a), illustrates this scenario. The splitting factor in this model is social pressure regarding the formation and maintenance of close relations. The normal factor is the perceived attractiveness of the potential partner. For low values of social pressure (e.g., nonsalience of social norms), the desired closeness of the relationship was a direct function of perceived attractiveness. For high values of social pressure (e.g., against forming a relationship with a stranger but for maintaining a relationship once it is formed), however, there were two attractors corresponding to preferred closeness—

one for a distant relationship, the other for a close relationship.

This scenario is also illustrated in a model of attitudes that focuses on the relationship between the positivity of information concerning an issue and the person's resultant attitude regarding the issue (Latané & Nowak, 1994). The splitting factor in this model is the importance of the issue to the person. For low values (i.e., unimportant issues), there is an approximately linear relationship between information positivity and the attitude. For high values (important issues), however, the relation between information positivity and attitude takes a nonlinear form. When the information concerning an attitude object is predominantly negative, only a negative attractor exists for the person's attitude. When the information increases in positivity, a second attractor appears corresponding to a positive attitude. As the positivity of information increases yet further, the negative attractor disappears and only the positive attractor exists.

Changing Attractor Types

Control parameters can change not only the position and number of attractors but also the types of attractors for a system's dynamics. The simplest and most commonly employed example of this role of control parameters involves the *logistic equation* (cf. Schuster, 1984). The logistic equation involves one variable, x, whose values change in time, and one control parameter, r, which is a coefficient of the equation. The equation specifies what the next value of x will be depending on the current value of x. There are two opposing components of this dependence. First, the higher the value of x at a given moment, the higher the value of x at the next moment—specifically, x_{n+1} equals x_n multiplied by the value of r. Second, the higher the value of x at a given moment, the *lower* the value of x at the next moment—specifically, x_{n+1} equals $(1 - x_n)$ multiplied by the value of r. The combined effect of these two forces is expressed as $x_{n+1} = rx_n(1 - x_n)$, where x_n is the value of the dynamical variable at one moment, x_{n+1} is the value of a dynamical variable at the next moment, and r is the control parameter (the crucial variable influencing changes of x over time). The equation thus expresses conflicting tendencies in a phenomenon and has been employed as a model of many natural science phenomena.

Conflict involving the coexistence of two opposing tendencies is well recognized in social psychology. In the approach–avoid situation (Miller, 1944), for example, movement toward a goal increases both approach and avoidance tendencies. In analogous fashion, theory and research concerning achievement motivation have identified two concerns, the desire for success and the fear of failure, that combine in different ways to produce resultant motivation (e.g., Atkinson, 1964). Theory and research on thought and action suppression, meanwhile, suggest that attempts at suppression activate an ironic process that works at cross-purposes with the attempted suppression (Wegner, 1994). Many other phenomena in psychology can be framed in terms of conflicting tenden-

cies or forces (e.g., impulse vs. self-control, short-term vs. long-term self-interest, egoism vs. altruism, and autonomy vs. social identity).

The temporal pattern of the dynamical variable, x, is totally dependent on the value of the control parameter, r. For low values of r, x will always converge on a single value that represents a fixed-point attractor of the system. For a higher range of values of r, x will oscillate between a set of values, with the period (i.e., number of values through which x oscillates) increasing with the value r. In this case, x follows the scenario of a periodic attractor. For a yet a higher range of values of r, the evolution of x becomes very complex, with the exact sequence of values never repeating. In this case, the attractor for x becomes chaotic.

Exact empirical tests of the application of the logistic equation to social psychological phenomena have yet to be devised. Theoretical models, however, have been proposed in such domains as consumer behavior, close relationships, and the dynamics of approach–avoid conflict (cf. Nowak & Vallacher, 1998a). The dynamics of judgment and decision making might also conform to the logistic equation. In this application, the complexity of relevant information might function as the control parameter, with increases in complexity promoting systematic changes in the attractor defining a person's mental process. Relatively simple information might promote convergence on a stable value (e.g., a decision or judgment)—a fixed-point attractor. As the information becomes increasingly complex, however, the stable value may give way to periodic movement between alternative decisions or evaluative states, neither of which provides for stability in the cognitive–affective system. This pattern might be experienced as ambivalence or conflict between different assessments of the appropriate course of action. Yet further increases in information complexity might promote multiperiodic or even chaotic trajectories of thought and feeling. This dynamic pattern might be manifest as indecision or perhaps confusion, and experienced emotionally as heightened arousal, agitation, and anxiety.

The logistic equation, in sum, shows that the same system can display entirely different patterns of behavior (e.g., fixed-point, periodic, and chaotic attractors) depending on changes in the values of a control parameter. There is a clear distinction between external influences or perturbations that change of value of x and control parameters that change the pattern of the temporal evolution of x. The logistic equation also highlights the possibility that whenever the dynamics of a system are governed by opposing forces, the system has the potential for very complex dynamics, even if the system consists of a very small number of variables.

COHERENCE IN SOCIAL PSYCHOLOGICAL SYSTEMS

Despite the enormous complexity of human minds and social groups, psychological systems at different levels of social reality display self-organization and the emergence of higher-order properties and processes. Emergence, however, is not a one-step process. To the contrary, once global properties and higher-order processes come to characterize a system, they effectively become lower-level units that can become integrated with respect to yet higher-order properties. In principle, this process of progressive integration can proceed until the system as a whole can be characterized in terms of a single macro-level state that functions in accordance with a single process. In reality, the tendency toward progressive integration is likely to "stall" at a level that promotes differentiation rather than unification in the phenomenon at issue.

The critical factor dictating the level at which integrative processes stall is coherence (cf. Nowak et al., 2000; Simon & Holyoak, 2002; Thagard, 1989). A set of elements is coherent to the extent that they are consistent in their implications for thought, emotion, or behavior, depending on the phenomenon. Thus, a set of thoughts relevant to social judgment is coherent if they collectively convey an unequivocal evaluation of someone, whereas a set of low-level act identities is coherent if they are sufficiently coordinated to promote effective performance of the action under a higher-level (comprehensive) act identity. In both cases, the challenge of attaining higher-order coherence may stall the process at a level well beneath that of global evaluation or action mastery. Instead, social judgment will reflect a differentiated view of the target (cf. Kunda & Thagard, 1996) and action identification will be characterized by the prepotence of various lower-level act identities (Vallacher & Wegner, 1985, 1989).

Coherence and the potential for progressive integration may have a wide range of application in social psychology, providing a principle that links otherwise distant phenomena. To date, however, research designed to illuminate the building of psychological structures in this fashion is limited to a few domains. Next, we discuss how coherence is relevant to emotion, the emergence and maintenance of self-concept, and the formation of public opinion.

Emotion and Affective Experience

Perhaps the most basic and widely accepted division of psychological processes is the tripartite distinction among action, cognition, and emotion. Numerous and diverse accounts have been forwarded to depict how these systems interact, but to a large extent these accounts all suggest that emotion serves to signal the state of the other systems or the quality of coordination between them (e.g., Carver & Scheier, 1999; Csikszentmihalyi, 1990; Festinger, 1957; Heider, 1958; Higgins, 2000; Kruglanski & Webster, 1996; Mandler, 1975; Simon, 1967; Thagard & Nerb, 2002; Vallacher & Nowak, 1999; Vallacher, Wegner, & Somoza, 1989; Winkielman & Cacioppo, 2001). In general, negative emotion signals incoherence: poor processing fluency in perception, inconsistency or uncertainty in judgment, discrepancy between mental representations (e.g., goals, attitudes) and behavior, and poor coordination between mind and action. Depending on the theoretical account

and the system or systems at issue, the negative emotional states are sometimes diffuse (e.g., arousal and agitation) and sometimes fairly specific (e.g., guilt, disliking, and self-consciousness).

The flip side of the notion that incoherence promotes negative emotion is that systems strive for coherence (e.g., Kruglanski & Webster, 1996). Insight into the dynamic underpinnings of this scenario is provided by Lewenstein and Nowak (1989a, 1989b), who introduced self-control mechanisms in attractor neural networks (Hopfield, 1982). Attractor neural networks (a subclass of connectionist models) are programmable dynamical systems. In this approach, familiar (previously experienced, overlearned, etc.) stimuli are encoded as attractors of the network. In the vicinity of an attractor (i.e., when a familiar stimulus is perceived), the signals arriving at a given neuron from other neurons are consistent in dictating the state of the neuron. Coherent signaling, in other words, signifies an energy minimum (equilibrium) in the network. When the network is far from an attractor, however, the system is characterized by incoherence—the signals arriving a given neuron from other neurons dictate conflicting states of the neuron.

Network incoherence is used to construct a self-control feedback loop. The control parameter is the level of "noise" in the network, corresponding to the random component of neuron firing. Self-control is established by making the noise level dependent on the degree of network incoherence. With increasing levels of incoherence, the network increases its noise level, which in turn decreases the coherence of the network, and so on, in a self-perpetuating manner. This has the effect of making progressively stronger attractors inaccessible, until at some point no attractor can capture the system's dynamics. On the other hand, because familiar stimuli are coherent and hence produce low levels of noise, the feedback loop results in their correct recognition. This feedback loop, in sum, enables the network to regulate its own recognition process. The recognition process is maintained when the network is characterized by coherence (produced by familiar patterns) but is interrupted by high levels of noise when the network detects its own incoherence (produced by unfamiliar patterns).

This scenario can be used to interpret the nature of coherence in different psychological systems (Vallacher & Nowak, 1999). The units in an attractor network are not restricted to neurons but can take on a variety of identities, from cognitive elements (in the case of judgment) to individuals (in the case of group dynamics). Noise, in turn, corresponds to negative emotion signaling incoherence. Depending on the phenomenon, noise could thus represent a wide variety of emotional states, including diffuse arousal, acute anxiety, self-conscious emotions such as embarrassment or guilt, or negative affect toward a stimulus. The detection of incoherence, signaled by negative emotion, engages the self-regulation feedback loop to restore coherence.

Consider, for example, the operation of action systems. In a malfunctioning system or one that has been destabilized by external influences, the action elements are no longer coherent (i.e., they provide conflicting signals to one another) and thus cannot be coordinated into an effective pattern. This can happen, for example, when cues to action are ambiguous or conflicting, or when novel circumstances disrupt the normal course of action. This promotes negative emotion (e.g., anxiety and self-doubt) and has the effect of disassembling the action system into its lower-level elements, which can then be reassembled in a different configuration that restores coherence (and effectiveness) to the action.

This scenario does not imply that a person's affective state goes flat when coherence is achieved in a psychological system. Rather, the attainment of system coherence is associated with positive emotion. When an action system, for example, functions autonomously and has the potential for continuing self-organization, it may produce a special affective state commonly referred to as "flow" (Csikszentmihalyi, 1990). Because positive emotions signal coherence and effective self-regulation at a given level, moreover, they free up consciousness from monitoring the functioning of the system and thus allow attention to move upward to assemble higher-level systems. This perspective is interesting in light of recent consideration of independence and asymmetry with respect to positive and negative emotions (cf. Cacioppo, Gardner, & Bernston, 1997; Carver & Scheier, 1999; Higgins, 2000). Whereas negative emotions focus attention on the internal workings of an incoherent system, positive emotions are associated with progression of consciousness to higher levels of integration. In the context of action, for example, effective self-regulation at a basic level of mental representation (e.g., hitting a tennis ball) not only feels good but also directs attention to the action's more comprehensive meanings, such as its purpose, consequences, and implications (Vallacher & Wegner, 1987). Positive emotions, in effect, provide the hedonic basis for progressive integration in psychological systems.

Considered together, the progressive assembly of lower-order systems into higher-order systems and the repair of disrupted systems impart a dynamic quality to affective and cognitive experience. Each time one's attention is diverted from a higher-order system to lower-level elements and systems, there is a press for consciousness to reassemble the elements into a higher-order system. To the extent that the resultant coordination differs from the earlier pattern of interelement connections, the emergent higher-order system may be qualitatively different from the original system. With each enactment of the disruption-repair pattern, then, there is potential for the creation of a new higher-order system of self-regulation. From a dynamical perspective, the content of mind is open-ended and ever-changing, representing a constructive process that fosters adaptation to changing demands and conditions.

Self-Reflection and the Emergence of Self-Concept

The self is arguably the largest and most chronically accessible structure in a person's mental system. Every facet of personal experience is potentially relevant to a person's self-understanding, from the details of his or

her physical appearance to his or her self-perceived traits, values, and personal aspirations. The voluminous information relevant to the self that is encountered on a daily basis is highly diverse in content and valence, ranging from incidental events to consequential feedback from significant others and success versus failure in personal pursuits. The nonstop exposure to distinct, often inconsistent, pieces of self-relevant information would seem to undermine the formation of a stable and coherent sense of self. Yet people somehow manage to develop relatively certain and coherent conceptions of themselves with respect to higher-order constructs such as traits, skills, and goals.

In addressing this issue, Nowak and colleagues (2000) conceptualized the self-structure as a complex system composed of cognitive elements representing self-relevant information, with mechanisms of self-organization promoting coherence and stability in self-concept in much the same way that that such mechanisms promote social consensus among autonomous agents in society (Nowak et al., 1990). Although the elements of self-structure are diverse by many criteria, they can all be scaled with respect to evaluation. Nowak and colleagues assumed that the elements do not have a fixed valence but, rather, that they influence each other to adopt a common evaluative state. An element incongruent with neighboring (thematically related) elements may change its valence or change the valence of its neighbors, to establish evaluative coherence with the related elements. The recognition that one is distractable, for example, may take on positive rather than negative valence in the context of other self-perceived qualities that together convey an image of oneself as a creative scientist. This press for integration generates subsets of self-relevant information that have a shared evaluation. The self-structure thus becomes differentiated, with different regions stabilizing on different values of self-evaluation (e.g., Showers, 1992). A person may have a coherent and positive view of him- or herself as a scholar, for example, and an equally coherent but negative view of him- or herself as an athlete.

Nowak and colleagues (2000) developed a cellular automata model to simulate the emergence of local coherence in self-structure. In this model, the self-system consists of n elements, each reflecting a specific aspect of the self, which are represented as cells arranged on a two-dimensional grid (see Figure 32.2). The physical proximity of any two elements represents their degree of relatedness. Each element is characterized with respect to its current evaluation, which is either positive (denoted by light gray) or negative (dark gray). Some elements are more important than others and have greater weight in self-evaluation. An element's importance, which remains constant in the course of simulation, is denoted by its height. Each element influences and is influenced by its eight neighboring elements (four on the adjacent sides and four on the connecting diagonals).

In the course of simulation, an element chosen at random tries to adjust to its neighboring elements by checking how much influence it receives from them. This process involves weighting the valence of each neighbor by the neighbor's importance. The result of this computation is the weighted sum of evaluations of the neighboring elements. This result is then compared to the current valence of the element. If the element's valence agrees with the overall evaluation suggested by its neighbors, the valence does not change. If the valence of the element differs from the overall evaluation suggested by its neighbors, the element changes evaluation only if the combined weight of evaluation from the neighboring elements is greater than the element's own weighted evaluation. It is relatively easy for neighboring elements to change the evaluation of a relatively unimportant element, in other words, but it is difficult to change the evaluation of a more important piece of self-relevant information.

This process is repeated for another randomly chosen element, and then again for another element, and so on, until each element has been chosen. In the next simulation step, each element has a chance to adjust its state again. The simulation steps are repeated the system reaches an asymptote, indicating no further changes in the state of elements (i.e., static equilibrium) or a stable pattern of changes in the system (i.e., dynamic equilibrium).

As Figure 32.2 illustrates, the process of mutual adjustment among elements of self-relevant information promotes the emergence of clusters. Self-relevant information that is randomly distributed at the outset (top picture) forms well-defined domains composed of elements that share a common valence (bottom picture). The emergence of evaluatively coherent clusters is due to the local nature of influence among elements. The self-system also becomes more polarized in overall evaluation, with more negative elements switching to positive valence than vice versa. In a disordered system, the proportion of positive and negative elements in a region roughly corresponds to the proportion of positive and negative elements in the entire structure. Hence, any given element is likely to be surrounded by more positive than negative elements and thus is likely to experience greater influence in the positive direction. Once the self-
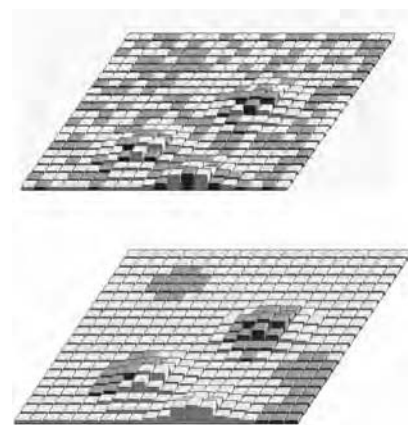
FIGURE 32.2. The emergence of coherence in self-structure (Nowak et al., 2000).

structure has become clustered, however, most elements are surrounded by elements of the same valence, so that only the elements on the border of a cluster are subjected to conflicting influences.

The emergence of locally coherent regions tends to stabilize the self-system. This is because each element in a coherent region supports the current state of the other elements. A change in one element results in joint influence of the other elements for the element to return to its original value. In effect, coherent regions function as attractors, where the value of each element is anchored in the value of the other elements. The joint influence of all the elements in a region promotes resistance to change, and enables an element to return to its original value if it is changed. In an incoherent region, in contrast, the current state of an element is supported by some elements but undermined by others. Hence, when an element is influenced from the outside, some of the surrounding elements help the element resist the influence, whereas other elements work in the direction of the influence. If an element is overwhelmed by outside influence, there will be little tendency for it to return to its original state because some of the neighboring elements are likely to support the new state.

Computer simulations provided support for coherence as a basis for the attracting tendencies of clusters. Integrated self-structures could withstand external influence and were able to rebound to their original state after being perturbed. Unintegrated self-structures yielded more readily to influence and displayed weaker tendencies toward restoration of their original state. Figure 32.2 also shows that although the proportion of positive elements increased, the negative elements that manage to survive tended to be more important and hence resistant to subsequent changes. This is consistent with research on the negativity effect in judgment (e.g., Cacioppo et al., 1997; Pratto & John, 1991; Skowronski & Carlston, 1989). Although positive information tends to be more prevalent than negative information in cognitive structures, the relatively few elements of negative information tend to be more important and hence more salient.

Despite the press for integration, evaluative coherence with respect to higher-order constructs such as traits, competencies, and roles may be difficult to achieve and maintain. An aspect of self-concept such as "honesty," for instance, may consist of conflicting self-views that render a unified self-evaluation impossible. Rather than having a global coherent self-view with respect to honesty, a person might differentiate this dimension into more basic dimensions, such as academic versus interpersonal honesty. If each of these lower-level assessments also prove to be internally inconsistent, the person's self-concept will be correspondingly more fragmented.

Specificity in self-assessment may reflect reality constraints and suggest precision in self-understanding, but it is unlikely to be personally satisfying. Because of the press for higher-order coherence, the person might experience periods of rumination as he or she attempts to integrate the conflicting thoughts, or, alternatively, he or she might engage in a variety of self-defense mechanisms (e.g., reinterpretation, discounting, and denial) that pro-vide subjective coherence with respect to the higher-order dimensions (Tesser et al., 1996). These tendencies are likely to be reinforced by the demands and expectations of social life, which typically emphasize broader categories of self-understanding. People describe and judge one another in terms of motives and traits rather than specific acts (cf. Kunda, 1999), after all, and decision making often pertains to basic goals and values rather than narrowly defined concerns. Hence, someone with an evaluatively mixed view of him- or herself with respect to traits and goals is likely to report relatively low self-concept certainty (Baumgardner, 1990; Campbell et al., 1996; Pelham, 1991; Vallacher, 1980; Vallacher, Nowak, Froehlich, & Rockloff, 2002)—even if the person is certain of what he or she is like with respect to lower-level aspects of self (Vallacher & Wegner, 1989). Such a person is also likely to display instability in the stream of self-evaluative thought, as inconsistent self-relevant thoughts rise and fall in salience on a short time scale (Vallacher, Nowak, Froehlich, & Rockloff, 2002).

Social Interaction and the Emergence of Group Structure

One would think that putting individuals together in groups would create a breeding ground for chaos and conflict. Particularly in an individualistic society like the United States, the potential for self-interest to color one's interactions with others might promote highly complex trajectories of group sentiment and behavior as the group members express and try to promote their respective idiosyncratic and egoistic agendas. Several decades of research on group dynamics, however, have pointed to a very different dynamic scenario, one that is similar in form to the emergence of coherence in individual minds. Across a wide range of contexts, groups of interacting individuals tend to become increasingly uniform over time in their opinions, beliefs, and recommended courses of action (cf. Arrow et al., 2000; Festinger, Schachter, & Back, 1950).

People in groups certainly attempt to influence one another. But such attempts usually represent more than a concern with forwarding one's own agenda, centering instead on a concern with achieving a common psychological state and platform for action. In the process of social interaction, group members adjust their opinions, mood, or behavior to promote consensus with one another. Even if the members have diverse opinions and initially disagree with one another, there is nonetheless a tendency for uniform opinions to emerge over time. The group-level product of local social interactions, however, rarely reflects the central tendency of members' individual initial opinions. Rather, the opinion that emerges as a result of social interaction tends to be more extreme than the average of group members' opinions (e.g., Myers & Lamm, 1976).

The tendency for groups to develop a common psychological state is largely adaptive, as it provides the social coordination and consensus necessary for social life and group action. But the tendency for groups to achieve uniformity has downsides as well. Under some condi-

tions, for example, it can promote mindless conformity to inaccurate assessments of reality (cf. Asch 1955; Sherif, 1936). The press for coherence in groups can also produce decisions and recommendations for action that are guided more by a concern for reaching consensus than for developing the best policy. Rather than examining all possible courses of action, people in the grip of "groupthink" (Janis, 1982) expend their mental energy on achieving and maintaining group solidarity and opinion unanimity. Once a group has achieved a common psychological state, moreover, anyone who expresses a contrary view is likely to experience enormous pressure to change. If the person doesn't cave in to such pressure, he or she is likely to be rejected (Schachter, 1951). This has clear implications for a host of phenomena, from peer pressure among adolescents to jury deliberations in cases in which a guilty verdict carries the death sentence (e.g., Hastie, Penrod, & Pennington, 1983).

Insight into the dynamics of coherence in groups is provided by social impact theory (Latané, 1981) and the implementation of a dynamical version of this model (Nowak et al., 1990). The theory holds that influence in a group reflects two basic functions—one describing the combined effect of different people on a single person, the other describing how a single person's influence is divided across different people. Three variables are common to both functions: the number of people influencing or being influenced, the respective strength of these people, and the immediacy of the people to one another. The influence of a group grows as a power function of the number of people, usually with an exponent of approximately 0.5. In other words, the joint effects of the group's influence grows as a square root of the number of group members. Strength represents the potential for influence and refers both to stable individual characteristics (e.g., social status) and to topic-relevant variables (e.g., relevant knowledge). Finally, influence depends on proximity and tends to decrease as a square of the distance.

In the dynamical model (Nowak et al., 1990), each individual is characterized by his or her opinion on a topic, persuasive strength, and position in a social space. For simplicity sake, individuals are commonly assumed to have one of two opinions on an issue (e.g., pro vs. con). The social group is modeled as a cellular automata consisting of n individuals located on a two-dimensional grid (see Figure 32.3). Each box in Figure 32.3 corresponds to an individual. The color of the box (light vs. dark gray) denotes the individual's opinion (e.g., light gray denotes pro, dark gray denotes con), and the height of the box corresponds to the individual's strength. Each individual discusses the issue with other group members to assess the degree of support for each position. The opinions of those who are closest and have the greatest strength are weighted most heavily. An individual's own position is also taken into consideration and is weighted most heavily by virtue of immediacy. Each individual adopts the opinion that is most prevalent in the process of social interaction. The strength of influence of each opinion is expressed as follows:
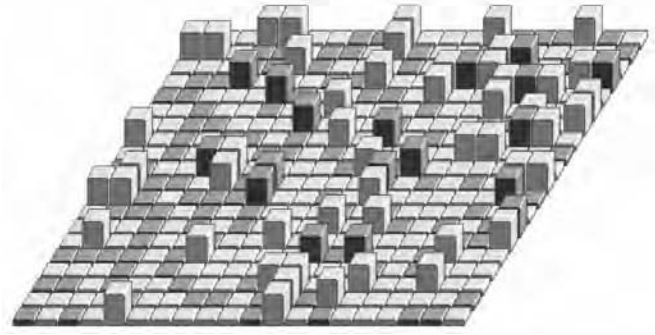
FIGURE 32.3a. Initial distribution of opinions in the simulated group.

$$I_i = \left(\sum_1^N \left(\frac{s_j}{d_{ij}^2} \right)^2 \right)^{1/2}$$

where I_i denotes total influence, s_j corresponds to the strength of each individual, and d_{ij} corresponds to the distance between individuals i and j.

In the simulations, one individual is chosen (usually at random), and influence is computed for each opinion in the group. If the resultant strength for an opinion position is greater than the strength of the individual's current position, his or her opinion changes to match the prevailing position. This process is performed for each individual. This procedure is repeated until there are no further changes, which typically requires several rounds of simulation, because a person who had previously changed his or her position to match that of his or her neighbors may revert to the original position if the neighbors change their opinions. Figure 32.3 presents representative results of the computer simulations. In Figure 32.3a, there is a majority of 60% (light gray) and a minority of 40% (dark gray). The majority and minority members are randomly distributed, and each group has the same relative proportions of strong and weak members (high vs. low boxes). Figure 32.3b shows the equilibrium
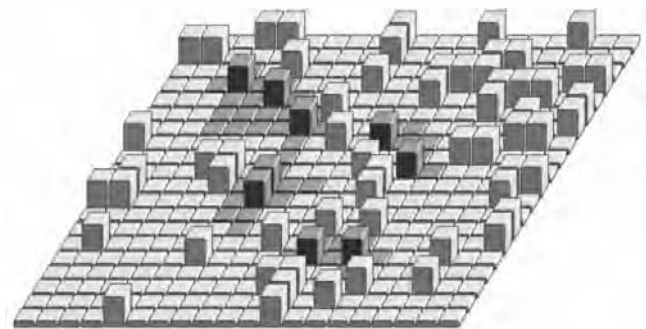
FIGURE 32.3b. Final equilibrium of opinions in the simulated group.

reached after six rounds of simulated discussion. Now the majority is 90% and the minority is 10%. The minority opinion has survived by forming clusters of like-minded people, and these clusters are largely formed around strong individuals.

These two group-level outcomes—*polarization and clustering*—are commonly observed in computer simulations (cf. Latané, Nowak, & Liu, 1994) and are reminiscent of well-documented social processes. As noted earlier, the average attitude in a group becomes polarized in the direction of the prevailing attitude as a result of group discussion. Polarization in the simulations reflects the greater influence of the majority opinion. In the initial random configuration (Figure 32.3a), the average proportion of neighbors holding a given opinion corresponds to the proportion of this opinion in the total group. This means that the average group member is surrounded by more majority than minority members, which results in more minority members being converted to the majority position than vice versa. Some majority members are converted to the minority position, however, because they happen to be located close to an especially influential minority member, or because more minority members happen to be at this location.

Clustering is also pervasive in social life. Attitudes have been shown to cluster in residential neighborhoods (Festinger et al., 1950), for example, and pronounced clustering has been observed for political beliefs, religions, clothing fashions, and farming techniques. Clustering reflects the relatively strong influence exerted by an individual's neighbors. When opinions are distributed randomly, the sampling of opinions through social interaction provides a reasonably accurate portrait of the distribution of opinions in the larger society. When opinions are clustered, however, the same sampling process will yield a highly biased result. Because the opinions of those in the nearby vicinity are weighted the most heavily, the prevalence of one's own opinion is likely to be overestimated. Hence, opinions that are in the minority in global terms can form a local majority. Individuals who hold a minority opinion are therefore likely to maintain this opinion in the belief that it represents a majority position.

Clustering occurs despite the press for coherence that is responsible for progressive integration in psychological systems. Three factors have been identified that effectively stall the integration process, preventing complete unification and hence preserving minority opinions in groups (Latané & Nowak, 1997; Lewenstein et al., 1993; Nowak, Lewenstein, & Frejlak, 1996). Individual differences, first of all, are indispensable to the survival of minority clusters. By counteracting the sheer number of majority opinions, strong leaders stop minority clusters from decaying. As a result of social influence, moreover, individual differences in strength tend to become correlated with opinions. This is because the weakest minority members will most likely adopt the majority position, so that the average strength of the remaining minority members will grow over time at the expense of the majority. This scenario provides an explanation for why individuals advocating minority positions are often more influential than those advocating majority positions (cf. Moscovici, Lage, & Naffrechoux, 1969).

The second factor is nonlinearity in attitude change. Abelson (1979) showed that when individuals move incrementally toward the opinions of their interaction partners, groups invariably become unified in their support of the majority opinion. In the model of dynamic social impact, however, attitudes change nonlinearly in accordance with a threshold function. Thus, individuals hold their opinion until social influence reaches a critical level, at which point they switch from one categorical position (e.g., pro) to the other (con). So whereas a linear change rule, which implies a normal distribution of opinions, promotes unification of opinions, a nonlinear change rule, which implies a bimodal distribution, can prevent complete unification and enable minority opinion to survive in clusters. Latané and Nowak (1994) have shown that a normal distribution tends to develop for relatively unimportant attitudes, but that a bimodal distribution is more often observed for attitudes of high personal importance. This suggests that consensus in a group can be achieved by decreasing the subjective importance of the topic.

The third factor is the geometry of the space in which individuals interact (Nowak, Latané, & Lewenstein, 1994). People do not communicate equally with all members of a group, nor are their interactions often random. In the cellular automata model, different communication patterns can be approximated with different geometries of social space. In the limiting case, geometry is lacking altogether and interactions occur randomly between people. Under these conditions, minority opinion decays rapidly and the group converges on the majority position. Other geometries have been used to capture different communication patterns, and these have been shown to have predictable consequences for the fate of minority opinions. In real social settings, of course, several different geometries are likely to coexist and determine the emergence of opinion structure in groups. Even in a small town, the ready availability of telephones, email, shopping malls, and common areas for recreation add many dimensions to the effective geometry in which interactions take place. The combined features of these geometries are certain to play significant roles in shaping the distribution of public opinion.

TOWARD COHERENCE IN SOCIAL PSYCHOLOGY

Social psychology is an ambitious area of scientific inquiry that has attempted to capture every conceivable nuance of human experience. But although the field's wide-ranging agenda is commendable, it has had the unintended effect of creating a highly fragmented discipline. With a topical landscape ranging from cardiovascular processes and momentary feelings to intergroup relations and societal change, the field has become populated by myriad theoretical orientations

and research strategies that have few points of contact with each other or with other scientific disciplines. The lack of theoretical and paradigmatic coherence has generated concern from several quarters in recent years (cf. Buss, 1995; Krueger & Funder, 2004; Vallacher & Nowak, 1994a). But lamenting this state of affairs is one thing and fixing it is quite another. Is it possible to establish theoretical integration for the complexity and vagaries of human social experience? Can the fragmented field of social psychology be put back together in a coherent fashion?

At first blush, it is easy to reframe diverse social psychological processes in dynamical terms. Human experience qualifies as a complex system, in that any mental, affective, or behavioral process can be analyzed with respect to myriad genetic, hormonal, dispositional, familial, situational, and cultural causes. And because these factors interact with one another over time to promote sustained temporal trajectories of thought, feeling, and action, social psychological phenomena can be investigated as nonlinear dynamical systems. The earliest formulations of social psychological issues were remarkably prescient in this regard. Such pioneers as James (1890), Cooley (1902), Mead (1934), Lewin (1936), and Asch (1946) all emphasized the multiplicity of interacting forces operating in individual minds and in social groups, the potential for sustained patterns of change resulting from such complexity, and the tendency for individuals and groups to strive for mental and interpersonal coherence.

The dynamical perspective, in fact, may prove *too* appealing for those seeking an integrative paradigm for social psychology. Such notions as self-organization, emergence, bifurcation, and chaos have an intuitive resemblance to many personal and interpersonal phenomena. It is tempting to note the penchant for spontaneous coordination of sentiments and actions in social groups, for example, or to suggest that attitudes emerge from the self-organization of specific thoughts. Although such intuitions are compelling, the success of the dynamical perspective will depend on the ability of this approach to go beyond metaphors, intuitive similarity, and general statements to generate explicit theoretical statements and testable hypotheses. With this concern in mind, we outline an approach specifically developed to frame and test hypotheses concerning the dynamical properties of personal and interpersonal experience. To establish the integrative and predictive utility of this approach, we conclude the chapter by summarizing the basic principles (to date) of dynamical social psychology.

Dynamical Minimalism

The subject matter of social psychology is obviously very complex, and a good theory must be able to account for this complexity. In the traditional approach to theory construction, the complexity of human thought and behavior is reflected in the complexity of the model, with many variables and complex interactions among them providing the starting point for an explanation of a phenomenon. The approach of *dynamical minimalism* (Nowak, 2004), in contrast, tries to construct models in such a way that the observed complexity emerges from the simplest possible assumptions rather than being inherent in the model itself. The focus of this approach is on identifying the minimal set of realistic principles and mechanisms capable of producing the phenomenon under investigation. This perspective often assumes simple, almost trivial, assumptions at the level of individual elements yet tries to reproduce the complexity of the phenomenon at the system level. Because the resultant theories provide simple explanations that nonetheless capture the complexity of human thought and behavior, this approach aims to maximize parsimony in theory construction without trivializing the phenomenon in question.

As emphasized in this chapter, simple rules governing the interactions among individual elements can generate very complex properties at the system level. For complexity to emerge from simple rules, though, the rules must interact over time. Thus, a simple theory of a complex phenomenon is necessarily dynamic in nature. Dynamical minimalism is the approach of choice, then, when the relations among elements are nonlinear and the phenomenon displays self-organization and emergence. It is less essential for phenomena governed by linear dependencies, in which the potential for emergence is minimal. Of course, one can develop dynamical theories that do not propose emergence. Such a theory may be valid if one's interest is the nature of the dynamics per se and the theory's assumptions are verified empirically. However, many relationships in social psychology, such as threshold phenomena, inverted-U relations, and statistical interactions, reflect nonlinearity and thus have the potential for emergence if the variables are embedded in a larger system that evolves over time (cf. Vallacher & Nowak, 1997).

The minimalist approach provides a new perspective on the relation between micro- and macrolevels of description. From the perspective of reductionism, the properties at higher levels of description can be directly reduced to properties of elements at lower levels. The rules observed at one level, in other words, directly correspond to the rules observed at another level. The relation between poverty and crime on the social level, for example, may be reduced to the relation between frustration and aggression operating at the level of individuals. In dynamical models, in contrast, the rules operating at one level may generate wholly different rules at a higher level. In the *society of self* model (Nowak et al., 2000), for example, basic rules regarding the integration of basic elements of self-knowledge have many interesting but unanticipated consequences for higher-level self-representation. The model assumes simply that each basic element acquires the prevailing valence of related elements. Repeated iterations of this basic rule have been shown to generate several interesting consequences at the global level of self-representation, such as the differentiation of self-structure into local regions of contrasting valence, and global properties (e.g., self-esteem) that

are relatively immune to external influences and that can rebound after being challenged.

Building scientific theories that capitalize on emergence leads to an apparent paradox. By definition, emergence refers to principles on the system level that cannot be derived by reasoning about the knowledge of the system's elements. In a system characterized by emergence, then, how can the knowledge of lower-level elements serve as an explanation of higher-level properties? The answer highlights the crucial role played by computer simulations in the dynamical approach. Computer models enable one to specify properties of elements and the rules of interaction among them. When the elements interact according to the specified rules, dynamics may be observed at the system level that were not assumed on the level of individual elements. Indeed, a primary reason for constructing computer simulations is to identify the emergent consequences of basic rules. Computer simulations thus allow for a theory formulated under one level of psychological reality to be tested at a different level of psychological reality.

Computer simulations are essential for another reason. Many specific properties of individual elements and their interactions have only minimal effects on properties at the system level. Indeed, the basic elements themselves are typically uninteresting, even trivial in nature. What matters are some very basic properties of the elements and the patterns of interaction among the elements. Consider, for example, the emergence of public opinion modeled with cellular automata by Nowak and colleagues (1990). As noted earlier, this model characterizes each individual in terms of only three properties: his or her location on a two-dimensional grid, his or her attitude (e.g., pro vs. con) on some topic, and his or her persuasive strength. Clearly, any individual is far more complex than this. Many nuances of individual variability (e.g., idiosyncratic traits), however, have little if any impact on the dynamics of public opinion formation in a social group. So although this model hardly does justice to the complexity of individuals, it captures the essential features that are responsible for the emergence of public opinion.

The goal of dynamical minimalism is to build a model that incorporates only the variables that are critical for the emergence of macrolevel properties. In computer simulation models, one can systematically vary the assumptions concerning different properties of elements and their interactions and observe which assumptions result in important changes at the macrolevel. The characteristics that do not have consequences at the system level can be disregarded and omitted from the model. In effect, computer simulations function as a sieve that distills the minimal set of components and their interactions that constitute the essence of the phenomenon of interest.

Although computer simulations play a pivotal role in dynamical minimalism, they cannot substitute for empirical verification of a theory's assumptions. To the contrary, computer simulations are useful in delimiting the crucial assumptions of a theoretical model and thus provide a roadmap for empirical efforts. In a common scenario, computer simulations of processes assumed to operate a lower level may be used to investigate the consequences of these processes at a higher level. These consequences, in turn, may function as hypotheses to be tested in empirical research. Dynamic social impact theory (Nowak et al., 1990), for example, was used to derive predictions concerning spatial–temporal patterns of social change processes. The existence of these hypothesized patterns were then assessed in statistical data concerning patterns of entrepreneurship and voting patterns in Poland following the fall of Communism in the 1990s (Nowak et al., 2005). In addition to testing a model's assumptions, empirical tests can also be used to refine the model. The refined model, in turn, can then be implemented in computer simulations, the results of which can provide further hypotheses to be tested empirically. This reciprocal loop between theory, computer simulation, and empirical research is instrumental to the progress of scientific social psychology.

Principles of Dynamical Social Psychology

The dynamical approach serves as both a useful heuristic and an integrative platform for the diverse phenomena comprising social psychology. Both these functions are reflected in the following principles, which summarize the primary themes conveyed in this chapter.

- *Principle 1. Intrinsic dynamics are inherent in social psychological systems.* Intrinsic dynamics characterize all levels of personal and social reality, from neural function to societal processes. Intrinsic dynamics are produced by multiple influences among the elements comprising the personal or social system. External factors influence thought, feelings, and action by interacting with the system's intrinsic dynamics. In some cases, intrinsic dynamics are manifest as resistance to change and thus promote stability rather than change in behavior. The characterization of intrinsic dynamics is a central task for social psychological research.
- *Principle 2. There is a direct connection between structure and dynamics in social psychological systems.* The interdependency among the variables comprising a system is expressed in the system's pattern of temporal evolution. Nonlinear relationships among system variables, for example, are necessary for chaotic dynamics. Knowing the structure of the system thus enables prediction of the system's behavior, and conversely, knowing the system's dynamics provides insight into the structure of the system.
- *Principle 3. Attractor dynamics provide for stability while allowing for constant change in social psychological systems.* The unperturbed dynamics of personal and interpersonal systems converge on states, or patterns of change between states, that can be described as attractors. There are three basic types of attractors: fixed-point, periodic, and chaotic. Although thought, emotion, and behavior can change on a moment-to-moment basis, identifying the attractors governing their dynamics allows one to characterize psychological processes in terms of stable points or patterns on which the dynamics converge.

• *Principle 4. Latent attractors represent possible equilibrium states for a system that have yet to be manifest*. External influences that do not have an immediate effect on the state of the system may nonetheless have important consequences for the long-term dynamics of the system by creating a latent attractor. The existence of latent attractors creates the potential for rapid change in the state of a system in response to small external influences.

• *Principle 5. Change in social psychological systems can refer to changes in state or changes in attractors*. External factors that change the momentary state of the system are likely to have only a transient effect if they move the system within the basin of attraction of the system's current attractor. They have more lasting but nonetheless reversible effects if they move the system to the basin of a different attractor. Lasting and irreversible changes in a system reflect changes in the type, number, and position of the system's attractors. Such changes are achieved by changing the control parameters of the system.

• *Principle 6. Personal causation can be understood as patterns of intrinsic dynamics and constraints on the effects of external influences*. As a first approximation, personal dispositions can be couched in terms of attractor dynamics. Personal causation is reflected in the rejection of influences that would move the system out of its attractor and in the system's tendency to return to the attractor if it is perturbed. Processes of self-regulation result in the creation of attractors (or repellors) for system dynamics.

• *Principle 7. Structure and order at the system level can result from interactions among elements at a lower level in a process referred to as self-organization*. No supervision or control by a higher-order structure is necessary for this to occur. Self-organization underlies the process of emergence in social psychological phenomena at different levels of personal and social reality. Emergence can only occur in a dynamical system in which elements interact over time. Systems characterized by nonlinear relations among elements are especially likely to demonstrate emergence. The principle of self-organization provides for coherence in social psychological processes.

• *Principle 8. The approach of dynamical minimalism is essential for building and testing dynamical theories*. The goal of dynamical minimalism is the construction of the simplest possible theories of complex phenomena. Very simple elements and simple rules of interaction among elements can result in very complex properties at the system level. This can happen only if the elements interact in time. Many specific features of elements and their interactions do not have important consequences at the system level and thus can be ignored in the description of elements. Computer simulations provide the primary tool for investigating the emergent properties of social psychological systems and for determining which features of the elements must be included in a minimalist account of a phenomenon. They also provide a tool for investigating the relation between the structure and dynamics of the system. Empirical research is necessary for verification of the model. Because of the concentration on basic principles, dynamical minimalism has the potential for building models that are general across otherwise diverse phenomena in social psychology.

THE TRAJECTORY AHEAD

Theory construction in social psychology can itself be viewed in dynamical terms, with individual researchers influencing one another over time in an attempt to achieve consensus on the nature of human experience. Despite the progressive coherence that emerges over time by virtue of self-organization, complex systems rarely attain complete integration. Hence, one should not expect a discipline as diverse and multifaceted as social psychology to reach a stable equilibrium, with a single set of immutable principles capturing all the nuances of personal and interpersonal function. Complex systems are inherently dynamic, continually evolving, and undergoing transformations by virtue of their intrinsic dynamics and in response to incoming information and outside influences. At this point in time, then, the promise of theoretical coherence is simply that—an optimistic extension of a current temporal trajectory. So although we anticipate that dynamical principles and methods will emerge as the paradigmatic foundation for social psychological science, we also anticipate that the field will display repeated episodes of disassembly and reconfiguration with respect to specific theories and research strategies in the years to come. This trajectory provides ironic testament to the viability and generality of the dynamical perspective on human experience.

NOTES

1. For depictions of specific research agendas spawned by this approach, the interested reader is referred to a special issue of *Personality and Social Psychology Review* (Vallacher, Read, & Nowak, 2002) devoted to this topic.
2. See Nowak et al. (2002) for a detailed discussion of these and other properties of fixed-point attractors.
3. This scenario is exemplified by the rapid improvement in Polish–Ukrainian relations after the 2004 "orange revolution" in Ukraine (see Nowak et al., in press).
4. The dynamics of a system may be governed by the superimposition of different periods. This is referred to as *multiperiodic evolution*. Patterns of human activity, for example, vary on a daily basis (e.g., wake and sleep), on a weekly basis (e.g., weekends vs. weekdays), and on an annual basis (e.g., seasonal activities). At each moment, a person's state is determined by its position in each of these periods. His or her activity thus depends on the time of day (e.g., lunch at noon), the day of the week (e.g., Saturday), and the time of year (e.g., Spring). The most complex case of multiperiodicity is quasi-periodic evolution, in which each of the system's variables exhibit different periods and these periods are not multiples of a single basic period. The temporal dependence of a dynamical variable that displays quasiperiodic motion might be very complicated and difficult to identify, as the superimposition of waves can create very irregular and complex shapes (cf. Nowak & Lewenstein, 1994).

REFERENCES

Abelson, R. P. (1979). Social clusters and opinion clusters. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives in social network research* (pp. 239–256). New York: Academic Press.

Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems*. Thousand Oaks, CA: Sage.

Asch, S. E. (1946). Forming impressions of personalities. *Journal of Abnormal and Social Psychology*, *41*, 258–290.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *19*, 31–35.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Baron, R. M., Amazeen, P. G., & Beek, P. J. (1994). Local and global dynamics of social relations. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 111–138). San Diego, CA: Academic Press.

Barton, S. (1994). Chaos, self-organization, and psychology. *American Psychologist*, *49*, 5–14.

Baumeister, R. F., Smart, L., & Bowden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of self-esteem. *Psychological Review*, *103*, 5–33.

Baumgardner, A. H. (1990). To know oneself is to like oneself: Self-certainty and self-affect. *Journal of Personality and Social Psychology*, *58*, 1062–1072.

Beek, P. J., & Hopkins, B. (1992). Four requirements for a dynamical systems approach to the development of social coordination. *Human Movement Science*, *11*, 425–442.

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York: Academic Press.

Brown, K. W., & Moskowitz, D. S. (1998). Dynamic stability of behavior: The rhythms of our interpersonal lives. *Journal of Personality*, *66*, 105–134.

Buder, E. H. (1991). A nonlinear dynamic model of social interaction. *Communication Research*, *18*, 174–198.

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological Inquiry*, *6*, 1–30.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*, 3–25.

Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavallee, L. F., & Lehman, D. R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology*, *70*, 141–156.

Cannon, W. B. (1932). *The wisdom of the body*. New York: Norton.

Carver, C. S., & Scheier, M. F. (1999). Themes and issues in the self-regulation of behavior. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 12, pp. 1–105). Mahwah, NJ: Erlbaum.

Carver, C. S., & Scheier, M. F. (2002). Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review*, *6*, 304–315.

Cervone, D. (2004). The architecture of personality. *Psychological Review*, *111*, 183–204.

Coleman, P. T. (2003). Characteristics of protracted, intractable conflict: Towards the development of a meta-framework—I. First paper in a three-paper series. *Peace and Conflict: Journal of Peace Psychology*, *9*, 1–37.

Coleman, P. T., Vallacher, R. R., Nowak, A., & Bui-Wrzosinska, L. (in press). Malignant social relations as an attractor: Presenting a dynamical model of seemingly intractable conflict. *American Behavioral Scientist*.

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research*, *5*, 221–235.

Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Davis, B. P., & Knowles, E. S. (1999). A disrupt-then-reframe technique of social influence. *Journal of Personality and Social Psychology*, *76*, 192–199.

Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven, CT: Yale University Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Dittman, A. T., & Llewellyn, L. G. (1969). Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology*, *11*, 98–106.

Durkheim, E. (1938). *The rules of sociological method*. Chicago: University of Chicago Press.

Eckmann, J. P., & Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Review of Modern Physics*, *57*, 617–656.

Eiser, J. R. (1994). *Attitudes, chaos, and the connectionist mind*. Oxford, UK: Blackwell.

Feigenbaum, M. J. (1978). Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, *19*, 25–52.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups*. Stanford, CA: Stanford University Press.

Fischer, K. W., & Bidell, T. R. (1997). Dynamic development of psychological structures in action and thought. In W. Damon (Series Ed.) & R. Lerner (Vol. Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 313–399). New York: Wiley.

Glass, L., & Mackey, M. C. (1988). *From clocks to chaos: The rhythms of life*. Princeton, NJ: Princeton University Press.

Gleick, J. (1987). *Chaos: The making of a new science*. New York: Viking-Penguin.

Goldstein, J. (1996). Causality and emergence in chaos and complexity theories. In W. Sulis & A. Combs (Eds.), *Nonlinear dynamics and human behavior* (pp. 161–190). Singapore: World Scientific.

Gottman, J. M. (1979). Detecting cyclicity in social interaction. *Psychological Bulletin*, *86*, 338–348.

Gottman, J. M., Murray, J., Swanson, C., & Tyson, R. (2004). *A general systems theory of marriage*. Cambridge, MA: MIT Press.

Gottman, J., Swanson, C., & Swanson, K. (2002). A general systems theory of marriage: Nonlinear difference equation modeling of marital interaction. *Personality and Social Psychology Review*, *6*, 326–340.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellot, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Journal of Personality and Social Psychology*, *109*, 3–25.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Guastello, S. J. (1995). *Chaos, catastrophe, and human affairs: Applications of nonlinear dynamics to work, organizations, and social evolution*. Mahwah, NJ: Erlbaum.

Haken, H. (1978). *Synergetics*. Berlin: Springer.

Hastie, R., Penrod, S. D., & Pennington, N. (1983). *Inside the jury*. Cambridge, MA: Harvard University Press.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Higgins, E. T. (1996). Ideals, oughts, and regulatory outcome focus: Relating affect and motivation to distinct pains and pleasures. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 91–114). New York: Guilford Press.

Higgins, E. T. (2000). Making a good decision: Value from "fit." *American Psychologist*, *55*, 1217–1230.

Holland, J. H. (1995). *Emergence: From chaos to order*. Reading, MA: Addison-Wesley.

Hopfield, J. J. (1982). Neural networks and physical systems with

emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA, 79*, 2554–2558.

Hsee, C. K, Abelson, R. P., & Salovey, P. (1991). The relative weighting of position and velocity in satisfaction. *Psychological Science, 2*, 263–266.

James, W. (1890). *Principles of psychology*. New York: Holt.

Janis, I. L. (1982). *Victims of groupthink* (2nd ed.). Boston: Houghton Mifflin.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.

Johnson, S. L., & Nowak, A. (2002). Dynamical patterns in bipolar depression. *Personality and Social Psychology Review, 6*, 380–387.

Kaplowitz, S. A., & Fink, E. L. (1992). Dynamics of attitude change. In R. L. Levine & H. E. Fitzgerald (Eds.), *Analysis of dynamic psychological systems* (Vol. 2, pp. 341–369). New York: Plenum Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kenrick, D. T., Li, N. P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision-rules and emergent social norms. *Psychological Review, 110*, 3–28.

Kernis, M. H. (1993). The roles of stability and level of self-esteem in psychological functioning. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 167–182). New York: Plenum Press.

Knowles, E. S., & Linn, J. A. (2004). Approach–avoidance model of persuasion: Alpha and omega strategies for change. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 117–148). Mahwah, NJ: Erlbaum.

Krueger, J. I., & Funder, D. C. (2004). Toward a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27*, 313–327.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review, 103*, 263–283.

Kunda, Z. (1999). *Social cognition*. Cambridge, MA: MIT Press

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*, 284–308.

Larsen, R. J. (1987). The stability of mood variability: A spectral analytic approach to daily mood assessments. *Journal of Personality and Social Psychology, 52*, 1195–1204.

Larsen, R. J., & Kasimatis, M. (1990). Individual differences in entrainment of mood to the weekly calendar. *Journal of Personality and Social Psychology, 58*, 164–171.

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*, 343–356.

Latané, B., & Bourgeois, M. J. (1996). Experimental evidence for dynamic social impact: The formations of subcultures in electronic groups. *Journal of Communication, 46*, 35–47.

Latané, B., & Nowak, A. (1994). Attitudes as catastrophes: From dimensions to categories with increasing involvement. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 219–249). San Diego, CA: Academic Press.

Latané, B., & Nowak, A. (1997). The causes of polarization and clustering in social groups. *Progress in Communication Sciences, 13*, 43–75.

Latané, B., Nowak, A., & Liu, J. (1994). Measuring emergent social phenomena: dynamism, polarization and clustering as order parameters of social systems. *Behavioral Science, 39*, 1–24.

Lewenstein, M., & Nowak, A. (1989a). Fully connected neural networks with self-control of noise levels. *Physics Review Letter, 62*, 225–229.

Lewenstein, M., & Nowak, A. (1989b). Recognition with self-control in neural networks. *Physical Review A, 40*, 4652–4664.

Lewenstein, M., Nowak, A., & Latané, B. (1993). Statistical mechanics of social impact. *Physics Review A, 45*, 703–716.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Lewis, M. (1997). Personality self-organization: Cascading constraints on cognition-emotion interactions. In A. Fogel, M. C. Lyra, & J. Valsiner (Eds.), *Dynamics and indeterminism in developmental and social processes* (pp. 193–216). Mahwah, NJ: Erlbaum.

Lorenz, E. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Science, 20*, 282–293.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808–817.

Mandler, G. (1975). *Mind and emotion*. New York: Wiley.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*, 954–969.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175–216.

McGrath, J. E., & Kelley, J. R. (1986). *Time and human interaction: Toward a psychology of time*. New York: Guilford Press.

Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Messick, D. M., & Liebrand, V. B. G. (1995). Individual heuristics and the dynamics of cooperation in large groups. *Psychological Review, 102*, 131–145.

Miller, N. E. (1944). Experimental studies of conflict. In J. M. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald Press.

Mischel, W., & Shoda, Y. (1995). A cognitive–affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.

Moscovici, S., Lage, E., & Naffrechoux, M. (1969). Influence of a consistent minority on the responses of a majority in a color perception task. *Sociometry, 32*, 365–380.

Myers, D. G., & Lamm, H. (1976). The group polarization hypothesis. *Psychological Bulletin, 83*, 602–627.

Newtson, D. (1994). The perception and coupling of behavior waves. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 139–167). San Diego, CA: Academic Press.

Nezlek, J. B. (1993). The stability of social interaction. *Journal of Personality and Social Psychology, 65*, 930–941.

Nowak, A. (2004). Dynamical minimalism: Why less is more in psychology. *Personality and Social Psychology Review, 8*, 183–192.

Nowak, A., Latané, B., & Lewenstein, M. (1994). Social dilemmas exist in space. In U. Schulz, W. Albers, & U. Mueller (Eds.), *Social dilemmas and cooperation* (pp. 114–131). Heidelberg: Springer-Verlag.

Nowak, A., & Lewenstein, M. (1994). Dynamical systems: A tool for social psychology? In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 17–53). San Diego, CA: Academic Press.

Nowak, A., Lewenstein, M., & Frejlak, P. (1996). Dynamics of public opinion and social change. In R. Hegselman & H. O. Pietgen (Eds.), *Modeling social dynamics: Order, chaos, and complexity* (pp. 54–78). Vienna: Helbin.

Nowak, A., Lewenstein, M., & Tarkowski, W. (1994). Repellor neural networks. *Physical Review E, 48*, 1491–1498.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review, 97*, 362–376.

Nowak, A., & Vallacher, R. R. (1998a). *Dynamical social psychology*. New York: Guilford Press.

Nowak, A., & Vallacher, R. R. (1998b). Toward computational social psychology: Cellular automata and neural network models of interpersonal dynamics. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 277–311). Mahwah, NJ: Erlbaum.

Nowak, A., & Vallacher, R. R. (2001). Societal transition: Toward a dynamical model of social change. In W. Wosinska, R. B. Cialdini, D. W. Barrett, & J. Reykowski (Eds.), *The practice of social influence in multiple cultures* (pp. 151–171). Mahwah, NJ: Erlbaum.

Nowak, A., Vallacher, R. R., Bui-Wrzosinska, L., & Coleman, P. T. (in press). Attracted to conflict: A dynamical perspective on malignant social relations. In A. Golec & K. Skarzynska (Eds.),

Understanding social change: Political psychology in Poland. Haauppague, NY: Nova Science.

Nowak, A., Vallacher, R. R., Kus, M., & Urbaniak, J. (2005). The dynamics of societal transition: Modeling non-linear change in the Polish economic system. *International Journal of Sociology*, *35*, 66–88.

Nowak, A., Vallacher, R. R., Tesser, A., & Borkowski, W. (2000). Society of self: The emergence of collective properties in self-structure. *Psychological Review*, *107*, 39–61.

Nowak, A., Vallacher, R. R., & Zochowski, M. (2002). The emergence of personality: Personality stability through interpersonal synchronization. In D. Cervone & W. Mischel (Eds.), *Advances in personality science* (Vol. 1, pp. 292–331). New York: Guilford Press.

Pelham, B. W. (1991). On confidence and consequences: The certainty and importance of self-knowledge. *Journal of Personality and Social Psychology*, *60*, 518–530.

Penner, L. A., Shiffman, S., Paty, A., & Fritzche, B. A. (1994). Individual differences in intraperson variability in mood. *Journal of Personality and Social Psychology*, *66*, 712–721.

Poincaré, H. (1908). *Science and method.* New York: Dover. (English translation, 1952)

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention grabbing power of negative information. *Journal of Personality and Social Psychology*, *61*, 380–391.

Queller, S. (2002). Stereotype change in a recurrent network. *Personality and Social Psychology Review*, *6*, 295–303.

Read, S. J., & Miller, L. C. (Eds.). (1998). *Connectionist models of social reasoning and social behavior.* Mahwah, NJ: Erlbaum.

Read, S. J., & Miller, L. C. (2002). Virtual personalities: A neural network model of personality. *Personality and Social Psychology Review*, *6*, 357–369.

Richards, D. (1990). Is strategic decision making chaotic? *Behavioral Science*, *35*, 219–232.

Ruelle, D. (1989). *Elements of differentiable dynamics and bifurcation theory.* New York: Academic Press.

Schachter, S. (1951). Deviation, rejection and communication. *Journal of Abnormal and Social Psychology*, *46*, 199–207.

Schuster, H. G. (1984). *Deterministic chaos.* Vienna: Physik Verlag.

Sherif, M. (1936). *The psychology of social norms.* New York: Harper.

Sherman, J. W., Stroessner, S. J., Loftus, S. T., & Deguzman, G. (1997). Stereotype suppression and recognition memory for stereotypical and nonstereotypical information. *Social Cognition*, *15*, 205–215.

Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra- and interpersonal dynamics. *Personality and Social Psychology Review*, *6*, 316–325.

Showers, C. J. (1992). Compartmentalization of positive and negative self-knowledge: Keeping bad apples out of the bunch. *Journal of Personality and Social Psychology*, *62*, 1036–1049.

Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, *103*, 219–240.

Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology Review*, *6*, 283–294.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29–39.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131–142.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology*, *70*, 893–912.

Story, A. L. (2004). Self-esteem and self-certainty: A mediational analysis. *European Journal of Personality*, *18*, 115–125.

Strogatz, S. (2003). *Sync: The emerging science of spontaneous order.* New York: Hyperion Books.

Swann, W. B., Jr. (1990). To be adored or to be known? The interplay of self-enhancement and self-verification. In E. T. Higgins &

R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Volume 2. Foundations of social behavior* (pp. 408–448). New York: Guilford Press.

Swann, W. B., Jr., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology*, *46*, 1287–1302.

Swann, W. B., Jr., Hixon, J. G., Stein-Seroussi, A., & Gilbert, D. (1990). The fleeting gleam of praise: Cognitive processes underlying behavioral reactions to self-relevant feedback. *Journal of Personality and Social Psychology*, *59*, 17–26.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 85–117). New York: Academic Press.

Tesser, A., & Achee, J. (1994). Aggression, love, conformity, and other social psychological catastrophes. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 96–109). San Diego, CA: Academic Press.

Tesser, A., Martin, L. L., & Cornell, D. P. (1996). On the substitutability of self-protective mechanisms. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 48–68). New York: Guilford Press.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*, 435–467.

Thagard, P., & Nerb, J. (2002). Emotional gestalts: Appraisal, change, and the dynamics of affect. *Personality and Social Psychology Review*, *6*, 274–282.

Thelen, E., & Smith, L. B. (Eds.). (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press/Bradford Books.

Thom, R. (1975). *Structural stability and morphogenesis.* New York: Addison-Wesley.

Tickle-Degnen, L., & Rosenthal, R. (1987). Group rapport and nonverbal behavior. *Review of Personality and Social Psychology*, *9*, 113–136.

Townsend, J. T., & Busemeyer, J. (1995). Dynamic representation of decision-making. In R. F. Port & T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 101–120). Cambridge, MA: MIT Press.

Turner, R. H., & Killian, L. M. (1957). *Collective behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Turvey, M. T. (1990). Coordination. *American Psychologist*, *4*, 938–953.

Turvey, M. T., & Carello, J. (1995). Some dynamical themes in perception and action. In R. F. Port & T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 373–402). Cambridge, MA: MIT Press.

Vallacher, R. R. (1978). Objective self-awareness and the perception of others. *Personality and Social Psychology Bulletin*, *4*, 63–67.

Vallacher, R. R. (1980). An introduction to self theory. In D. M. Wegner & R. R. Vallacher (Eds.), *The self in social psychology* (pp. 3–30). New York: Oxford University Press.

Vallacher, R. R., & Nowak, A. (1994a). The chaos in social psychology. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 1–16). San Diego, CA: Academic Press.

Vallacher, R. R., & Nowak, A. (1994b). The stream of social judgment. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 251–277). San Diego, CA: Academic Press.

Vallacher, R. R., & Nowak, A. (1997). The emergence of dynamical social psychology. *Psychological Inquiry*, *4*, 73–99.

Vallacher, R. R., & Nowak, A. (1999). The dynamics of self-regulation. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 12, pp. 241–259). Mahwah, NJ: Erlbaum.

Vallacher, R. R., & Nowak, A. (2000). Landscapes of self-reflection: Mapping the peaks and valleys of personal assessment. In A. Tesser, R. Felson, & J. Suls (Eds.), *Psychological perspectives on self and identity* (pp. 35–65). Washington, DC: American Psychological Association.

Vallacher, R. R., Nowak, A., Froehlich, M., & Rockloff, M. (2002b). The dynamics of self-evaluation. *Personality and Social Psychology Review*, *6*, 370–379.

Vallacher, R. R., Nowak, A., & Kaufman, J. (1994). Intrinsic dynamics of social judgment. *Journal of Personality and Social Psychology*, *66*, 20–34.

Vallacher, R. R., Nowak, A., Markus, J., & Strauss, J. (1998). Dynamics in the coordination of mind and action. In M. Kofta, G. Weary, & G. Sedlek (Eds.), *Personal control in action: Cognitive and motivational mechanisms* (pp. 27–59). New York: Plenum Press.

Vallacher, R. R., Nowak, A., & Miller, M. E. (2002). Social influence and group dynamics. In I. Weiner (Series Ed.) & T. Millon & M. J. Lerner (Vol. Eds.), *Handbook of psychology: Vol. 5. Personality and social psychology* (pp. 383–417). New York: Wiley.

Vallacher, R. R., Read, S. J., & Nowak, A. (Eds.). (2002). The dynamical perspective in social psychology. *Personality and Social Psychology Review* [Special issue], *6*.

Vallacher, R. R., & Wegner, D. M. (1985). *A theory of action identification*. Hillsdale, NJ: Erlbaum.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*, 3–15.

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*, 660–671.

Vallacher, R. R., Wegner, D. M., & Somoza, M. P. (1989). That's easy for you to say: Action identification and speech fluency. *Journal of Personality and Social Psychology*, *56*, 199–208.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.

Wegner, D. M., Vallacher, R. R., Kiersted, G., & Dizadji, D. (1986). Action identification in the emergence of social behavior. *Social Cognition*, *4*, 18–38.

Weidlich, W. (1991). Physics and social science: The approach of synergetics. *Physics Reports*, *204*, 1–163.

Weisbuch, G. (1992). *Complex systems dynamics*. Redwood City, CA: Addison-Wesley.

Wilson, T. D., Wheatley, T., Meyers, J. M., Gilbert, D. T., & Axsom, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *78*, 821–836.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media.

Zimbardo, P. G. (1970). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation, 1969* (Vol. 17, pp. 237–307). Lincoln: University of Nebraska.

Inclusion and Exclusion

Implications for Group Processes

JOHN M. LEVINE
NORBERT L. KERR

Many efforts have been made to elucidate the fundamental motives underlying human emotion, thought, and behavior. Recently, Fiske (2004) identified five core social motives that enhance people's integration into groups and thereby their physical and psychological survival. One of these motives—belonging—is "first among equals" because it serves as the foundation for the remaining four (understanding, controlling, self-enhancing, and trusting). Fiske defines the motive to belong as the need for strong, stable relationships with other people (cf. Baumeister & Leary, 1995). She notes that in addition to enhancing individual survival, the motive to belong also benefits the group by increasing the likelihood that members will cooperate and engage in coordinated action.

Although belonging has been recognized as an important human motive for decades, there has been an explosion of recent theoretical and empirical work on this topic, much of it stimulated by Baumeister and Leary's (1995) influential paper. These authors argued that there is a fundamental need to belong that requires frequent positive interactions with one or more others in the context of stable, supportive relationships.

Current interest in the need to belong is high, as indicated by several recent conferences and edited volumes on inclusion/exclusion (e.g., Abrams, Hogg, & Marques, 2005b; Williams, Forgas, & von Hippel, 2005), as well as a large number of journal articles, chapters, and books focusing on this topic. Researchers from several disciplines, including psychology (social, developmental, per-

sonality), biology, ethology, and economics, are addressing the origins of the need to belong and the consequences of striving to satisfy this need. These issues include, among others, individual differences in sensitivity to rejection (Romero-Canyas & Downey, 2005); causes and consequences of familial rejection (Fitness, 2005); coping strategies for stigma-based exclusion (Major & Eccleston, 2005); prosocial and antisocial reactions to ostracism (Williams & Govan, 2005); impact of exclusion on aggression and cognitive activity (Twenge & Baumeister, 2005); bullying as a form of exclusion (Juvonen & Gross, 2005); self-regulation in coping with exclusion (Pickett & Gardner, 2005); inclusion and exclusion as means of managing group composition (Levine, Moreland, & Hausmann, 2005); consequences of marginally fitting the group's prototype (Hogg, 2005); and threat of exclusion as inducement to cooperation (Ouwerkerk, Kerr, Gallucci, & van Lange, 2005).

Although sizable, the contemporary literature on inclusion/exclusion is limited in at least two respects. One limitation concerns the tendency to favor some levels of analysis over others. Most of the work on inclusion/exclusion has involved either the intrapersonal, the interpersonal, or the intergroup level of analysis—that is, either the individual's response to inclusionary/exclusionary opportunities/threats, inclusionary/exclusionary relations between members of dyads, or inclusionary/exclusionary relations between members of different groups that are more or less hostile to one another. In contrast, the intragroup level of analysis has

been neglected, with relatively little attention given to how social processes within groups influence and are influenced by inclusion/exclusion.

The second limitation concerns the tendency to focus on one side of the inclusion/exclusion relationship (the target or the source) while neglecting the other. This is particularly true in the case of dyads. Research on dyads has sought to explain why relationship partners desire inclusion and how they respond to exclusion, with much less attention to how and why relationship partners include or exclude one another. Until recently, work on inclusion/exclusion in intergroup contexts had precisely the opposite bias. Research on intergroup relations assumed (explicitly or implicitly) that problems between groups are due to strong majorities excluding weak minorities, and hence the key to increasing intergroup harmony is identifying the origins of this exclusion and designing ways to transform it into inclusion. Not surprisingly, this perspective stimulated much more attention to the source of exclusion (the majority) than to its target (the minority). In the last few years, however, increased research has focused on how minorities cope with the exclusion they encounter (e.g., Crocker & Quinn, 2000; Major & Eccleston, 2005).

The central premise of this chapter is that inclusion/exclusion underlies a large and disparate set of small group phenomena. In the following pages, we review and extend prior analyses of inclusion/exclusion with the goal of clarifying the dynamics of small, task-oriented groups. Though both individuals and groups can—at least in certain senses—be sources as well as targets of inclusion/exclusion, we focus on the more common case in which the group is the source and the individual (member) is the target. This focus reflects the common asymmetry in power between the two parties, in which the group typically has more power to include/exclude than does the individual, and the individual typically is more sensitive to being included/excluded than is the group (Abrams, Hogg, & Marques, 2005a).

THE GROUP MEMBER AS THE TARGET OF INCLUSION/EXCLUSION

Humans strive for inclusion in small groups and, once included, seek to avoid exclusion. This assertion clearly reflects—but implies more than—the *need to belong* proposed by Baumeister and Leary (1995). They present a reasoned and persuasive argument that the need to belong is a fundamental human need. But they also suggest that this need is not simply for any kind of positive bond with others (like the positive unit relations of balance theory) but, rather, requires interpersonal bonds of both high quantity and high quality. That is, infrequent contacts or casual associations with others are not sufficient. There are indications that, for most of us, the need to belong can largely be satisfied through a few intimate personal relationships (Baumeister & Leary, 1995, pp. 513–515). However, this need can also be satisfied in other contexts. In particular, membership or inclusion in small

groups often provides interpersonal contacts that are both frequent and close enough (e.g., as in families) to satisfy the need to belong (Hornsey & Jetten, 2004). Interestingly, however, people also routinely worry about their inclusion/exclusion in groups that provide only infrequent contact (e.g., a high school graduating class) or relatively weak interpersonal bonds (e.g., an informal discussion group). In this chapter, our analysis of the individual's drive to obtain group inclusion (and avoid group exclusion) will thus incorporate both the need to belong, as defined by Baumeister and Leary, and other (perhaps less urgent) motivations to maintain group memberships.

Roots of the Drive for Group Inclusion

Why is being and remaining a group member so important to us? There are three general answers to this question, which should be viewed as complementary rather than mutually exclusive.

The Drive for Group Inclusion as a Residue of Evolution

It has become not only acceptable but fashionable in social psychology to suggest that patterns of human social behavior are, at least to some degree, the results of long-term evolutionary processes (Buss & Kenrick, 1998; Caporael, 2001; Neuberg, Smith, & Asher, 2000). In that vein, a number of scholars (e.g., Barchas, 1986; Baumeister & Leary, 1995; Caporael & Baron, 1997; Kurzban & Leary, 2001) have suggested that propensities to seek, maintain, and value group memberships have been strongly selected for during human evolution. The evidence for such claims—as for many assertions of evolutionary psychology—is by no means conclusive, but these claims remain intriguing nonetheless. Several arguments can be made for the role of evolution in the drive for group inclusion.

One argument stresses the likely adaptive value of group memberships for our early ancestors. A number of essential reproductive tasks (e.g., finding mates, caring for young) and survival tasks (e.g., hunting, defense, winning and holding territories) can be pursued more effectively by groups than by individuals, so it makes sense that a strong tendency to join and remain in groups was selected for during evolution. Other arguments are based on anthropological and comparative evidence. For example, there are no fossil records suggesting that *homo sapiens* ever lived other than in groups (Caporael, Dawes, Orbell, & Van de Kragt, 1989). Moreover, the fact that our closest hominid relatives (e.g., gorillas and chimpanzees) live in groups suggests that this preference for collective social organization may have predated the beginnings of our species. The apparent universality of collective social organization in modern (as well as primitive) human societies is likewise consistent with an innate drive to affiliate. Finally, there is evidence linking biological mechanisms to social affiliation and disaffiliation (e.g., opioid production and social bonding, Panksepp, Siviy, & Normansell, 1985; pain center activation in the brain following social ostracism, Eisenberger,

Lieberman, & Williams, 2003; MacDonald & Leary, 2005; see also Dickerson & Kemeny, 2004).

An interesting, although indirect, argument involves the relationship between the evolution of sociality (the tendency for members of a species to live, work, and associate with one another) and the existence and regulation of group boundaries. A long-standing puzzle for those interested in explaining human cooperation is how sociality evolved when so many environments have the structure of *social dilemmas* (cf. Orbell & Dawes, 1981), in which pursuing immediate self-interest and survival (e.g., not sharing one's catch with others, taking more than one's share of a dwindling resource) is inimical to collective survival (e.g., keeping the whole clan fed, stretching scarce resources as long as possible). A number of theoretical models (e.g., Boyd, Gintis, Bowles, & Richerson, 2003; Boyd & Richerson, 1992; Hayashi & Yamagishi, 1998; Kameda, Takezawa, & Hastie, 2003) suggest that the evolution of cooperative behavior is facilitated by creating and maintaining group boundaries and tying cooperative behavior to group membership. For example, a strategy that involves sticking with partners who cooperated in the past but abandoning partners who failed to cooperate does better than many alternative strategies in a selective-play environment, where individuals can choose with whom to affiliate (de Vos & Zeggelink, 1997; Yamagishi, Hayashi, & Jin, 1994). It has also been suggested that intragroup cooperation emerges more readily through group selection processes (1) when there are competing groups in the environment than when there are not (Boyd et al., 2003) and (2) when group members recognize and punish those who act uncooperatively than when they do not (Boyd et al., 2003; Kameda et al., 2003; Price, Cosmides, & Tooby, 2002). The ample evidence that humans are often cooperative (e.g., Dawes, Van de Kragt, & Orbell, 1990; Marwell & Ames, 1979; Van Vugt & Van Lange, 2006) supports (at least indirectly) the arguments that the collective value of cooperative behavior created selective pressures for humans to seek group membership, to maintain group boundaries, to value the welfare of fellow group members, to feel obligated to reciprocate cooperation, to be attentive to those who fail to cooperate, and to be highly punitive toward such defectors. As this analysis suggests, group membership has costs as well as rewards, and the need for inclusion thus stems, at least in part, from the competitive demands of group life.

The Drive for Group Inclusion as a Residue of Early Experience

The role of evolution in the drive for inclusion was also posited by Bowlby (1969), who argued for the existence of an evolutionarily adaptive attachment system that keeps infants and their caregivers in close proximity to one another. According to Bowlby, early attachment experiences are important because they affect children's perceptions of themselves (as worthy or unworthy) and others (as dependable or undependable), which in turn influence how they think, feel, and act toward others.

In recent years, increasing attention has been given to adult attachment in the context of romantic relationships and groups (Mikulincer & Shaver, in press). For example, Hazan and Shaver (1987) identified three adult attachment styles (secure, anxious–ambivalent, and avoidant) that predict people's thoughts, feelings, and behaviors in romantic relationships (Simpson & Rholes, 1998). Moreover, Smith, Murphy, and Coats (1999) used adult attachment theory to explain people's psychological ties to groups. Arguing that attachment to groups, like attachment to caregivers, has a strong evolutionary basis, Smith and his colleagues developed a questionnaire measure of group attachment with a two-dimensional structure: (1) Attachment Anxiety (how much a person fears group rejection) and (2) Avoidance (how much a person wants and values group closeness). They then demonstrated that this measure predicted several important group outcomes, including group-relevant emotions, perceived social support, and plans to leave.

Regardless of the utility of attachment theory as an explanation for the impact of early experience on reaction to later inclusion/exclusion, it is clear that social isolation and parental neglect are aversive and potentially lethal to children (cf. Bowlby, 1980; Harlow, Harlow, & Suomi, 1971). If the presence/absence of nurturant social contact is powerfully rewarding/punishing to children, then any stimuli consistently associated with such contact should acquire secondary reinforcing properties. Such conditioning could create strong connections between narrow but powerful needs for social attachment in early life and a variety of common social stimuli or situations. For example, if a parent were to withdraw attention and nurturance whenever a child was disobedient, then the act (or even the prospect) of disobedience might well acquire aversive properties, which in turn would lead to anxiety and to actions designed to reduce (or avoid) this unpleasant state (Baumeister & Tice, 1990). Through secondary conditioning processes, the act (or prospect) of disobedience in group settings may elicit very similar reactions. Recent research (Abrams, de Moura, Frings, & Rutland, 2005) nicely documents developmental trends in awareness of such relevant social stimuli and situations (e.g., social norms).

The Drive for Group Inclusion as a "Rational" Choice

The forgoing discussion suggests that the drive for group inclusion is based on primitive (unconscious) mechanisms, such as evolution or conditioning. However, inclusion can be important to group members for another reason as well, namely, because of a calculated decision that it facilitates their ability to achieve specific individual goals. Any number of such goals can only be achieved in group settings (e.g., winning a basketball championship, getting a team leadership award). Clearly, the more a person values such goals, the more he or she would want to obtain and maintain group membership.

One ubiquitous individual goal is obtaining social status or approval, and group inclusion is a clear indicator that one enjoys the esteem of fellow group members. Moreover, when standards of social evaluation link high

status to particular or multiple group affiliations (e.g., belonging to a particular fraternity or sorority; belonging to a large number of campus groups), such memberships can elicit esteem from those outside as well as inside the group. Consistent with this idea, Leary and Baumeister (2000) have argued in their *sociometer theory* that an individual's level of self-esteem is a subjective monitor of how much he or she is socially valued. From this perspective, the more functional group memberships are for obtaining positive social evaluations, the more important group inclusion/exclusion should be for self-regard.

Another perspective on the relationship between group membership and self-esteem involves the self-enhancement function of social comparison (e.g., Goethals & Darley, 1977; Tesser, 1988; Wills, 1991). The fundamental idea here is that the valence and intensity of people's feelings about themselves are affected by how they compare to others on valued dimensions. Although much of this work has focused on the affective consequences of interpersonal comparisons in dyadic settings, other work has dealt with self–other comparisons in group settings or with ingroup–outgroup comparisons (e.g., Levine & Moreland, 1987; Luhtanen & Crocker, 1991). This latter type of comparison has been emphasized by social identity theorists (e.g., Tajfel & Turner, 1986), who argue that people have high self-esteem to the extent that their ingroup is superior to outgroups on valued dimensions (see Abrams & Hogg, 2001). To exploit this potential source of self-esteem, people should seek to establish and maintain memberships in only certain groups, namely, those that are obviously "superior" to other groups. Absent the option for such memberships, people should try to produce ingroup superiority in other ways, such as encouraging intergroup competition that the ingroup is likely to win, perceiving the ingroup as superior to the outgroup on new dimensions, and so on (Jackson, Sullivan, Harnish, & Hodge, 1996; Tajfel & Turner, 1986).

Rather than focusing on the goal of self-enhancement, the original formulation of social comparison theory (Festinger, 1950, 1954) emphasized the goal of self-evaluation. According to Festinger, people want an accurate appraisal of their abilities and opinions, which, in the absence of objective standards, can only be obtained through social comparison (cf. Laughlin & Ellis, 1986). The general notion that people use information obtained from others to validate their perception of reality is widely held in social psychology (see Hardin & Higgins, 1996; Levine & Higgins, 2001; Tindale, Meisenhelder, Dykema-Engblade, & Hogg, 2001). A recent example of this emphasis is the *subjective uncertainty reduction hypothesis*, which posits that people self-categorize as group members in order to reduce uncertainty about the social world and their place in it (Hogg, 2000; Hogg & Abrams, 1993). The goal of reducing uncertainty, or gaining understanding (Fiske, 2004), is thus a powerful individual motive that, in most cases, can only be satisfied through group membership. For this reason, uncertainty reduction, like self-esteem, can be a "rational" reason for seeking group inclusion.

The Detection System for Inclusion/Exclusion

Our discussion so far has focused on why people seek inclusion (and avoid exclusion) in groups. We now turn to the question of how they behave to achieve these goals. Although much of our analysis is informed by Baumeister and Leary's (1995) treatment of the *need to belong* within high contact, caring relationships, our analysis also applies to other motives for group inclusion, such as "rational" efforts to attain individual goals.

According to Baumeister and Leary, the need to belong has several basic features. First, the need involves avoiding exclusion as well as seeking inclusion (although MacDonald & Kingsbury, 2006, have recently suggested that these might be independent motives). Second, one's level of success in fulfilling the need elicits strong emotional and cognitive responses (e.g., positive emotions to satisfaction and negative emotions to frustration). Third, need satisfaction is characterized by both satiation (i.e., satisfaction of the need diminishes motivation for additional relationships) and substitution (i.e., one relationship can replace another). Finally, the need energizes and directs behavior—one plans and acts to fulfill the need.

Leary and Baumeister (2000) suggest that the power of the need to belong has led to the development of an internal system—the *sociometer*–for monitoring satisfaction of this need. Three of the proposed properties of that system are crucial to our current analysis. First, humans are highly sensitive to information about their level of acceptance/rejection by others. Second, people are more sensitive to decrements than to increments of belongingness. "Granted, we feel good when we think we are valued or loved, but most people seem to feel far worse after learning that they are devalued or hated" (Leary & Baumeister, 2000, p. 41) This asymmetry in sensitivity to negative versus positive information appears to be a general phenomenon, with several possible origins (e.g., the survival value of avoiding danger, the relative novelty and diagnosticity of negative events; Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Third, this monitoring system operates automatically (i.e., continuously, involuntarily, and unconsciously). Several recent findings by Williams and his colleagues support these claims (see Williams, 2007). They have shown that quite rudimentary cues of ostracism (e.g., being excluded in a ball toss game) activate the same regions in the brain that are activated for another primitive and automatic monitoring system, namely, physical pain (Eisenberger & Lieberman, 2005; Eisenberger et al., 2003). In addition, they have demonstrated that initial aversive reactions to cues of ostracism are not mediated by ostensibly relevant situational or personality factors (Williams & Zadro, 2005). For example, participants were just as upset by ostracism from members of a disliked group (viz. members of the Ku Klux Klan) as from a control group of liked ingroup members (Gonsalkorale & Williams, in press).

It is important to note that this monitoring system is not only sensitive to current signs of group exclusion but also to settings or behaviors associated with past in-

stances of exclusion. The confederates in Asch's (1956) classic conformity studies did not say or do anything overt to suggest that a deviant (but correct) response would result in exclusion. Nevertheless, participants reported that expressing disagreement with these confederates was uncomfortable—for many, sufficiently uncomfortable to prompt conformity. Thus, rather than being simply reactive, the individual's detection system is also proactive, seeking to identify and avoid settings or behaviors that are "dangerous" because of their association with past exclusion.

Responses to Threats of Exclusion

Assuming that (for whatever reason) an individual perceives that he or she is vulnerable to group exclusion, how might the person react to this threat? Unfortunately, there is no simple answer to this question. Some studies report reflexive behaviors, such as hostile and antisocial reactions to exclusion (see Twenge & Baumeister, 2005). Other studies report more planful behaviors, including acts of "behavioral supplication," such as prosocial behavior, greater effort at group tasks, and higher conformity (see Williams & Govan, 2005).

The latter category of responses is relatively easy to explain. If one's standing in a group is threatened, a rational actor would be expected to choose a course of action that is likely to reduce the threat. A major strategy for doing so is changing the self, for example, by altering the behavior that gave rise to the threat. In addition, an actor might try to change the group (e.g., by persuading members that one's actions are benign or even functional) or leave the group (e.g., by joining another group that is more likely to value one's contributions) (cf. Festinger, 1950). Which of these strategies one uses depends, of course, on a number of factors, including their availability in the situation, their perceived efficacy, one's habitual coping responses, and the attractiveness of alternative group memberships.

The former category of responses, hostility and antisocial behavior, is harder to explain, because it runs the risk of eliciting even more exclusion from the group. Hostile and antisocial responses to the threat of exclusion have been explained in several ways. Williams and Govan (2005), for example, suggest that the threat of exclusion endangers not only the need to belong but other fundamental needs as well (viz., the need to maintain positive self-esteem, the need for control, the existential need for a meaningful existence). From this perspective, behavior that is dysfunctional for reestablishing belongingness (e.g., aggressing against excluders) might well be functional for satisfying other needs (e.g., enhancing one's sense of control). An alternative explanation has been offered by Baumeister and DeWall (2005), who suggest that exclusionary acts or signals void an implicit social contract stipulating that the group will meet one's belongingness needs as long as one regulates one's behavior to satisfy group standards. If one receives feedback from the group suggesting that one's best efforts at self-regulation are not working, one may lose the motiva-

tion to self-regulate and instead act in an impulsive manner. It is also possible, of course, for hostile and antisocial behaviors to be used in a conscious and instrumental fashion to signal that the group may pay a price if it persists in acts of exclusion.

In an effort to explain both "positive" and "negative" responses to threats of exclusion, MacDonald, Kingsbury, and Shaw (2005) offer the interesting argument that responses to threats of exclusion derive from and mimic primitive responses to threats of physical pain (e.g., Gray & McNaughton, 2000). When a physical threat is not immediate and severe, the usual reaction is *fearful avoidance*, which involves attempts at supplication tinged with cautious, defensive distancing. However, when the threat is immediate and severe, *panic* ensues, which in turn prompts one of three basic survival responses—fight (e.g., aggression, verbal insult), flight (e.g., withdrawal, leaving the group), or freezing (lethargic unresponsiveness; cf. Twenge, Catanese, & Baumeister, 2003).

The ideas of MacDonald and colleagues (2005) and Williams and Govan (2005) suggest that when the threat of exclusion is mild and remote and the individual has both the motivation and ability to assess his or her behavioral options, responses to threat will be thoughtful attempts to avoid or reduce exclusion. In contrast, when the threat of exclusion is severe and immediate, impulsive, panic-like reactions may occur, some of which may involve hostility and antisocial behavior. Such reactions may also occur with relatively mild and remote threats, if the individual lacks the motivation and/or ability to engage in a thoughtful analysis of his or her options.

Threat of Exclusion and Group Processes

To this point, we have argued that individuals are acutely sensitive to inclusion and exclusion in group settings. In so doing, we have discussed the factors that cause individuals to desire group inclusion, the system that individuals use to detect inclusion/exclusion, and the responses that individuals make to threats of exclusion. In the following sections, we use the notion of inclusion/exclusion to analyze three important group topics—conformity, group decision making and problem solving, and behavior in mixed-motive settings.

Conformity: The Role of Normative Influence

That individuals desire inclusion and fear exclusion in group settings is consistent with a substantial amount of theoretical and empirical work on majority influence, or conformity. This phenomenon is typically studied by leading a lone participant to believe that several people unanimously disagree with his or her position on a factual or opinion issue and then assessing the extent to which the participant shifts toward the majority position. It has long been argued that conformity derives from two general motives—the desire to reduce uncertainty and the desire to be accepted (Deutsch & Gerard, 1955; Jones & Gerard, 1967; Kelley, 1952; Thibaut & Strickland, 1956). The best-known analysis of these motives was of-

fered by Deutsch and Gerard (1955), who defined informational influence as "influence to accept information obtained from another as *evidence* about reality" and normative influence as "influence to conform with the positive expectations of another" (p. 629). Informational influence is based on the desire to hold accurate beliefs or opinions, coupled with the majority's real or presumed ability to provide valid information. Normative influence is based on the desire to gain or maintain social approval, coupled with the assumption that the majority will respond more favorably to conformity than to deviance. Thus, desire for inclusion and fear of exclusion is a fundamental feature of normative influence.[1]

Consistent with the idea that individuals anticipate negative responses for deviance, evidence indicates that opposition from a majority on a judgment task is a stressful experience and that conformity can reduce this stress (Back & Bogdonoff, 1964; Costell & Leiderman, 1968; Gerard, 1961). In addition, a powerful mechanism for reducing conformity, namely, the presence of a social supporter who agrees with one's position, has been shown to reduce anxiety (Allen & Levine, 1968; Asch, 1952). These data, however, do not provide unequivocal evidence that anxiety in group pressure situations is based on fear of punishment for deviance. This anxiety may be attributable, at least in part, to participants' concern about the validity of their position—a concern that underlies informational (rather than normative) influence.

More convincing evidence is provided by research indicating that people who dissent from group consensus both expect to receive and actually do receive more negative evaluations from other members than do people who conform. Regarding anticipation of punishment, Gerard and Rotter (1961) found that 76% of participants who conformed very little or not at all on a line-judgment task expected negative group evaluation, whereas 71% of participants who conformed a great deal expected positive evaluation (see Jones & Gerard, 1967). Regarding actual receipt of punishment, many studies (reviewed below) demonstrate that people who dissent from group consensus are indeed disliked by other members (Levine, 1989; Marques, Abrams, Paez, & Hogg, 2001). And, after one has been rejected or ostracized by the group, subsequent conformity to the group has been shown to increase (Williams, Cheung, & Choi, 2000).

There is also reason to believe that the effectiveness of social support for attenuating the power of the majority is based, at least in part, on an ally's ability to reduce fear of majority retaliation for deviance. For example, Allen (1975) reported an unpublished study showing that 69% of participants who deviated alone from a group expected to be rejected, whereas only 30% of those who had social support held this expectation. Social support may decrease fear of majority retaliation for several reasons. First, even superficial and transitory relationships with supporters may produce feelings of "inclusion" that innoculate participants against the threat of majority retaliation. Second, participants may believe that their partners will siphon off at least some of the hostility that would otherwise be directed toward them alone. Finally, participants may believe that if they have a partner, the

majority will be more likely to attribute their nonconformity to the validity of their position than to a personal idiosyncrasy, which in turn will reduce the majority's anger at such behavior. Consistent with a normative explanation of social support, evidence indicates that supporters who are disliked by the majority are relatively ineffective in reducing conformity (at least on certain kinds of issues), presumably because they may increase, rather than decrease, the punishment directed toward participants (Boyanowsky & Allen, 1973). Moreover, supporters whose presence is unknown to the majority (and hence give the impression that participants are isolated) have little impact on conformity (Allen, 1975).

To the extent that fear of group punishment underlies conformity, we would expect conformity to be higher under some conditions than others. For example, conformity should be higher when group members are working for a common goal than for individual goals. This is because people in common goal (interdependent) groups should fear that deviance on their part will be seen as a threat to group goal attainment and hence will be severely punished, whereas people in individual goal (noninterdependent) groups should not have these expectations. Consistent with this reasoning, conformity is generally higher in interdependent than in noninterdependent groups (e.g., Deutsch & Gerard, 1955). As might be expected, people in interdependent groups are less likely to conform if their behavior reduces, rather than enhances, the group's probability of success (Sakurai, 1975).

Another factor that should increase normative influence is surveillance by other group members. Because others can only deliver rewards and punishments contingent upon one's behavior if they can somehow monitor this behavior, people ought to be more concerned about others' reactions (and more likely to show normative influence) when their behavior is public rather than private. Consistent with this reasoning, people conform more when their responses are known to other group members than when they are not (e.g., Asch, 1956; Deutsch & Gerard, 1955; Insko, Drenan, Solomon, Smith, & Wade, 1983).

Group attractiveness is also likely to affect normative influence, because both acceptance and rejection should be more potent reinforcers when one cares about the group than when one does not. Although some studies indicate that people conform more to attractive than to unattractive groups (e.g., Festinger, Gerard, Hymovitch, Kelley, & Raven, 1952; Lott & Lott, 1961), this finding is not always obtained (see Allen, 1965; Turner, 1991). Jones and Gerard (1967) argued that the impact of group attractiveness on conformity may depend on such factors as the individual's level of acceptance (i.e., status) in the group, the likelihood that conformity will be instrumental in increasing or maintaining acceptance, and the individual's alternatives outside the group (cf. Dittes & Kelley, 1956; Harvey & Consalvi, 1960; Hollander, 1958). Although the interactive effects of these variables have not been adequately tested, there is suggestive evidence that people conform to an attractive group primarily when they feel insecure about being accepted by the

group and believe that conformity will increase their acceptance (Walker & Heyns, 1962).

In discussing the role of normative influence in conformity, it is important to distinguish between public compliance (overt behavioral change toward the majority's position) and private acceptance (covert attitudinal change toward the majority's position). Although Deutsch and Gerard (1955) did not discuss the relationship between normative/informational influence and compliance/acceptance, subsequent authors have generally assumed that normative influence produces compliance but not acceptance, whereas informational influence produces both compliance and acceptance (e.g., Allen, 1965; Wood, 1999). This assumption is consistent with models of social influence in dyadic relationships proposed by Kelman (1958, 1961) and Raven and his colleagues (e.g., French & Raven, 1959; Raven & Kruglanski, 1970). For example, Raven posited six types of unilateral social influence—reward, coercive, expert, referent, legitimate, and informational. His reward and coercive power (which produce public but not private conformity and require surveillance by the influence source) closely parallel Deutsch and Gerard's normative influence. His informational power (which produces both public and private conformity and requires neither surveillance nor any particular perception of the influence source) closely parallels Deutsch and Gerard's informational influence. Raven's remaining three types of power are less clear-cut in this regard, although legitimate power has normative connotations, whereas expert and referent power have informational ones (cf. Eagly & Chaiken, 1993).

The relationship between normative/informational influence and compliance/acceptance is complicated by several factors. First, although the conceptual distinction between public compliance and private acceptance seems straightforward, methodological problems arise in operationally defining these two types of influence (Allen, 1965). For example, it is often assumed that a person who publicly agreed with a majority at time 1 but privately disagreed at time 2 was showing only compliance at time 1. However, it is possible that acceptance did in fact occur at time 1 but then dissipated, perhaps because of forgetting old information or receiving new information. In contrast, it is often assumed that a person who agreed with a majority at both time 1 and time 2 was showing acceptance as well as compliance at time 1. However, it is possible that the person was only complying at time 1 and subsequently came to believe what he or she said as a function of dissonance reduction or self-perception. Second, the binary distinction between compliance and acceptance greatly oversimplifies how individuals can respond to majority pressure (e.g., Allen, 1965; Nail, 1986; Willis, 1965). Nail (1986), for example, identified eight types of potential influence, based on crossing initial agreement–disagreement, final public congruence–noncongruence, and final private agreement–disagreement. Third, and most important, there are reasons to question the assumption that normative and informational influence produce qualitatively different types of conformity. As Chaiken, Wood, and

Eagly (1996) argued, "social influence theories need to de-couple the motives inducing change from the processes through which change occurs and from the kinds of influence outcomes obtained" (p. 724).

One challenge to traditional notions of normative/informational influence and compliance/acceptance was offered by self-categorization theorists (e.g., Abrams & Hogg, 1990; Turner, 1991; Turner & Oakes, 1989). In brief, they argue that social influence only occurs if a target person perceives that a member of the ingroup disagrees with his or her position and if the source's position represents (is prototypical of) the ingroup norm. This perspective differs from traditional analyses of social influence in several ways, for example, by suggesting that group formation is necessary for influence to occur and that subjective uncertainty is not based on stimulus ambiguity per se but, rather, on the unexpected disagreement with similar others. Most important for our purposes, self-categorization theory denies the distinction between normative and informational influence. Instead, this distinction "is replaced by the idea that the basic influence process is one where the normative position of people categorized as similar to self tends to be subjectively accepted as valid. The validity of information is (psychologically) established by ingroup norms" (Turner, 1991, p. 171). An implication of the self-categorization perspective is that the distinctions between public and private responding and between compliance and acceptance are less relevant for ingroup than for outgroup influence (cf. Abrams & Hogg, 1990).

This effort to merge informational and normative influence, while interesting, is not entirely convincing from our point of view. One reason is that the analysis does not apply to outgroup influence, in which the target and source belong to different groups. This point was acknowledged by Turner and Oakes (1989) in their discussion of the difference between "influence" and "compliance." They noted that, whereas ingroups produce influence, "outgroups produce compliance, i.e., it is people with whom one does not expect to agree, whom one cannot be influenced by, that must resort to coercion, force and power to change behavior (p. 254) (see also Mugny & Perez, 1991). In addition, the self-categorization analysis fails to acknowledge the possibility that, at least under certain conditions, ingroups use normative influence techniques (e.g., threats of sanction) to shape members' behavior. Based on our earlier discussion of factors that stimulate desire for inclusion and fear of exclusion, such techniques are likely to be particularly effective when employed by ingroups. Thus, although self-categorization theorists are correct that informational and normative influence are sometimes intertwined, we believe that maintaining the conceptual distinction between them is worthwhile.

Traditional conceptualizations of normative/informational influence and compliance/acceptance have also been questioned by Wood (1999). Regarding normative influence, Wood argued that outcomes provided by others can prompt *either* public change that is limited to situations involving surveillance by the source *or* private change that endures in the absence of surveillance. The

former type of influence, which occurs when outcomes that others provide are not highly important, involves limited cognitive processing of relevant information and transitory attitude change. The latter form of influence, which occurs when such outcomes are important, involves substantial processing and enduring attitude change. Because this relationship between level of processing and public/private change also obtains when other goals are salient (e.g., desire to hold a valid position), Wood concluded that "the distinction between public and private agreement is not a reliable indicator of the motives underlying participants' responses in social influence research" (p. 563).[2] Nevertheless, Wood argued for maintaining the distinction between normative and informational goals, because people in everyday settings behave differently when motivated by groups to obtain acceptance and to hold a valid position (see also Chaiken et al., 1996).[3]

Finally, it is worth considering the possibility that the desire for accuracy, which is assumed to underlie informational influence and plays an important role in many other phenomena (Hardin & Higgins, 1996; Levine & Higgins, 2001), may be a secondary motive that derives from the desire for acceptance. In other words, the motive to hold accurate views on factual and opinion issues, while very real, may be important precisely because it facilitates an individual's ability to adapt to group life and thereby attain inclusion and avoid exclusion. This perspective is consistent with evolutionary analyses of individual motivation for inclusion discussed earlier, which emphasize the critical importance of group membership for individual survival. To gain group acceptance and avoid rejection, an individual must behave in ways that other members see as instrumental to attaining collective goals. One such behavior is holding "accurate" views on issues of importance to the group—in most cases, views that agree with group consensus. Failure to hold such views elicits strong reactions from other members, including pressure to conform, negative evaluations, and various kinds of punishment (Festinger, 1950; Hogg, 2005; Levine, 1989). Given this reward structure, the desire to hold accurate (i.e., consensual) positions has adaptive *social* consequences for individuals (cf. Brewer, 2004).

In concluding this section, we do not wish to leave the impression that fear of group sanctions for deviance invariably leads to conformity. One exception, noted earlier, occurs when the presence of a social supporter reduces anxiety about punishment. In addition, fear of group sanctions can produce responses other than conformity. These include leaving the group, remaining in it but redefining one's group-relevant identity (e.g., by emphasizing simultaneous memberships in other highly valued groups), or engaging in collective action designed to protect oneself or alter the group via minority influence (Hart & Van Vugt, 2006; Hogg, 2005; Levine & Kaarbo, 2001; Sani & Reicher, 1998). While these strategies are often difficult to implement, for example, because of low intergroup permeability in the case of exit (Ellemers, 1993) or majority resistance in the case of minority influence (Wood, Lundgren, Ouellette, Busceme, &

Blackstone, 1994), they are nonetheless used. Finally, people are sometimes insulated from fear of group sanctions by other powerful motives, such as the desire to establish personal distinctiveness (Brewer, 1991; Hornsey & Jetten, 2004; Lemaine, 1974). Interestingly, this desire may be based, at least in part, on the assumption that the courage one shows by defying the group in the short run will elicit its approval in the long run.

Group Decision Making and Problem Solving

The preceding section focused on the reactions of an attitudinal or behavioral minority to the actions or mere existence of a corresponding majority. In this section, our focus shifts to group contexts in which there is genuine interaction between group members in the pursuit of some common goal (e.g., reaching a decision, solving a problem).

One clear implication of our analysis of inclusion/exclusion is that there should be a *strength-in-numbers* (i.e., majority influence) effect in such contexts (cf. Hastie & Kameda, 2005). As we noted earlier, small (minority) factions are more susceptible to threats of exclusion by large (majority) factions than vice versa. Hence, all else being equal, decision alternatives or problem solutions that enjoy high member support at the beginning of group interaction should have a disproportionately high likelihood of being adopted by the group during interaction. And, indeed, a substantial literature—much of it employing Davis's (1973) *social decision scheme* model—has documented that majority (or even plurality, cf. Kerr, 1992) factions are very likely to prevail in decision-making groups (for reviews, see Baron & Kerr, 2003; Stasser, 1999; Stasser, Kerr, & Davis, 1989). For example, in juries, one can accurately predict final verdicts most of the time from simply knowing the predeliberation preferences of group members (e.g., Kalven & Zeisel, 1966; Kerr, 1981).

But can such reliable strength-in-numbers effects be confidently attributed to normative pressure based on minority members' fear of exclusion for deviance? Clearly, other interpretations are possible. The most common (and plausible) alternative interpretation is that majorities exert informational influence (Burnstein, 1982; Stasser et al., 1989). This could occur because larger factions can muster a greater number of novel arguments for their position (Burnstein & Vinokur, 1977), can appropriate more "airtime" and thereby dominate group discussion (e.g., Hawkins, 1962), and/or can profit from a consensus heuristic, which presumes that majority positions are likely to be correct (e.g., Darke et al., 1998).

Although informational influence can certainly play a role in the strength-in-numbers effect (Stasser et al., 1989), several lines of evidence suggest that it is not the only, or even the primary, determinant of this phenomenon. It has been found, for example, that majorities actually use less than their proportional share of airtime (e.g., Hawkins, 1962) but nonetheless exert more than their proportional impact on group decisions. Stated differently, the relative size of majority factions "overpredicts"

majority members' speaking time and "underpredicts" their ability to exert influence (Stasser et al., 1989). In addition, at least on issues involving continuous response dimensions, positions that are relatively distant from other group members' positions (i.e., more deviant) are less influential (Davis, 1996).

Support for a normative (as opposed to informational) interpretation of the strength-in-numbers effect can also be gleaned from reports of people who held minority positions in decision-making groups. For example, work on "groupthink" (Janis, 1982) has shown that people who find themselves in a minority position often go along with the majority because of fear of sanctions for deviance. In describing the unsuccessful 1961 Bay of Pigs operation, Janis (1982) noted, "The sense of group unity concerning the advisability of going ahead with the CIA's invasion plan appears to have been based on superficial appearances of complete concurrence, achieved at the cost of self-censorship of misgivings by several of the members" (p. 39). In explaining his failure to dissent from administration policy during this period, Arthur Schlesinger (1965) said, "In the months after the Bay of Pigs I bitterly reproached myself for having kept so silent during those crucial discussions in the Cabinet Room, though my feelings of guilt were tempered by the knowledge that a course of objection would have accomplished little save to gain me a name as a nuisance" (p. 255). Additional anecdotal support for the pressures that deviates feel is provided by memoirs of jurors who took minority positions (e.g., Burnett, 2002; Thornton, Menendez, Wrightsman, Posey, & Scheflin, 1995; Timothy, 1975). In addition to lamenting the fatigue associated with long deliberations and the difficulties of developing counterarguments, holdout jurors complain about the implicit and explicit exclusionary messages they received in retaliation for blocking group unanimity.

Further evidence against an informational interpretation of majority influence is that features of group tasks moderate the impact of majorities in just the manner one would expect if that impact were based on normative influence (cf. Kerr, 2001). As Festinger (1950) noted (see also Laughlin, 1980; Laughlin & Ellis, 1986), for some judgments there exist nonsocial criteria for validation, which place such judgments in the realm of *physical reality*. Thus, deciding whether San Diego is the most populous city in California can be determined through objective ("physical") means, such as consulting census records. But other judgments (e.g., deciding whether San Diego is the most beautiful city in California) cannot be made using consensually accepted objective criteria, because such criteria do not exist. Instead, these kinds of judgments must be made on the basis of social consensus (i.e., what other people believe about the beauty of cities in general and San Diego in particular). It is precisely these kinds of judgments, which involve *social reality*, for which minority positions are most difficult to defend in the face of majority opposition, because the minority cannot appeal to objective criteria to bolster its position. Minorities holding deviant positions on issues of social (as opposed to physical) reality should therefore feel particularly vulnerable to exclusionary sanctions if they dis-

sent from majority opinion. If so, majorities should exert greater influence on issues of social than physical reality (i.e., on judgmental than on intellective tasks), and indeed they do (Laughlin, 1999; Laughlin & Ellis, 1986; see also Crano & Hannula-Bral, 1994; Wood et al., 1994).[4]

There is also reason to believe that majority influence on judgmental issues has normative (as opposed to informational) underpinnings. In a group decision-making study investigating the bases of social influence, Kaplan and Miller (1987) found that, whereas content analyses of discussions of an intellective issue showed evidence of more informational than normative influence, parallel analyses of discussions of a judgmental issue showed just the opposite pattern, namely, more normative than informational influence. This difference was stronger in groups using a unanimity than a majority decision rule, perhaps because it was harder to reach consensus in the former case and hence groups relied more heavily on their "default" influence mode. Kaplan and his colleagues have also identified other factors that increase the likelihood that normative (as opposed to informational) messages will be used in decision-making groups and will prove effective (for reviews, see Kaplan, 1989; Kaplan & Wilke, 2001). Two of the most important of these are the group's dominant goal and the personal orientation of its members. In general, normative messages are more prevalent and more influential when the group is motivated to achieve positive socioemotional relations (as opposed to high-quality decisions) and when members are oriented toward group harmony (as opposed to task performance).

It is also the case that group decision-making procedures that force members to take sides early and publicly (e.g., by asking for a show of hands before beginning discussion) alter the style of group deliberation from a rational exchange and weighing of information (evidence-driven deliberation) to a confrontation between antagonistic factions (verdict-driven deliberation) (Hastie, Penrod, & Pennington, 1983; Kerr & MacCoun, 1985). Likewise, all else being equal, groups are less likely to rely on informational argumentation (and more likely to rely on normative argumentation) when a clear majority position exists than when it does not (Parks & Nelson, 1999).

Of course, majorities do not always prevail in decision-making contexts, and the conditions under which they fail are revealing. On intellective tasks (e.g., simple arithmetic problems, semantic associations), one-person minorities holding positions that are deviant but clearly "correct" within the group's shared conceptual system do more than resist majority influence—they typically convince the majority to adopt their position (a *truth-wins* social decision scheme) (Laughlin & Ellis, 1986; Laughlin, Kerr, Davis, Halff, & Marciniak, 1975). One interpretation of this effect is that the minority member's strong conviction about the validity of his or her position reduces fear of group sanctions for deviance. Moreover, on tasks that contain both judgmental and intellective elements (e.g., vocabulary tests—Laughlin et al., 1975; memory tests—Hinsz, 1990), a person who advocates a deviant but "correct" position in the face of majority opposition also can prevail, as long as he or she receives

support from another group member (a *truth-supported wins* social decision scheme). Thus, just as social support can reduce susceptibility to majority pressure in conformity settings by reducing fear of exclusion for deviance, so the presence of a supporter in decision-making settings may blunt the power of the majority by weakening its normative impact. It is also noteworthy that even when a group member knows the demonstrably correct answer to a problem, he or she is less likely to prevail (because of failure to mention the answer or argue vigorously for it) when the costs of exclusion are raised (viz., when the member has low status in a group that must continue to work together; Torrance, 1954).

As noted earlier, an important condition for normative influence is the majority's ability to monitor (and punish) deviant behavior. The impact of surveillance on social influence in interacting groups has been studied by comparing (1) individuals' private opinions prior to discussion, (2) the group decision that emerges during discussion, and (3) individuals' private opinions following discussion. Many such comparisons have been made in group polarization studies (see, e.g., Kerr, Davis, Meek, & Rissman, 1975; Myers & Lamm, 1976), and a common pattern has emerged: In nearly all cases, a "rebound" effect occurs, such that prediscussion opinions become more polarized during discussion and then shift back toward moderation after discussion. For example, in jury settings, some jurors who initially think the defendant is innocent but nonetheless join a unanimous guilty verdict privately assert, after deliberation, that they still believe in the defendant's innocence (e.g., Kerr et al., 1976). This pattern implies that at least some of the group polarization effect (i.e., the increased extremitization from prediscussion to discussion) reflects compliance to those holding the more prevalent (guilty) position in the group, motivated by the desire to avoid exclusion.

Other aspects of the group polarization literature are also consistent with this conclusion. Although there is substantial evidence that informational influence contributes to polarization effects (e.g., Burnstein & Vinokur, 1977; Ebbesen & Bowers, 1974), many findings challenge the sufficiency of this process. For example, if polarization is motivated by a desire to avoid deviance and its social costs, then merely permitting the exchange of group member preferences—without the opportunity for discussion and mutual persuasion—should be sufficient to produce polarization. And, indeed, although omitting group discussion does attenuate the polarization effect somewhat, it does not eliminate it (e.g., Baron et al., 1996; Baron & Roper, 1976; Myers, 1982). Several social comparison models of group polarization have been advanced, all of which are broadly consistent with a normative, or deviance-avoidance, explanation of polarization. For example, the position that individuals take on risky-shift items appears to represent a compromise between their (bold and risky) ideals and what they perceive to be the group norm, as predicted by the *pluralistic-ignorance theory* of polarization. If and when they discover that others (making the same compromise) define a descriptive group norm that is actually more risky than they had presumed, they shift their preferences in the norma-

tively preferred direction (cf. Pruitt, 1971). And as we have seen in a number of paradigms, when individuals discover that at least one other group member takes a position closer to their own ideal, they feel liberated to take a position somewhat further from the group norm in the direction of that ideal (Pruitt, 1971).

To this point, we have assumed that the robust strength-in-numbers effect in groups is attributable—in large part—to compliance motivated by deviates' desire to avoid social exclusion. But another, more subtle, mechanism may also be at work. That is, the minority's fear of exclusion may increase the majority's ability to produce *informational*, as well as normative, influence. In other words, minority members' concern about exclusion may increase the likelihood that they will be genuinely persuaded by majority members' arguments. This could occur if fear of exclusion motivates minority members to attend to, elaborate, and value majority members' arguments. Although direct evidence for these processes is not available, several lines of research are consistent with them (e.g., Crano & Chen, 1998; De Dreu & De Vries, 1996; Mackie, 1987; Schulz-Hardt, Frey, Luethgens, & Moscovici, 2000; Tesser, Campbell, & Mickler, 1983).

Behavior in Mixed-Motive Settings

We now turn our attention to cases in which group members' individual interests are not wholly compatible with one another or with collective interests. This state of negative interdependence clearly characterizes settings with explicit intragroup conflict, such as social dilemmas. Perhaps less obviously, it also characterizes most group performance contexts. Even though group members typically share a common interest in achieving a high level of joint performance, the costs (in effort, time, and lost opportunities) required to achieve such performance can create (or exacerbate) conflicts of interest among members (e.g., how hard one should work on the group's behalf; cf. Kerr, 1986). We argue that motivation to avoid social exclusion can powerfully shape members' behavior in such settings. We discuss first the less obvious case of group performance and then the more obvious case of social dilemmas.

GROUP PERFORMANCE

There are pervasive social expectations about how group members should behave on collective tasks. In some cases, these expectations deal with members' effort, that is how hard they try. In other cases, these expectations deal with members' competence, that is how well they perform vis-à-vis some standard of quality. As discussed below, evidence of laziness or incompetence reduces a member's social desirability and value to a group. Hence, if concern about exclusion motivates behavior in group settings, we would expect people to work harder and perform better when their behavior can be evaluated by others than when it cannot.

One line of relevant work concerns social facilitation, or the finding that others' presence sometimes energizes

performance (Zajonc, 1965). Social facilitation is often explained in terms of evaluation apprehension. According to this explanation, the presence of others causes people to become concerned about how their performance will be evaluated, which increases their level of drive or arousal. This in turn increases their task-relevant motivation (Cottrell, 1972; see also Blascovich, Mendes, Hunter, & Salomon, 1999; Geen, 1989; Harkins & Szymanski, 1987). It is important to note that the enhanced arousal produced by others' presence only facilitates performance on simple (well-learned) tasks. On difficult tasks, it has just the opposite effect, namely, inhibiting performance.

Concern about evaluation also contributes to social loafing, or the finding that others' presence sometimes deenergizes performance (Latane, Williams, & Harkins, 1979; see also Karau & Williams, 1993). Research on social loafing shows that people often expend less effort when working collectively rather than coactively or individually. This effect is often explained by positing that the ability to "hide in the crowd" in collective situations reduces group members' feeling of evaluation apprehension, which in turn reduces their task-relevant effort. To the extent that apprehension-based arousal enhances performance on simple tasks but impairs performance on difficult tasks, as argued previously, group members who believe that their individual performance cannot be evaluated should perform relatively poorly on simple tasks (because they are too relaxed) but relatively well on complex tasks (because they are not too tense) (cf. Jackson & Williams, 1985). Evidence on the role of task complexity in social loafing provides some support for these hypotheses (Karau & Williams, 1993).

Other research is also consistent with the notion that performance in group settings is influenced by members' concerns about how others evaluate them. For example, the performance of workers in the classic Hawthorne studies (e.g., Landsberger, 1958) was affected more by the presence of observers per se than by all the workplace environment manipulations examined. In addition, the consistent failure of members of face-to-face brainstorming groups to produce as many ideas as members of nominal groups can be attributed, at least in part, to evaluation apprehension. Anxiety that one's contributions may be negatively evaluated by others contributes to reluctance to speak up in brainstorming groups (e.g., Camacho & Paulus, 1995).

The impact of evaluation apprehension on performance in groups should, in principle, be moderated by members' dependence on one another for meeting their belongingness needs. Thus, we would expect stronger conformity to others' performance expectations in groups that are highly cohesive, elicit strong identification, and have high levels of member interaction or interdependence. Indeed, group productivity tends to be higher in groups that are more cohesive (Mullen & Copper, 1994) and elicit stronger identification (Haslam, 2004), especially when those groups have high levels of member interaction or interdependence (Gully, Devine, & Whitney, 1995). In a similar vein, tendencies to socially loaf are moderated by group valence (Karau & Williams,

1993), with less loafing in higher-valence groups. The occasional exceptions to these rules seem to occur when the group does *not* expect members to do their best. So, for example, when groups of workers hold antiproduction norms and punish "rate busters," high cohesiveness results in reduced, rather than enhanced, performance (Roethlisberger & Dickson, 1939; Schachter, Ellertson, McBride, & Gregory, 1951; Seashore, 1954).

Concern about others' evaluations can also influence how group members respond to common social interaction norms, such as the ubiquitous equity norm, which stipulates that group members' outcomes should be proportional to their inputs. Several studies show that when one member's contributions are not essential (i.e., are "dispensable") to the group's success, that person's task-relevant motivation is undermined (e.g., Kerr, 1983; Kerr & Bruun, 1983). But if group members expect to be rewarded equally, then the equity norm would prescribe that they contribute equally to the group product. Hence, those who contribute less than others (e.g., don't work as hard) would violate the equity norm and invite negative group sanction (and also produce guilt if the equity norm has been internalized). Thus, even when one's efforts are dispensable to the group's success, a person is unlikely to reduce his or her effort if others can monitor this effort and are likely to interpret low effort as failure to do one's fair share (Harkins & Petty, 1982). Conversely, when an ostensibly capable group member is clearly *not* doing his or her fair share, other group members will reduce their own efforts rather than tolerate the inequity of carrying a "slacker," even when such behavior produces costs for the self and the group (cf. Kerr, 1983). Such tendencies to match fellow members' effort levels are accentuated by conditions that increase concerns with social relations in the group (e.g., anticipated future interaction; Groenenboom, Wilke, & Wit, 2001).

SOCIAL DILEMMAS

Group members' decisions about how hard to work on group tasks are in many ways parallel to their decisions about how much to cooperate in mixed-motive situations, such as social dilemmas (Kerr, 1983, 1986). In social dilemmas, a personally beneficial or group-defecting choice yields higher outcomes to the individual than does a cooperative choice no matter what other members of the group do, yet universal defection produces worse outcomes for individuals and the group than does universal cooperation (Orbell & Dawes, 1981). Hence, it is not surprising that many of the behavior patterns identified earlier in the group performance literature are paralleled in the social dilemma literature. Just as effort in performance settings is enhanced by attraction to and identification with the group, so cooperation in social dilemmas varies positively with attraction to and identification with the group (e.g., Kramer & Brewer, 1984; Sattler, 1998), as well as the publicness (i.e., surveillance potential) of one's behavior (e.g., DeCremer & Bakker, 2003; although see Kerr, 1999). Moreover, just as there is a general expectation that group members should work hard on group tasks, so there is a general expectation

that group members should sacrifice self-interest when the welfare of the group is at stake. (As noted earlier, such group-serving norms could well have evolutionary roots.) And such sacrifice indeed occurs. Striking anecdotal evidence indicates that members of military units sometimes sacrifice their personal safety, and even their lives, to aid their comrades (e.g., Fussel, 2003; Stouffer, 1949). More prosaic experimental evidence confirms the power of (often implicit) cooperative norms (e.g., Hertel & Kerr, 2001; Van Vugt & Hart, 2004; Zdaniuk & Levine, 2001). Uncooperative group members are punished in one way or another (Dawes, McTavish, & Shaklee, 1977; Fehr & Gächter, 2002), and removing the threat of such exclusion for uncooperative behavior increases the likelihood of such behavior (Kerr, 1999; Masclet, Noussair, Tucker, & Villeval, 2003; Ouwerkerk et al., 2005), as does reducing a person's relative concern for the group as a whole versus some particular member of the group (Batson et al., 1995).

Other studies suggest that cooperative behavior is shaped by compliance with more specific social interaction norms (Kerr, 1995). One example is the powerful positive effect of group discussion on cooperation in social dilemmas. This effect stems, at least in part, from commitments to cooperate that group members make during discussion and then feel bound to honor (Chen & Komorita, 1994; Kerr & Kaufman-Gilliland, 1994). Likewise, evidence indicates that many patterns of cooperative choice reflect the prescriptions associated with norms of fairness or reciprocity (e.g., Komorita, Parks, & Hulbert, 1992; van Dijk & Vermunt, 2000; Wit, Wilke, & Oppewal, 1992).

We are not suggesting, of course, that group members never act selfishly, putting their own interests above the group's interests and violating normative expectations. There is no question that group members have other concerns besides inclusion (e.g., minimizing effort, maximizing short-term personal payoffs). But at least some of these violations of normative expectations may be reactions to the group's perceived violation of the implicit contract that conformity will be rewarded by inclusion. For example, if a member feels that the group (or its leadership) is not treating him or her in a respectful manner (e.g., by acting autocratically, by refusing to listen to his or her concerns), then the individual will no longer feel obligated to meet the group's expectations (e.g., Branscombe, Spears, Ellemers, & Doosje, 2002; De Cremer, 2002, 2003; Tyler & Degoey, 1995; Van Vugt, Jepson, Hart, & De Cremer, 2004). Likewise, a member's willingness to work for or contribute to the group will be undercut if the individual is convinced that his or her inclusionary needs will not be met by the group (Kerr, Harris, Messe, Poulsen, & Seok, 2005; Twenge, Baumeister, De Wall, Ciarocco, & Bartels, in press).

THE GROUP AS THE SOURCE OF INCLUSION/EXCLUSION

In the previous section, we focused on the group member as the *target* of inclusion/exclusion. In this section, we turn our attention to the group as the *source* of inclusion/exclusion. First we discuss the impact of collective goals on the monitoring system that groups use to assess individuals' contributions to collective goals, as well as the nature of this system. Next we consider the consequences of this monitoring system for how groups behave toward individuals, focusing primarily on exclusive reactions.

Collective Goals and Group Monitoring Systems

A fundamental property of groups is that they seek to achieve collective goals. Many typologies of group goals have been offered over the years. For example, Mackie and Goethals (1987) suggested that groups (as well as individuals) want to attain (1) utilitarian goals involving tangible outcomes, such as food and shelter; (2) knowledge goals involving valid information about how the world works; and (3) identity goals involving accurate self-knowledge and positive self-concept. In addition to such general goals, groups may seek to achieve a variety of specific task goals. Extrapolating from McGrath's (1984) circumplex model of group tasks, these latter goals include generating ideas or plans, solving problems with correct answers or deciding issues involving preferences, resolving conflicts of viewpoint or interest, and performing tasks that have objective standards and resolving conflicts of power. Moreover, to attain collective goals of any sort, groups need to satisfy a variety of "maintenance goals" that are critical to group survival (Cartwright & Zander, 1968). These include gaining and retaining members, managing role assignments, facilitating cohesion, enforcing norms, and so on.

Given that groups have goals, it is hardly controversial to suggest that they will pay attention to factors that facilitate and inhibit the attainment of these goals. Of the many potential determinants of group goal attainment, members' behaviors are typically quite important. Therefore, groups are likely to spend substantial time and energy monitoring these behaviors. In our earlier discussion of individuals' need for group inclusion, we noted that the monitoring system people use to assess their social status is asymmetrical (i.e., more sensitive to rejection than acceptance) and automatic (i.e., continuous, involuntary, and unconscious). We would argue that groups use a similar system to monitor individuals, although this similarity is greater in regard to system asymmetry than to automaticity. Just as individuals are particularly sensitive to group rejection, so groups are more sensitive to individuals' negative than positive contributions to collective goal attainment (Blanton & Christie, 2003). However, whereas groups are typically highly vigilant in monitoring individuals' contributions, their monitoring systems frequently do not fulfill the criteria of automaticity. Instead, these systems often have "controlled" features, such as routinized performance reviews carried out by designated group representatives. As discussed below, the outcome of the group monitoring system influences the form, valence, and magnitude of group responses to individuals (typically some form of

exclusion in the case of negative contributions and some form of inclusion in the case of positive contributions).

This analysis is consistent with recent evolutionary explanations of why and when groups exclude members (e.g., Kurzban & Leary, 2001; Neuberg et al., 2000). According to these explanations, because humans live in interdependent, cooperative groups, various forms of exploitative, or "selfish," behavior pose serious threats to group (and ultimately individual) welfare. These threats, in turn, have caused humans to evolve "mechanisms to identify individuals who threaten or hinder successful group functioning, to label them as such, to motivate group members to withhold group benefits from them, and to separate such individuals from the group if necessary" (Neuberg et al., 2000, p. 36).[5]

This evolutionary perspective suggests that certain kinds of individual behaviors may operate as "generic" threats, eliciting exclusion from a wide range of groups (cf. Neuberg et al., 2000; Stangor & Crandall, 2000). Of particular importance are behaviors that directly undermine the group's ability to function effectively. Some of these behaviors reflect lack of task-relevant knowledge or skills, as when new members do not understand the requirements of the overall group task or cannot perform the subtasks they are assigned. Other behaviors reflect lack of task-relevant motivation, as when individuals fail to work hard (socially loaf) because they believe others will take up the slack (Kerr, 1983). Social loafing is just one of several "misbehaviors" that undermine a group's ability to achieve its goals because they violate the implicit (sometimes explicit) social contract stipulating that group members should behave cooperatively to achieve shared goals. Other behaviors that undermine group effectiveness include being an "imposter," that is, publicly claiming a group identity while violating the expectations associated with this identity (Hornsey & Jetten, 2003), and failing to honor generic social norms such as reciprocity, promise/commitment keeping, property rights, and fair exchange (Kerr, 1995). It is difficult to think of cases, for example, in which cheating and stealing within the group are not punished (Kurzban & Leary, 2001; Neuberg et al., 2000). Of particular importance are violations of the norms prescribing that group members should sacrifice personal welfare in favor of group welfare (at least up to a point) and should show loyalty to their ingroup, especially when intergroup conflict is high (cf. Hertel & Kerr, 2001; Van Vugt & Hart, 2004; Zdaniuk & Levine, 2001). Disloyalty is typically viewed as the worst form of deviance and hence elicits the harshest penalties (Hogg, Fielding, & Darley, 2005; Levine & Moreland, 2002).

Other normative violations pose indirect but still potentially real threats to group effectiveness. These include acting in an unpredictable fashion (Kurzban & Leary, 2001) and violating implicit "residual rules" regarding appropriate dress, speech, cleanliness, interaction distance, and so on (Scheff, 1966), Whereas these examples (and the ones mentioned previously) deal with violations of behavioral norms, exclusion can also be triggered by various "abnormal" bodily conditions, such as physical disabilities or deformities and skin disorders (e.g., Neuberg et al., 2000; Thornhill & Gangestad, 1993). According to evolutionary explanations of exclusion, these conditions elicit negative responses because they imply inability to reciprocate benefits or produce viable offspring (in the case of disabilities or deformities) or danger of parasitic infection (in the case of skin disorders).

Social psychologists have been particularly interested in one kind of indirect threat to group effectiveness, namely, deviations from group consensus on opinion issues. In an influential early formulation, Festinger (1950) argued that group members want to attain opinion uniformity for two reasons: (1) desire to validate opinions that are not based on physical reality (social reality motive) and (2) desire to move toward (i.e., achieve) group goals (group locomotion motive). These two motives bear some resemblance to those that underlie informational and normative influence, respectively (Deutsch & Gerard, 1955). However, as discussed in more detail later, Festinger argued that both the social reality and group locomotion motives have a normative component in that individuals who impede the group's ability to satisfy either motive are perceived as "dangerous" and hence elicit some form of exclusion.

Although certain behaviors on the part of an individual may have generic threat-inducing properties, this does not mean that a given behavior is equally threatening to all groups at all times. Instead, the perceived threat value of a behavior varies as a function of the perceived probability that it will seriously undermine the group's ability to achieve a valued goal, and this probability judgment, in turn, can be influenced by numerous features of the behavior itself, the individual who emits the behavior, and the group context in which the behavior occurs (Levine, 1989; Levine & Moreland, 2002). For example, holding constant the value of the group goal, violation of a norm is likely to be perceived as more threatening if the violation is seen as directly (rather than indirectly) inhibiting goal attainment, if it is clear-cut (rather than ambiguous), and if it occurs repeatedly (rather than once). In addition, violation of a norm is likely to be perceived as more threatening if it is done by a member who has high (rather than low) status, has been in the group for a long (rather than a short) time, and seems unmotivated (rather than motivated) to achieve the group goal. Finally, violation of a norm is likely to be perceived as more threatening if the group is performing poorly (rather than well), is understaffed (rather than adequately staffed), and is in a competitive (rather than cooperative) relationship with an outgroup.

This line of argument has two important implications. First, it suggests that the nature of norm violations is often not "perceptually given" but, rather, is constructed during an inferential process.[6] That is, the group uses information about the violation itself, the person who commits the violation, and the group context in which the violation occurs to make attributions about the person and to judge the likely impact of his or her behavior on group goal attainment (Levine, 1989; see also Hogg et al., 2005; Stangor & Crandall, 2000). This inferential process can

be more or less complex, depending on the amount and clarity of information that the group has access to and decides to use. Second, because the inferential process occurs at the group level, it can involve a range of social processes, including majority and minority influence, group polarization, leadership, coalition formation, negotiation, and so on (Levine, 1989). The complexity of these processes makes it difficult, if not impossible, to predict the outcome of the group inferential process simply by knowing what individual members think about the norm violation prior to group interaction. Because these same processes operate when groups make decisions about how individuals should be treated, it is also useful to adopt a group-level perspective when analyzing the nature of exclusive responses.

A similar analysis is relevant to cases in which individuals adhere to, rather than violate, group norms and hence elicit inclusive, rather than exclusive, reactions. Evidence indicates, for example, that group members (1) believe a slider (who shifts from disagreement to agreement with the majority position) has less genuine belief in the correctness of his or her final position and more desire to be liked than does a mode (who always agrees with the majority position) and (2) evaluate the slider more negatively than the mode (Levine, Saxe, & Harris, 1976). Recent research in which a minority faction becomes a majority faction as the result of majority members' conversion to the minority position yields consistent findings (Prislin, Levine, & Christensen, 2006). Here, former minority (current majority) members expressed more liking for converts who changed their position for genuine reasons (i.e., because they were convinced by persuasive arguments) than for superficial or unexplained reasons. Thus, it appears that group members make attributions about the motives of conformers, as well as deviates, and these attributions influence their reactions to these individuals (cf. Morris & Miller, 1975).

This is not to say, of course, that group members do not sometimes appraise and respond to norm violations (and norm adherence) on their own. If, as we have suggested, individual group members are very alert to risks of exclusion, then deviation by others may often be deterred or reversed by exclusionary actions or signals (e.g., commenting on the deviant behavior) from members acting on their own (e.g., Fehr & Gächter, 2002). Such independent individual actions can invoke group norms, identify behavior as deviant, and remind the deviate of the potential for more concerted group action. Such individual-level appraisals and responses may be more or less consistent across members, depending on the degree to which they share appraisal biases and response preferences. To the extent members differ on these dimensions, norm violators can receive "mixed messages" about the appropriateness of their behavior.

Group Inclusion and Exclusion

As suggested earlier, the outcome of the group monitoring system—a judgment about the individual's contribution to group goal attainment—plays an important role in determining the form, valence, and magnitude of the

group's behavior toward the individual. In general, negative contributions elicit some form of exclusion, and positive contributions elicit some form of inclusion (Festinger, 1950; Levine et al., 2005; Mackie & Goethals, 1987). It is not the case, however, that deviance from group norms is always perceived negatively and hence elicits exclusion. Although groups have a default assumption that conformity to norms enhances, and deviance inhibits, collective welfare, they often desire (though do not require) that some members violate these norms by performing exceptionally well (Blanton & Christie, 2003). In such cases, positive deviance (i.e., norm violation in the "right" direction) can facilitate group goal attainment and hence elicit strong inclusion (cf. Schmitt, Silvia, & Branscombe, 2000).[7] Reactions to positive deviance have also been analyzed from the perspective of social identity theory and the Subjective Group Dynamics (SGD) model (e.g., Abrams, Marques, Bown, & Dougill, 2002; Abrams, Marques, Bown, & Henson, 2000; Hogg et al., 2005). This line of research, which is discussed in more detail later, indicates that reactions to positive ingroup deviates (who take a position consistent with, but more extreme than, the modal ingroup position) are most favorable when an intergroup context is salient and the ingroup's positive self-image is threatened.

Because social psychologists have devoted most of their attention to how groups respond to opinion deviance, it is appropriate to begin our discussion of group inclusion/exclusion by reviewing work on this topic. According to Festinger (1950), regardless of whether the social reality or group locomotion motive is dominant, differences of opinion among group members produce pressures toward uniformity, which in turn produce communication designed to reduce these differences. Communication can resolve opinion differences in two ways—either the group convinces the deviate to move toward its position or the deviate convinces the group to move toward his or her position. In addition, the group can redefine its boundaries by rejecting the deviate. Festinger made several predictions about the determinants of communication pressure in groups, the likelihood of communication to particular members, the amount of opinion change that communication produces, and the rejection of deviates. For example, he suggested that the likelihood of communication to particular individuals increases as a function of their perceived disagreement with the group, the group's desire that they remain members, and the group's belief that communication will alter their opinions. In addition, he suggested that group rejection of deviates increases as a function of their perceived disagreement with the group, the relevance of the issue to group function, and the group's cohesiveness. These and related hypotheses received partial support in studies conducted by Festinger and his colleagues (see reviews by Levine, 1989; Turner, 1991).

Perhaps the best known of these studies is Schachter's (1951) experiment on reaction to opinion deviance. In this study, participants were assigned to small groups that were either high or low in cohesiveness and then asked to discuss a topic that was either relevant or irrele-

vant to the group's purpose. Each group contained several naive participants and three confederates. The mode agreed with participants throughout the discussion. The slider began by disagreeing but gradually shifted to agreement. And the deviate consistently disagreed throughout the discussion. Results indicated that the deviate received the most overall communication, the slider received the next most, and the mode received the least. Moreover, the temporal pattern of communication to the three confederates differed. Communication to the mode remained uniform over the course of the discussion, communication to the slider decreased, and communication to the deviate tended to increase in all conditions except one (high-cohesive-relevant), where a final decrease occurred. Surprisingly, however, the greatest total communication to the deviate occurred in the low-cohesive-irrelevant condition. After the discussion, the three confederates were evaluated by the rest of the group. In general, the deviate was rejected more than the slider and mode, who were liked about equally. Finally, group cohesiveness and issue relevance affected deviate rejection as predicted, although their impact varied as a function of how rejection was measured (sociometric responses vs. committee nominations).

A number of more recent experiments have also investigated factors that affect rejection of opinion deviates in small groups (see Levine, 1989; Levine & Thompson, 1996). Evidence indicates that both the extremity and content of the deviate's position influence group reactions. Extreme deviates receive more communications and less favorable evaluations than do moderate deviates (e.g., Hensley & Duval, 1976; Levine & Ranelli, 1978). Moreover, the reasons that deviates give for taking their position influence how they are evaluated (e.g., Levine & Ruback, 1980), and deviates taking positions that are consistent with the wider Zeitgeist are liked better than those taking positions that are inconsistent (Maass, Clark, & Haberkorn, 1982; Paicheler, 1976). In addition, when deviates change their position over time, those shifting toward the majority position are generally liked better than those shifting away (e.g., Levine & Ranelli, 1978; Levine et al., 1976). Evaluation of shifting deviates is also affected by the social pressures ostensibly acting on them (e.g., the number of other people who support their position), presumably because these pressures affect the attributions that are made for their behavior (e.g., Levine, Sroka, & Snyder, 1977).

We have focused so far on studies in which group members' responses to deviates presumably were influenced by their desire to validate or evaluate the correctness of their opinions. Other studies have investigated how group members respond to deviates who interfere with group locomotion toward valued goals. These studies have operationalized deviance as either opinion disagreement or overt disruption of group performance (e.g., failure to follow instructions). It has been found, for example, that the greater the deviates' interference with group locomotion and the greater their perceived responsibility for this interference, the less they are liked (e.g., Berkowitz & Howard, 1959; Jones & deCharms, 1957; Miller & Anderson, 1979; Wiggins, Dill, &

Schwartz, 1965). Deviates' status within the group has also been shown to be important. Hollander (1958, 1960) argued that people who conform to group norms and demonstrate task-relevant competence gain status, or idiosyncrasy credit, which in turn allows them to deviate from group norms (but see Bray, Johnson, & Chilstrom, 1982). Other studies indicate that a deviate's status and degree of interference with group locomotion interact in determining majority members' reactions. High-status deviates tend to be severely punished for major interference with group locomotion toward valued goals, mildly punished for minor interference, and highly rewarded for facilitating locomotion toward these goals (e.g., Alvarez, 1968; Suchner & Jackson, 1976; Wiggins et al., 1965). In general, high-status persons who interfere with group locomotion are less likely to be defined as deviates than are low-status persons. However, if they are identified as deviates, they often receive severe punishment (e.g., Giordano, 1983; Hollander & Willis, 1967; Wahrman, 1970).

To understand reaction to deviance, it is important to consider the group context in which deviance occurs. Three potentially important features of this context are the group's norms regarding deviance, the group's desire for consensus, and social influence among majority members. Most groups have norms that prohibit deviance (e.g., Janis, 1982; Roethlisberger & Dickson, 1939). But some groups tolerate or even encourage deviance. These latter reactions can occur for several reasons, including members' desire to uphold a value system fostering freedom of expression, develop creative solutions to group problems, or demarcate the boundaries of tolerable behavior within the group (cf. Coser, 1962; McAuliffe, Jetten, Hornsey, & Hogg, 2003). Evidence indicates, for example, that deviates are liked better when the group has a norm favoring originality rather than objectivity (Moscovici & Lage, 1978). In addition to explicit norms regarding the acceptability of deviance, group members' desire for consensus, or collective cognitive closure, can affect reaction to deviance. Thus, Kruglanski and Webster (1991) found more rejection of deviates (and more acceptance of conformers) when the group's costs for failing to attain consensus were high rather than low. Finally, social influence among group members can affect how deviates are treated. Evidence indicates that people who see a fellow member react negatively toward a deviate are more likely to react negatively themselves than are people who do not have this experience (Wheeler & Caggiula, 1966; see also Dedrick, 1978), suggesting that exclusion itself can emerge as a group norm.

In recent years, substantial theoretical and empirical attention has been devoted to clarifying how reaction to opinion (and other forms of) deviance is influenced by the intergroup context in which it occurs. This work, which adopts a social identity approach, differs from that discussed above in a fundamental way, namely, by downplaying the importance of group members' face-to-face interaction and behavioral interdependence and focusing instead on their collective self-definition as members of social categories (Marques, Abrams, Paez, & Hogg, 2001). Initial work on reaction to deviance from

this perspective focused on the "black sheep effect," the finding that unlikable ingroup members are evaluated more negatively, and likable ones more positively, than similar outgroup members (e.g., Marques, Yzerbyt, & Leyens, 1988). Researchers working in this tradition interpret negative evaluations of unlikable ingroup members as a group protection strategy designed to maintain positive group identity by rejecting those who threaten this identity (Marques & Paez, 1994; but see Eidelman & Biernat, 2003).

Current social identity research on reaction to deviance is framed in terms of the SGD model (see reviews by Abrams, de Moura, Hutchison, & Viki, 2005; Marques, Abrams, Paez, & Hogg, 2001). This model is based on the assumption that group members are motivated to maximize both the positive distinctiveness of their ingroup vis-à-vis outgroups and the relative validity of their ingroup's norms. These goals, in turn, are satisfied through a combination of intergroup differentiation (regarding descriptive norms) and intragroup differentiation (regarding prescriptive norms). In addition to the research on reaction to positive ingroup deviates described earlier, a substantial amount of work on reaction to negative ingroup deviates (who threaten one or both of the goals mentioned above) is consistent with the SGD model. Evidence indicates, for example, that such deviates are more likely to be rejected (1) by people who identify strongly with the ingroup (e.g., Branscombe, Wann, Noel, & Coleman, 1993; Coull, Yzerbyt, Castano, Paladino, & Leemans, 2001; Hutchison & Abrams, 2003), particularly when their own prototypicality is threatened (Schmitt & Branscombe, 2001); (2) when deviance occurs in an intergroup context (Matheson, Cole, & Majka, 2003; but see Marques & Yzerbyt, 1988); (3) when the status of the ingroup is insecure (Marques, Abrams, & Serodio, 2001); (4) when there is lack of perceived consensus among ingroup members (Marques, Abrams, & Serodio, 2001); (5) when ingroup members feel a sense of accountability to one another (Marques, Abrams, Paez, & Martinez-Taboada, 1998); and (6) when the deviate challenges prescriptive ingroup norms (Bown & Abrams, 2003; Scheepers, Branscombe, Spears, & Doosje, 2002). This and related work provides strong support for the notion that social identity threat (Branscombe, Ellemers, Spears, & Doosje, 1999) plays an important role in group reaction to deviance.

Group responses to deviance can vary in form, valence, and intensity. For example, Festinger (1950) distinguished between communication directed toward changing the deviate's opinion and redefinition of the group's boundary by rejecting the deviate (see also Marques, Abrams, & Serodio, 2001; Schachter, 1951). A similar distinction was made by Orcutt (1973), who distinguished between inclusive and exclusive reactions to deviates. In the former case, the group wants to change the opinion of a person who seems responsive to persuasion. In the latter case, the group wants to exclude a person who seems unresponsive. Orcutt argued that the more serious the deviate's dissent is perceived to be, the more likely it will elicit an exclusive reaction. Yet another typology of reaction to deviance was offered by Israel

(1956), who argued that overt hostility toward a deviate is different from rejection of that person. Whereas overt hostility is used as a sanction to produce conformity, rejection is a last-ditch response that occurs after efforts to produce conformity fail. According to Israel, rejection can take three forms: expelling the deviate from the group, isolating the deviate from all interaction within the group, and depriving the deviate of the normal privileges of group membership (cf. Eidelman, Silvia, & Biernat, 2006; Williams, 2001).

An interesting example of how groups seek to alter the public (but not necessarily private) opinions of deviates was reported by Janis (1982) in his book on groupthink. He noted that self-appointed "mindguards" urged members who disagreed with the dominant group position to keep their concerns to themselves. Dissenters who had the temerity to question the group's position were subjected to strong conformity pressures, including disparaging remarks about their loyalty and competence. For example, during the Johnson administration, "everyone in the hierarchy, including every senior official, was subjected to conformity pressures, which took the form of making those who openly questioned the escalation policy the butt of an ominous epithet: 'I am afraid he's losing his effectiveness' " (Janis, 1982, pp. 114–115). Events in the years since indicate that advisors to more recent presidents have also felt pressure to go along with official policy.

Responses to deviance can be clarified by considering how groups use inclusion/exclusion to mark individuals' role transitions across different phases of group membership (Levine et al., 2005). Rather than simply being members or nonmembers of groups, people can relate to groups in more differentiated ways (i.e., as prospective members, newcomers, oldtimers, and ex-members). Given that group membership varies along an ingroup–outgroup continuum, group inclusion/exclusion can involve not only managing transitions across the group's external boundary but also managing transitions across internal boundaries that demarcate different roles members can play.

Levine and colleagues (2005) discuss three psychological processes relevant to inclusion/exclusion—evaluation, commitment, and role transition (cf. Levine & Moreland, 1994; Moreland & Levine, 1982). Evaluation involves efforts by the group and the individual to assess the past, present, and probable future rewardingness of their own and alternative relationships. Thus, a group's evaluation of an individual is high to the extent that his or her past, present, and anticipated future contributions to group goal attainment are greater than those of other people who were, are, or might be associated with the group. Evaluation affects feelings of commitment between the group and the individual, which can rise or fall over time as a function of each party's assessment of the other's rewardingness. The greater the group's commitment to an individual, the more likely it will be to engage in inclusive behaviors and the less likely it will be to engage in exclusive behaviors.

Role transitions are an important consequence of commitment, because they signal major changes in the

group's and the individual's relationship. These transitions occur when the two parties' commitment levels rise or fall to their decision criteria, which are specific levels of commitment indicating that it is time for the individual to undergo a role transition (i.e., move from one phase of group membership to another). Four basic role transitions can be identified (entry, acceptance, divergence, and exit). Entry reflects promotion from prospective member to new member, and acceptance reflects promotion from new member to full member. Because these role transitions indicate the individual's movement toward the core of the group (full membership), they signal the group's inclusion of the person. In contrast, divergence reflects demotion from full member to marginal member, and exit reflects demotion from marginal member to ex-member. Because these role transitions indicate the individual's movement away from the core of the group, they signal the group's exclusion of the person.

Groups use a variety of behaviors to indicate their inclusion/exclusion of an individual. Whereas some of these behaviors explicitly communicate the group's desire that the individual undergo a specific role transition, others do not. Some of these latter behaviors signal the group's level of commitment to the individual, based on its evaluation of his or her contributions. Positive behaviors (e.g., praise, personal warmth, bestowal of resources) generally indicate high commitment and thereby inclusion. In contrast, negative behaviors (e.g., criticism, personal coldness, withdrawal of resources) generally indicate low commitment and thereby exclusion. An important way in which groups indicate inclusion is by paying attention to the individual's arguments and thereby enabling him or her to exert influence. This type of inclusion has been studied under the rubric of minority influence. Although space considerations preclude a review of the extensive literature on minority influence (see DeDreu & DeVries, 2001; Martin & Hewstone, 2001; Wood et al., 1994), two issues relevant to our present concerns are worth noting. First, minority influence is hard to produce, and minorities are influential only under certain conditions (i.e., compared to majorities, minorities produce less public and direct private influence, although they can produce more indirect private influence) (Wood et al., 1994). Second, even when minorities succeed in exerting influence, they are viewed by majorities with ambivalence at best and hostility at worst (Moscovici, 1976).

Other behaviors that indicate inclusion/exclusion of an individual involve the group's decision criteria for role transitions. Groups signal inclusion by adopting decision criteria that (1) facilitate promotion from prospective to new member and from new to full member and (2) inhibit demotion from full to marginal member and from marginal to ex-member. Conversely, groups signal exclusion by adopting criteria with the opposite consequences—inhibiting promotion from prospective to new member and from new to full member and facilitating demotion from full to marginal member and from marginal to ex-member.

We have argued that groups generally behave in an exclusive (and negative) manner toward those who retard the attainment of collective goals. This tendency was recently documented in a meta-analytic summary of the strength of a large number of social psychological phenomena (Richard, Bond, & Stokes-Zoota, 2003). A mean effect size of .60 was reported for the finding that "people who deviate from a group are rejected by that group," as compared to mean effect sizes of .32 for the category of 27 group phenomena and .21 for all 474 social psychological phenomena in the sample. Given this strong tendency to punish deviates, the question arises as to whether all group members are equally likely to play the role of "enforcer." The answer seems to be no. In fact, there appears to be a bimodal distribution of enforcers. In some cases, they are high-status people, such as leaders (Hogg et al., 2005). In other cases, they are low-status people, such as peripheral group members (cf. Noel, Wann, & Branscombe, 1995).

Exclusive reactions toward deviance can vary in intensity. Mild examples of exclusion include frowns, derisive laughter, disparaging remarks, and withdrawal of privileges and responsibilities. Extreme examples include pressure to wear humiliating symbols (e.g., the scarlet letter "A"), threats of physical injury, and expulsion from the group. It is plausible to assume that the intensity of exclusive behaviors depends on the deviate's perceived interference with group goal attainment. This relationship makes sense if groups are particularly motivated to punish individuals who cause major harm (and deter others who might be inclined to behave the same way). However, an alternative possibility should be considered, namely, that groups *generally* prefer strong exclusive behaviors, because such behaviors send clear signals about the group's negative evaluation of deviance and hence are likely to influence the target's (and others') future actions.

In evaluating this latter hypothesis, it is important consider some potential costs associated with strong exclusive behaviors. For example, strong behaviors are more time-consuming and expensive to produce than are mild behaviors (firing an employee is much harder than frowning when he speaks). This is true, in part, because strong behaviors typically require greater coordination of group members' actions than do mild behaviors. In addition, strong behaviors make a greater impression on recipients than do mild behaviors, and sometimes this impression is not the one the group intended. Thus, a group that publicly criticizes a member for poor performance in an effort to increase his motivation may instead increase his anxiety to the point that he (and, perhaps, others) cannot function at all. Strong exclusive behaviors may also undermine a member's intrinsic motivation to work on the group task and make the group less attractive (Darley, 2001). When the person's contributions are indispensable to the group, for example, when loss of the person would lead to understaffing, the risks of alienating her may exceed the benefits associated with altering her behavior (Hogg et al., 2005).

The costs associated with strong exclusive behaviors often lead groups to use a cautious strategy in which harsh actions are taken only if milder ones fail (Festinger, 1950; Hogg et al., 2005). In fact, however, such steady es-

calation is rare. Because of their experiences in groups, people carry a set of expectations about how other members will behave toward them (e.g., regarding eye contact or personal greetings). Even small deviations from these expectations in the direction of reduced intimacy/acceptance (absent some benign attribution for such deviations) are likely to be interpreted as signaling exclusionary danger. Thus, the group rarely needs to make an overt threat to get a member's attention and prompt behavioral change—subtle and mild exclusionary signals will usually suffice. This argument is consistent with evidence, discussed earlier, regarding the sensitivity of individuals' monitoring system for detecting signs of group exclusion.

Notwithstanding the fact that groups generally respond to members' deviance with exclusive (and negative) behavior, it is not the case that negative behaviors are always used to punish negative contributions to group goal attainment and positive behaviors are always used to reward positive contributions. Groups sometimes use positive responses as "carrots" for individuals making negative contributions and negative responses as "sticks" for individuals making positive contributions. For example, carrots may be used when a group fears alienating an unproductive, but potentially useful, member, whereas sticks may be used when a group seeks to motivate a productive, but seemingly complacent, member. In addition, groups sometimes "pull their punches" with individuals making negative contributions for other reasons, including the desire to foster innovation or demonstrate tolerance (Levine, 1989; see also Alvaro & Crano, 1997; Crano & Chen 1998).

Our analysis to this point has assumed that a group's inclusive/exclusive reaction to an individual (1) is based on a (more or less rational) calculation of the person's contribution to group goal attainment and (2) has the instrumental goal of maintaining or increasing the group's ability to reach its goals. While these assumptions are plausible in many cases, they do not explicitly consider the role of emotion in influencing the direction and magnitude of the group's behavior. In particular, they do not address the possibility that inclusive/exclusive reactions to an individual may be attributable, in whole or in part, to the group's affective responses to the person or his or her actions. These responses may arise for several reasons. For example, a group may have a strong negative emotional response to a person because his behavior closely resembles an evolutionarily wired threat (cf. Kurzban & Leary, 2001; Neuberg et al., 2000). Several current models of sociality (e.g., Boyd & Richerson, 1992; Price et al., 2002) suggest a strong punitive reaction to behavior that can hurt the group or violates its norms, and experimental work demonstrates that people are willing to bear an extra cost simply to punish such behavior (Fehr & Gächter, 2002). Alternatively, a group may have a strong positive emotional response to a person because she was merely present when the group received a reward or possesses some valued trait. Such affective responses may influence the group's behavioral reactions to the individual by altering or even overriding

behavioral tendencies based on the "rational calculus" model proposed previously.

Emotional responses can influence a group's behavior toward an individual in another way as well. That is, emotions arising from the social dynamics of the group can affect how the person is treated by increasing the intensity of the inclusive or exclusive reaction that he or she elicits via one of the mechanisms discussed earlier, or, when no reaction tendency exists, by creating one. Examples include cases of extreme group adulation (e.g., responses to Hitler at Nazi rallies) as well as extreme group hostility (e.g., lynchings in the American South). Several mechanisms may underlie such effects, including social facilitation (Bond, 2000; Zajonc, 1965), emotional contagion (Hatfield, Cacioppo, & Rapson, 1994), behavioral contagion (Wheeler, 1966), and deindividuation (Diener, 1980; Postmes & Spears, 1998). These mechanisms illustrate additional ways in which group-level processes can influence inclusion/exclusion.

CONCLUSIONS

In this chapter, we used the notion of inclusion/exclusion to explain a range of small-group phenomena. In so doing, we adopted two different but complementary perspectives—one focusing on the individual member as the target of inclusion/exclusion, and the other focusing on the group as the source of inclusion/exclusion. Although there are differences in how groups and individuals deal with inclusion/exclusion (e.g., groups generally have more power to include/exclude than do individuals, and individuals generally are more sensitive to inclusion/exclusion than are groups), there are also striking parallels between the two perspectives. It may be useful, therefore, to suggest some general principles that help to integrate these perspectives.

First, there are strong commonalities in key motives that underlie individuals' desire for inclusion and groups' desire that individuals adhere to normative standards. In both cases, evolutionary factors are probably important. In addition, rational self-interest, based on previous experience, plays a critical role. Because of the rewards typically associated with inclusion and the costs typically associated with exclusion, individuals believe that inclusion generally facilitates personal goal attainment. Similarly, because of the rewards typically associated with normative behavior and the costs typically associated with deviant behavior, groups believe that normative behavior generally facilitates collective goal attainment.

Second, given that these motives produce strong desires for inclusion and normative behavior, both individuals and groups are acutely sensitive to information relevant to the satisfaction of their needs. This sensitivity, in turn, is reflected in similar kinds of monitoring systems by which individuals and groups assess their need satisfaction. One feature of these systems is asymmetry, that is, greater sensitivity to negative than to positive information. A second feature is automaticity, that is, continu-

ous, involuntary, and unconscious operation (although this is clearer in the case of individuals than groups).

Third, notwithstanding the fact that individuals generally desire to be included and groups generally desire that individuals behave normatively, the strength of these motives can vary across situations. Critical determinants of motivational strength are the salience and importance of momentary goals and the perceived impact of inclusion (from the individual's perspective) or normative behavior (from the group's perspective) on these goals. Thus, the more the individual wants to attain a particular personal goal and the more he or she believes that group inclusion is instrumental for doing so, the more motivated he or she will be to achieve or maintain inclusion. Similarly, the more the group wants to attain a particular collective goal and the more it believes that normative behavior is instrumental for doing so, the more motivated it will be to ensure that members engage in such behavior. This analysis suggests that, at least under certain conditions, individuals and groups make cognitive appraisals of one another's likely contributions to goal attainment rather than only responding in an automatic manner to evolutionarily wired cues.

Finally, this distinction between automatic and controlled processes in perception of the benefits of inclusion and normative behavior (and the costs of exclusion and deviance) is mirrored by a similar distinction regarding the processes underlying individuals' and groups' behavioral responses to these perceptions. As we argued earlier, the perceived benefits/costs of inclusion/exclusion for the individual and normative/deviant behavior for the group energize and direct both parties' behavior. For example, an individual who perceives that he or she is (or might be) excluded will be motivated to act, as will a group that perceives that its collective goal is (or might be) thwarted by a deviate. In both cases, these threat-induced behaviors will vary along a continuum from automatic/reflexive to controlled/planful. Factors that influence where a particular behavior falls on this continuum include the perceived intensity of the threat, the accessibility of habitual coping responses, and the availability of cognitive (and other) resources for selecting and producing responses.

We believe that our analysis of group processes in terms of inclusion/exclusion is distinctive in at least two ways. First, it seeks to provide a conceptual framework for understanding a wide range of small-group phenomena. With the notable exception of investigators working within the social identity tradition, contemporary researchers have been reluctant to offer integrative accounts of group processes, preferring instead to develop circumscribed explanations of specific phenomena. Although such analyses are often very insightful, their narrow scope has had the unfortunate consequence of fragmenting the field of group processes into seemingly unrelated subfields (Levine & Moreland, 1998). The disadvantages of such fragmentation are many (cf. De Dreu & Levine, 2006), and hence efforts to develop broader conceptual frameworks for understanding groups are worthwhile. Second, our perspective emphasizes the role of social needs in explaining group processes. Over the last few decades, research in social psychology has been dominated by a focus on how cognitive needs (e.g., to reduce uncertainty, to maintain consistent beliefs) influence people's thoughts, feelings, and actions (e.g., Kruglanski, 1989). This general perspective has yielded a wealth of important insights and will no doubt continue to generate interesting research. In the last few years, however, social psychologists in general and group researchers in particular have begun paying more attention to social needs, which include "the needs to be accepted by, connected with, and of value to other persons and social groups" (Brewer, 2004, p. 108). We applaud this increasing balance in how group processes are conceptualized and hope that our analysis will stimulate further work on the role that social needs (in conjunction with cognitive needs) play in such processes.

In closing, a few caveats are in order regarding the theoretical analysis we proposed and the empirical work we cited. As implied in the previous paragraph, we are not suggesting that inclusion/exclusion provides a simple and sovereign explanation of group processes. However, we are suggesting that by taking this perspective seriously, researchers will gain substantial leverage in predicting and explaining how groups and their members behave. In citing empirical work relevant to our arguments, we were selective rather than exhaustive. In addition to the literatures we discussed (conformity, group decision making and problem solving, behavior in mixed-motive settings, reaction to deviance), several other lines of work are also amenable to analysis emphasizing inclusion and exclusion. These include research on coalition formation (e.g., the fact that larger-than-necessary coalitions sometimes form because of the desire not to exclude people) and negotiation (e.g., the fact that outcome distributions sometimes reflect negotiators' concerns about their present and future relationships with their opponents). It is also important to acknowledge that whereas the studies we cited were consistent with our conceptualization, many were not designed to provide direct tests of the role of inclusion/exclusion in group processes. Because control conditions needed to rule out alternative explanations were usually not included in these studies, the evidence cited for many of our arguments should be viewed as suggestive rather than definitive. Finally, in some cases, there is little or no research bearing on our arguments. For example, neither the specific cues that individuals use to detect group exclusion nor the social dynamics that underlie group decisions about how to handle deviates have been systematically investigated. Although the absence of such research leaves several facets of our theoretical perspective untested, we believe this perspective suggests many interesting questions for future research.

ACKNOWLEDGMENTS

NOTES

1. Although Deutsch and Gerard specified that "another" can refer to a group, a person, or the self, analyses of normative influence have generally assumed that the source of influence is a group (Wood, 1999).

2. Mackie and her colleagues make the related point that group-mediated persuasion can involve either superficial or intensive cognitive processing (e.g., Mackie & Queller, 2000).

3. Wood (1999) distinguished between two forms of "normative" influence, one based on desire to obtain others' approval and the other based on desire to align oneself with valued reference groups and against derogated groups. We focus on the former kind of normative influence in this chapter. (See also Cialdini & Goldstein, 2004.)

4. Notwithstanding these findings, it is important to note that the distinction between social and physical reality may be more apparent than real. Perhaps matters of "physical reality" are thought to have objectively correct answers only because people assume that everyone perceives them in the same way (cf. Hardin & Higgins, 1996; Moscovici, 1976).

5. It is also possible, of course, for individual group members to exclude others who do not directly undermine the attainment of group goals, because doing so yields personal benefits, such as self-enhancement, cognitive simplicity, and anxiety management (Crocker, Major, & Steele, 1998). These individual-level explanations for exclusion are not the primary focus of our present discussion.

6. As noted earlier, violation of some "generic" norms may elicit immediate and strong reaction that does not depend on the sort of inferential process described here.

7. This is not to say, of course, that high achievers are always liked. When their accomplishments threaten perceivers' self-esteem or induce envy, they may be disliked (e.g., Feather, 1994; Schmitt, Silva, & Branscombe, 2000; Tesser, 1988).

REFERENCES

Abrams, D., de Moura, G. R., Frings, D., & Rutland, A. (2005, July). *People without: The definition of deviance and rejection of groups and their members*. Paper presented at the 14th General Meeting of the European Association of Experimental Social Psychology, Wurzburg, Germany.

Abrams, D., de Moura, G. R., Hutchison, P., & Viki, G. T. (2005). When bad becomes good (and vice versa): Why social exclusion is not based on difference. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 161–190). New York: Psychology Press.

Abrams, D., & Hogg, M. A. (1990). Social identification, self-categorization and social influence. *European Review of Social Psychology*, *1*, 195–228.

Abrams, D., & Hogg, M. A. (2001). Collective identity: Group membership and self-conception. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 425–460). Malden, MA: Blackwell.

Abrams, D., Hogg, M. A., & Marques, J. M. (2005a). A social psychological framework for understanding social inclusion and exclusion. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 1–23). New York: Psychology Press.

Abrams, D., Hogg, M. A., & Marques, J. M. (Eds.). (2005b). *The social psychology of inclusion and exclusion.* New York: Psychology Press.

Abrams, D. Marques, J., Bown, N., & Dougill, M. (2002). Anti-norm and pro-norm deviance in the bank and on the campus: Two experiments on subjective group dynamics. *Group Processes and Intergroup Relations*, *5*, 163–182.

Abrams, D., Marques, J. M., Bown, N., & Henson, M. (2000). Pro-norm and anti-norm deviance within and between groups. *Journal of Personality and Social Psychology*, *78*, 906–912.

Allen, V. L. (1965). Situational factors in conformity. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 133–175). New York: Academic Press.

Allen, V. L. (1975). Social support for nonconformity. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 1–43). New York: Academic Press.

Allen, V. L., & Levine J. M. (1968). Social support, dissent, and conformity. *Sociometry*, *31*, 138–149.

Alvarez, R. (1968). Informal reactions to deviance in simulated work organizations: A laboratory experiment. *American Sociological Review*, *33*, 895–912.

Alvaro, E. M., & Crano, W. D. (1997). Indirect minority influence: Evidence for leniency in source evaluation and counterargumentation. *Journal of Personality and Social Psychology*, *72*, 949–964.

Asch, S. E. (1952). *Social psychology*. New York: Prentice-Hall.

Asch, S. E. (1956). Studies of independence and submission to group pressure: I. A minority of one against a unanimous majority. *Psychological Monographs*, *70*(Whole No. 417).

Back, K. W., & Bogdonoff, M. D. (1964). Plasma lipid responses to leadership, conformity, and deviation. In P. H. Leiderman & D. Shapiro (Eds.), *Psychobiological approaches to social behavior* (pp. 24–42). Stanford, CA: Stanford University Press.

Barchas, P. (1986). A sociophysiological orientation to small groups. In E. Lawler (Ed.), *Advances in group processes* (Vol. 3, pp. 209–246). Greenwich, CT: JAI Press.

Baron, R. S., Hoppe, S. I., Kao, C. F., & Brunsman, B., Linneweh, B., & Rogers, D. (1996). Social corroboration and opinion extremity. *Journal of Experimental Social Psychology*, *32*, 537–560.

Baron, R. S., & Kerr, N. L., (2003). *Group process, group decision, group action* (2nd ed.). Buckingham, UK: Open University Press.

Baron, R. S., & Roper, G. (1976). Reaffirmation of social comparison views of choice shifts: Averaging and extremity effects in an autokinetic situation. *Journal of Personality and Social Psychology*, *33*, 521–530.

Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., & Aldeguer, C. M. R. (1995). Empathy and the collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology*, *68*, 619–631.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Baumeister, R. F., & DeWall, N. C. (2005). The inner dimension of social exclusion: Intelligent thought and self-regulation among rejected persons. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 53–73). New York: Psychology Press.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.

Baumeister, R. F., & Tice, D. M. (1990). Anxiety and social exclusion. *Journal of Social and Clinical Psychology*, *9*, 165–195.

Berkowitz, L., & Howard, R. C. (1959). Reactions to opinion deviates as affected by affiliation need (n) and group member interdependence. *Sociometry*, *22*, 81–91.

Blanton, H., & Christie, C. (2003). Deviance regulation: A theory of action and identity. *Review of General Psychology*, 7, 115–149.

Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, K. (1999). Social "facilitation" as challenge and threat. *Journal of Personality and Social Psychology*, *71*, 68–77.

Bond, C. F., Jr. (2000). Social facilitation. In A. E. Kazdin (Ed.), *En-

cyclopedia of psychology (Vol. 7, pp. 338–340). Washington, DC: American Psychological Association.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and loss: Vol. 3. Loss: Sadness and depression*. New York: Basic Books.

Bown, N. J., & Abrams, D. (2003). Despicability in the workplace: Effects of behavioral deviance and unlikeability on the evaluation of in-group and out-group members. *Journal of Applied Social Psychology*, *33*, 2413–2426.

Boyanowsky, E. O., & Allen, V. L. (1973). Ingroup norms and self-identity as determinants of discriminatory behavior. *Journal of Personality and Social Psychology*, *25*, 408–418.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, *100*, 3531–3535.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*, 171–195.

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity* (pp. 35–58). Oxford, UK: Blackwell.

Branscombe, N. R., Spears, R., Ellemers, N., & Doosje, B. (2002). Intragroup and intergroup evaluation effects on group behavior. *Personality and Social Psychology Bulletin*, *28*, 744–753.

Branscombe, N. R., Wann, D. L., Noel, J. G., & Coleman, J. (1993). In-group or out-group extremity: Importance of the threatened social identity. *Personality and Social Psychology Bulletin*, *19*, 381–388.

Bray, R. M., Johnson, E., & Chistrom, J. T., Jr. (1982). Social influence by group members with minority opinions: A comparison of Hollander and Moscovici. *Journal of Personality and Social Psychology*, *43*, 78–88.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Brewer, M. B. (2004). Taking the social origins of human nature seriously: Toward a more imperialist social psychology. *Personality and Social Psychology Review*, *8*, 107–113.

Burnett, D. G. (2002). *Trial by jury*. New York: Knopf.

Burnstein, E. (1982). Persuasion as argument processing. In H. Brandstätter, J. H. Davis, & G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 103–124). London: Academic Press.

Burnstein, E., & Vinokur, A. (1977). Persuasive argumentation and social comparison as determinants of attitude polarization. *Journal of Experimental Social Psychology*, *13*, 315–32.

Buss, D. M., & Kenrick, D. T. (1998). Evolutionary social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 982–1026). Boston: McGraw-Hill.

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology*, *68*, 1071–1080.

Caporael, L. R. (2001). Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology*, *52*, 607–628.

Caporael, L. R., & Baron, R. M. (1997). Groups as the mind's natural environment. In J. A Simpson & D. T. Kenrick (Eds.), *Evolutionary social psychology* (pp. 317–344). Hillsdale, NJ: Erlbaum.

Caporael, L. R., Dawes, R. M., Orbell, J. M., & Van de Kragt, A. J. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, *12*, 683–739.

Cartwright, D., & Zander, A. (1968). *Group dynamics: Research and theory* (3rd ed.). New York: Harper.

Chaiken, S., Wood, W., & Eagly, A. H. (1996). Principles of persuasion. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 702–742). New York: Guilford Press.

Chen, X. P., & Komorita, S. S. (1994). The effects of communica-

tion and commitment in a public goods dilemma. *Organizational Behavior and Human Decision Processes*, *60*, 367–386.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621.

Coser, L. A. (1962). Some functions of deviant behavior and normative flexibility. *American Journal of Sociology*, *68*, 172–181.

Costell, R. M., & Leiderman, P. H. (1968). Psychophysiological concomitants of social stress: The effects of conformity pressure. *Psychosomatic Medicine*, *30*, 28–310.

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185–236). New York: Holt.

Coull, A., Yzerbyt, V. Y, Castano, E., Paladino, M.-P., & Leemans, V. (2001). Protecting the ingroup: Motivated allocation of cognitive resources in the presence of threatening ingroup members. *Group Processes and Intergroup Relations*, *4*, 327–339.

Crano, W. D., & Chen, X. (1998). The leniency contract and persistence of majority and minority influence. *Journal of Personality and Social Psychology*, *74*, 1437–1450.

Crano, W. D., & Hannula-Bral, K. A. (1994). Context/categorization model of social influence: Minority and majority influence in the formation of a novel response norm. *Journal of Experimental Social Psychology*, *30*, 247–276.

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 504–553). Boston: McGraw-Hill.

Crocker, J., & Quinn, D. M. (2000). Social stigma and the self: Meanings, situations, and self-esteem. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma* (pp. 153–183). New York: Guilford Press.

Darke, P. R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H. P., & Hazlewood, J. D. (1998). Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin*, *24*, 1205–1215.

Darley, J. (2001). Social comparison motives in ongoing groups. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 334–351). Malden, MA: Blackwell.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review*, *80*, 97–125.

Davis, J. H. (1996). Group decision making and quantitative judgments: A consensus model. In E. H. Witte & J. H. Davis (Eds.), *Understanding group behavior, Vol. 1: Consensual action by small groups* (pp. 35–59). Hillsdale, NJ: Erlbaum.

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, *35*, 1–11.

Dawes, R. M., Van de Kragt, A. J., & Orbell, J. M. (1990). Cooperation for the benefit of us—Not me, or my conscience. In J. J. Mansbridge (Ed.), *Beyond self-interest* (pp. 97–110). Chicago: University of Chicago Press.

De Cremer, D. (2002). Respect and cooperation in social dilemmas: The importance of feeling included. *Personality and Social Psychology Bulletin*, *28*, 1335–1341.

De Cremer, D. (2003). Noneconomic motives predicting cooperation in public good dilemmas, The effect of received respect on contributions. *Social Justice Research*, *16*, 367–377.

De Cremer, D., & Bakker, M. (2003). Accountability and cooperation in social dilemmas: The influence of others' reputational concerns. *Current Psychology: Developmental, Learning, Personality, Social*, *22*, 155–163.

De Dreu, C. K. W., & De Vries, N. K. (1996). Differential processing and attitude change following majority versus minority arguments. *British Journal of Social Psychology*, *35*, 77–90.

De Dreu, C. K. W., & De Vries, N. K. (Eds.). (2001). *Group consensus and minority influence: Implications for innovation*. Oxford, UK: Blackwell.

De Dreu, C. K W., & Levine, J. M. (2006). Bridging social psychology and the organizational sciences. In P. A. M. Van Lange (Ed.), *Bridging social psychology: Benefits of transdisciplinary approaches* (pp. 341–346). Mahwah, NJ: Erlbaum.

Dedrick, D. K. (1978). Deviance and sanctioning within small groups. *Social Psychology*, *41*, 94–105.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and information social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*, 629–636.

de Vos, H., & Zeggelink, E. (1997). Reciprocal altruism in human social evolution: The viability of reciprocal altruism with a preference for "old helping partners." *Evolution and Human Behavior*, *18*, 261–278.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*, 355–391.

Diener, E. (1980). Deindividuation: The absence of self-awareness and self-regulation in group members. In P. B. Paulus (Ed.), *Psychology of group influence* (pp. 209–242). Hillsdale, NJ: Erlbaum.

Dittes, J. E., & Kelley, H. H. (1956). Effect of different conditions of acceptance upon conformity to group norms. *Journal of Abnormal and Social Psychology*, *53*, 100–107.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

Ebbesen, E. B., & Bowers, R. J. (1974). Proportion of risky to conservative arguments in a group discussion and choice shifts. *Journal of Personality and Social Psychology, 29*, 316–327.

Eidelman, S., & Biernat, M. (2003). Derogating black sheep: Individual or group protection? *Journal of Experimental Social Psychology*, *39*, 602–609.

Eidelman, S., Silvia, P. J., & Biernat, M. (2006). Responding to deviance: Target exclusion and differential devaluation. *Personality and Social Psychology Bulletin*, *32*, 1153–1164.

Eisenberger, N. I., & Lieberman, M. D. (2005). Why it hurts to be left out: The neurocognitive overlap between physical and social pain. In K. D. Williams, J. P. Forgas, J. P., & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 109–126). New York: Psychology Press.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt: An fMRI study of social exclusion. *Science*, *302*, 290–292.

Ellemers, N. (1993). The influence of socio-structural variables on identity management strategies. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 4, pp. 27–57). Chichester, UK: Wiley.

Feather, N. (1994). Attitudes toward high achievers and reactions to their fall: Theory and research concerning tall poppies. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 26, pp. 1–74). San Diego, CA: Academic Press.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137–140.

Festinger, L. (1950). Informal social communication. *Psychological Review*, *57*, 271–282.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Festinger, L., Gerard, H. B., Hymovitch, B., Kelley, H. H., & Raven, B. (1952). The influence process in the presence of extreme deviates. *Human Relations*, *5*, 327–346.

Fiske, S. T. (2004). *Social beings: A core motives approach to social psychology.* Hoboken, NJ: Wiley.

Fitness, J. (2005). Bye bye, black sheep: The causes and consequences of rejection in family relationships. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 263–276). New York: Psychology Press.

French, J. R. P., Jr., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Ann Arbor: University of Michigan Press.

Fussell, P. (2003). *The boys' crusade.* New York: Random House.

Geen, R. G. (1989). Alternative conceptions of social facilitation. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 15–51). Hillsdale, NJ: Erlbaum.

Gerard, H. B. (1961). Disagreement with others, their credibility, and experienced stress. *Journal of Abnormal and Social Psychology*, *62*, 559–564.

Gerard, H. B., & Rotter, G. S. (1961). Time perspective, consistency of attitude and social influence. *Journal of Abnormal and Social Psychology*, *62*, 565–572.

Giordano, P. C. (1983). Sanctioning the high-status deviant: An attributional analysis. *Social Psychology Quarterly*, *46*, 329–342.

Goethals, G. R., & Darley, J. M. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 259–278). Washington, DC: Hemisphere.

Gonsalkorale, K., & Williams, K. D. (in press). The KKK won't let me play: Ostracism by a despised outgroup hurts. *European Journal of Social Psychology*.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety*. New York: Oxford University Press.

Groenenboom, A., Wilke, H. A. M., & Wit, A. P. (2001). Will we be working together again? The impact of future interdependence on group members' task motivation. *European Journal of Social Psychology*, *31*, 369–378.

Gully, S. M., Devine, D. J., &Whitney, D. J. (1995). A meta analysis of cohesion and performance: Effects of levels of analysis and task interdependence. *Small Group Research*, *26*, 497–520.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 3, The interpersonal context* (pp. 28–84). New York: Guilford Press.

Harkins, S. G., & Petty, R. E. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology*, *43*, 1214–1229.

Harkins, S. G., & Szymanski, K. (1987). Social loafing and social facilitation: New wine in old bottles. In C. Hendrick (Ed.), *Group processes and intergroup relations* (pp. 167–188). Newbury Park, CA: Sage.

Harlow, H. F., Harlow, M. K., & Suomi, S. J. (1971). From thought to therapy: Lessons from a primate laboratory. *American Scientist*, *59*, 538–549.

Hart, C. M., & Van Vugt, M. (2006). From fault line to group fission: Understanding membership changes in small groups. *Personality and Social Psychology Bulletin, 32*, 392–404.

Harvey, O. J., & Consalvi, C. (1960). Status and conformity to pressure in informal groups. *Journal of Abnormal and Social Psychology*, *60*, 182–187.

Haslam, S. A. (2004). *Psychology in organizations* (2nd ed.). London: Sage.

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, *112*, 494–508.

Hastie, R., Penrod, S. D., & Pennington, N. (1983). *Inside the jury*. Cambridge, MA: Harvard University Press.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Hawkins, C. (1962). Interaction rates of jurors aligned in factions. *American Sociological Review*, *27*, 689–691.

Hayashi, N., & Yamagishi, T. (1998). Selective play, Choosing partners in an uncertain world. *Personality and Social Psychology Review*, *2*, 276–289.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511–524.

Hensley, V., & Duval, S. (1976). Some perceptual determinants of perceived similarity, liking, and correctness. *Journal of Personality and Social Psychology*, *34*, 159–168.

Hertel, G., & Kerr, N. L. (2001). Priming and in-group favoritism: The impact of normative scripts in the minimal group paradigm. *Journal of Experimental Social Psychology*, *37*, 316–324.

Hinsz, V. B. (1990). Cognitive and consensus processes in group recognition memory performance. *Journal of Personality and Social Psychology*, *59*, 705–718.

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology*, *11*, 223–255.

Hogg, M. A. (2005). All animals are equal but some animals are more equal than others: Social identity and marginal member-

ship. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 243–261). New York: Psychology Press.

Hogg, M. A., & Abrams, D. (1993). Towards a single-process uncertainty-reduction model of social motivation in groups. In M. A. Hogg & D. Abrams (Eds.), *Group motivation: Social psychological perspectives* (pp. 173–190). London: Harvester-Wheatsheaf.

Hogg, M. A., Fielding, K. S., & Darley, J. (2005). Fringe dwellers: Processes of deviance and marginalization in groups. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 191–210). New York: Psychology Press.

Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review, 65*, 117–127.

Hollander, E. P. (1960). Competence and conformity in the acceptance of influence. *Journal of Abnormal and Social Psychology, 61*, 365–369.

Hollander, E. P., & Willis, R. H. (1967). Some current issues in the psychology of conformity and nonconformity. *Psychological Bulletin, 68*, 62–76.

Hornsey, M. J., & Jetten, J. (2003). Not being what you claim to be: Imposters as sources of group threat. *European Journal of Social Psychology, 33*, 639–657.

Hornsey, M. J., & Jetten, J. (2004). The individual within the group: Balancing the need to belong with the need to be different. *Personality and Social Psychology Review, 8*, 248–264.

Hutchison, P., & Abrams, D. (2003). Ingroup identification moderates stereotype change in reaction to ingroup deviance. *European Journal of Social Psychology, 33*, 497–506.

Insko, C. A., Drenan, S., Solomon, M. R., Smith, R., & Wade, T. J. (1983). Conformity as a function of the consistency of positive self-evaluation with being liked and being right. *Journal of Experimental Social Psychology, 19*, 341–358.

Israel, J. (1956). *Self-evaluation and rejection in groups: Three experimental studies and a conceptual outline*. Uppsala, Sweden: Almqvist & Wiksell.

Jackson, J. M., & Williams, K. D. (1985). Social loafing on difficult tasks: Working collectively can improve performance. *Journal of Personality and Social Psychology, 49*, 937–942.

Jackson, L. A., Sullivan, L. A., Harnish, R., & Hodge, C. N. (1996). Achieving positive social identity: Social mobility, social creativity, and permeability of group boundaries. *Journal of Personality and Social Psychology, 70*, 241–254.

Janis, I. L. (1982). *Groupthink* (2nd ed.). Boston: Houghton Mifflin.

Jones, E. E., & deCharms, R. (1957). Changes in social perception as a function of the personal relevance of behavior. *Sociometry, 20*, 75–85.

Jones, E. E., & Gerard, H. B. (1967). *Foundations of social psychology*. New York: Wiley.

Juvonen, J., & Gross, E. F. (2005). The rejected and the bullied: Lessons about social misfits from developmental psychology. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 155–170). New York: Psychology Press.

Kalven, H., & Zeisel, H. (1966). *The American jury*. Boston: Little, Brown.

Kameda, T., Takezawa, M., & Hastie, R. (2003). The logic of social sharing: An evolutionary game analysis of adaptive norm development. *Personality and Social Psychology Review, 7*, 2–19.

Kaplan, M. F. (1989). Task, situational, and personal determinants of influence processes in group decision making. In E. J. Lawler & B. Markovsky (Eds.), *Advances in group processes* (pp. 87–105). Greenwich CT: JAI Press.

Kaplan, M. F., & Miller, C. E. (1987). Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology, 53*, 306–313.

Kaplan, M. F., & Wilke, H. (2001). Cognitive and social motivation in group decision making. In J. P. Forgas, K. D. Williams, & L. Wheeler (Eds.), *The social mind: Cognitive and motivational aspects of interpersonal behavior* (pp. 406–428). New York: Cambridge University Press.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*, 681–706.

Kelley, H. H. (1952). Two functions of reference groups. In G. E. Swanson, T. Newcomb, & E. Hartley (Eds.), *Readings in social psychology* (pp. 410–414). New York: Holt.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution, 2*, 51–60.

Kelman, H. C. (1961). Processes of opinion change. *Public Opinion Quarterly, 25*, 57–78.

Kerr, N. L. (1981). Social transition schemes: Charting the group's road to agreement. *Journal of Personality and Social Psychology, 41*, 684–702.

Kerr, N. L. (1983). Motivation losses in task-performing groups: A social dilemma analysis. *Journal of Personality and Social Psychology, 45*, 819–828.

Kerr, N. L. (1986). Motivational choices in task groups: A paradigm for social dilemma research. In H. A. M. Wilke, D. M. Messick, & C. G. Rutte (Eds.), *Experimental social dilemmas* (pp. 4–27). Frankfurt am Main: Lang GmbH.

Kerr, N. L. (1992). Group decision making at a multialternative task: Extremity, interfaction distance, pluralities, and issue importance. *Organizational Behavior and Human Decision Processes, 52*, 64–95.

Kerr, N. L. (1995). Norms in social dilemmas. In D. Schroeder (Ed.), *Social dilemmas: Perspectives on individuals and groups* (pp. 31–47). Westport, CT: Praeger.

Kerr, N. L. (1999). Anonymity and social control in social dilemmas. In M. Foddy, M. Smithson, S. Schneider, & M. A. Hogg (Eds.), *Resolving social dilemmas* (pp. 103–119). Philadelphia: Psychology Press.

Kerr, N. L. (2001). Is it what one says or how one says it?: Style vs. substance from an SDS perspective. In C. D. W. De Dreu & N. K. De Vries (Eds.), *Group consensus and minority influence: Implications for innovation* (pp. 201–228). Oxford, UK: Blackwell.

Kerr, N. L., Atkin, R., Stasser, G., Meek, D., Holt, R., & Davis, J. H. (1976). Guilt beyond a reasonable doubt: Effects of concept definition and assigned decision rule on the judgments of mock jurors. *Journal of Personality and Social Psychology, 34*, 282–294.

Kerr, N. L., & Bruun, S. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*, 78–94.

Kerr, N. L., Davis, J. H., Meek, D., & Rissman, A. (1975). Group position as a function of member attitudes: Choice shift effects from the perspective of social decision scheme theory. *Journal of Personality and Social Psychology, 35*, 574–593.

Kerr, N. L., Harris, D., Messé, L. M., Poulsen, J., & Seok, D. H. (2005). *Social ostracism by one's coworkers II: Does rejection affect the Köhler motivation gain?* Unpublished manuscript, Michigan State University.

Kerr, N. L., & Kaufman-Gilliland, C. M. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology, 48*, 349–363.

Kerr, N. L., & MacCoun, R. J., (1985). The effects of jury size and polling method on the process and product of jury deliberation. *Journal of Personality and Social Psychology, 48*, 349–363.

Komorita, S. S., Parks, C. D., & Hulbert, L. G. (1992). Reciprocity and the induction of cooperation in social dilemmas. *Journal of Personality and Social Psychology, 62*, 607–617.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology, 46*, 1044–1057.

Kruglanski, A. (1989). *Lay epistemics and human knowledge: Cognitive and motivational biases*. New York: Plenum Press.

Kruglanski, A. W., & Webster, D. M. (1991). Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of environmental noise. *Journal of Personality and Social Psychology, 61*, 212–225.

Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin*, *127*, 187–208.

Landsberger, H. A. (1958). *Hawthorne revisited*. Ithaca, NY: Cornell University Press.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*, 822–832.

Laughlin, P. R. (1980). Social combination processes of cooperative problem-solving groups on verbal intellective tasks. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1, pp. 127–155). Hillsdale, NJ: Erlbaum.

Laughlin, P. R. (1999). Collective induction: Twelve postulates. *Organizational Behavior and Human Decision Processes*, *80*, 50–69.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology*, *22*, 177–189.

Laughlin, P. R., Kerr, N. L., Davis, J. H., Halff, H. M., & Marciniak, K. A. (1975). Group size, member ability, and social decision schemes on an intellective task. *Journal of Personality and Social Psychology*, *31*, 522–535.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 1–62). San Diego, CA: Academic Press.

Lemaine, G. (1974). Social differentiation and social originality. *European Journal of Social Psychology*, *4*, 17–52.

Levine, J. M. (1989). Reaction to opinion deviance in small groups. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 187–231). Hillsdale, NJ: Erlbaum.

Levine, J. M., & Higgins, E. T. (2001). Shared reality and social influence in groups and organizations. In F. Butera & G. Mugny (Eds.), *Social influence in social reality: Promoting individual and social change* (pp. 33–52). Bern, Switzerland: Hogrefe & Huber.

Levine, J. M., & Kaarbo, J. (2001). Minority influence in political decision-making groups. In C. K. W. De Dreu & N. K. De Vries (Eds.), *Group consensus and minority influence: Implications for innovation* (pp. 229–257). Oxford, UK: Blackwell.

Levine, J. M., & Moreland, R. L. (1987). Social comparison and outcome evaluation in group contexts. In W. P. Smith & J. C. Masters (Eds.), *Social comparison, social justice, and relative deprivation, Theoretical, empirical, and policy perspectives*. (pp. 105–127). Hillsdale, NJ: Erlbaum.

Levine, J. M., & Moreland, R. L. (1994). Group socialization: Theory and research. *European Review of Social Psychology*, *5*, 305–336.

Levine, J. M., & Moreland, R. L. (1998). Small groups. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 415–469). Boston: McGraw-Hill.

Levine, J. M., & Moreland, R. L. (2002). Group reactions to loyalty and disloyalty. In S. R. Thye & E. J. Lawler (Eds.), *Group cohesion, trust and solidarity: Advances in group processes* (Vol. 19, pp. 203–228). Oxford, UK: Elsevier Science.

Levine, J. M., Moreland, R. M., & Hausmann, L. (2005). Managing group composition: Inclusive and exclusive role transitions. In D. Abrams, M. A. Hogg, & J M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 137–160). New York: Psychology Press.

Levine, J. M., & Ranelli, C. J. (1978). Majority reaction to shifting and stable attitudinal deviates. *European Journal of Social Psychology*, *8*, 55–70.

Levine, J. M., & Ruback, R. B. (1980). Reaction to opinion deviance: Impact of a fence straddler's rationale on majority evaluation. *Social Psychology Quarterly*, *43*, 73–81.

Levine, J. M., Saxe, L., & Harris, H. J. (1976). Reaction to attitudinal deviance: Impact of deviate's direction and distance of movement. *Sociometry*, *39*, 97–107.

Levine, J. M., Sroka, K. R., & Snyder, H. N. (1977). Group support and reaction to stable and shifting agreement/disagreement. *Sociometry*, *40*, 214–224.

Levine, J. M., & Thompson, L. (1996). Conflict in groups. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 745–776). New York: Guilford Press.

Lott, A. J., & Lott, B. E. (1961). Group cohesiveness, communication level, and conformity. *Journal of Abnormal and Social Psychology*, *62*, 408–412.

Luhtanen, R., & Crocker, J. (1991). Self-esteem and intergroup comparisons: Toward a theory of collective self-esteem. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 211–236). Hillsdale, NJ: Erlbaum.

Maass, A., Clark, R. D., III, & Haberkorn, G. (1982). The effects of differential ascribed category membership and norms on minority influence. *European Journal of Social Psychology*, *12*, 89–104.

MacDonald, G., & Kingsbury, R. (2006). Does physical pain augment anxious attachment? *Journal of Social and Personal Relationships*, *23*, 291–304.

MacDonald, G., Kingsbury, R., & Shaw, S. (2005). Adding insult to injury: Social pain theory and response to social exclusion. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 77–90). New York: Psychology Press.

MacDonald, G., & Leary, M. R. (2005). Why does social exclusion hurt?: The relationship between social and physical pain. *Psychological Bulletin*, *131*, 202–223.

Mackie, D. M. (1987). Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology*, *53*, 41–52.

Mackie, D. M., & Goethals, G. R. (1987). Individual and group goals. In C. Hendrick (Ed.), *Group processes* (pp. 144–166). Newbury Park, CA: Sage.

Mackie, D. M., & Queller, S. (2000). The impact of group membership on persuasion: Revisiting "who says what to whom with what effect?." In D. J. Terry & M. A. Hogg (Eds.), *Attitudes, behavior, and social context: The role of norms and group membership* (pp. 135–155). Mahwah, NJ: Erlbaum.

Major, B., & Eccleston, C. P. (2005). Stigma and social exclusion. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 63–87). New York: Psychology Press.

Marques, J. M., Abrams, D., Paez, D., & Hogg, M. A. (2001). Social categorization, social identification, and rejection of deviant group members. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 400–424). Malden, MA: Blackwell.

Marques, J. M., Abrams, D., Paez, D., & Martinez-Taboada, C. (1998). The role of categorization and in-group norms in judgments of groups and their members. *Journal of Personality and Social Psychology*, *75*, 976–988.

Marques, J. M., Abrams, D., & Serodio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*, 436–447.

Marques, J. M., & Paez, D. (1994). The "Black Sheep Effect": Social categorization, rejection of ingroup deviates, and perception of group variability. *European Review of Social Psychology*, *5*, 37–68.

Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology*, *18*, 287–292.

Marques, J. M., Yzerbyt, V. Y., & Leyens, J.-P. (1988). The "Black Sheep Effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*, 1–16.

Martin, R., & Hewstone, M. (2001). Conformity and independence in groups: Majorities and minorities. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 209–234). Malden, MA: Blackwell.

Marwell, G., & Ames, R. E. (1979). Experiments on the provision of public goods, I, Resources, interest, group size, and the free rider problem. *American Journal of Sociology*, *84*, 1335–1360.

Masclet, D., Noussair, C., Tucker, S., & Villeval, M. (2003). Mone-

tary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review*, *93*, 366–380.

Matheson, K., Cole, B., & Majka, K. (2003). Dissidence from within: Examining the effects of intergroup context on group members' reactions to attitudinal opposition. *Journal of Experimental Social Psychology*, *39*, 161–169.

McAuliffe, B. J., Jetten, J., Hornsey, M. J., & Hogg, M. A. (2003). Individualist and collectivist group norms: When its OK to go your own way. *European Journal of Social Psychology*, *33*, 57–70.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

Mikulincer, M., & Shaver, P. R. (in press). *Attachment processes in adulthood*. New York: Guilford Press.

Miller, C. E., & Anderson, P. D. (1979). Group decision rules and the rejection of deviates. *Social Psychology Quarterly*, *42*, 354–363.

Moreland, R. L., & Levine, J. M. (1982). Socialization in small groups: Temporal changes in individual-group relations. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 15, pp. 137–192). New York: Academic Press.

Morris, W. N., & Miller, R. S. (1975). Impressions of dissenters and conformers: An attributional analysis. *Sociometry*, *38*, 327–339.

Moscovici, S. (1976). *Social influence and social change.* London: Academic Press.

Moscovici, S., & Lage, E. (1978). Studies in social influence IV: Minority influence in a context of original judgments. *European Journal of Social Psychology*, *8*, 349–365.

Mugny, G., & Perez, J. A. (1991). *The social psychology of minority influence* (V. W. Lamongie, Trans.). Cambridge, UK: Cambridge University Press.

Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin*, *115*, 210–227.

Myers, D. G. (1982). Polarizing effects of social interaction. In H. Brandstatter, J. H. Davis, & G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 125–161). London: Academic Press.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*, 602–27.

Nail, P. R. (1986). Toward an integration of some models and theories of social response. *Psychological Bulletin*, *100*, 190–206.

Neuberg, S. L., Smith, D. M., & Asher, T. (2000). Why people stigmatize: Toward a biocultural framework. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma* (pp. 31–61). New York: Guilford Press.

Noel, J. G., Wann, D. L., & Branscombe, N. R. (1995). Peripheral ingroup membership status and public negativity toward outgroups. *Journal of Personality and Social Psychology*, *68*, 127–137.

Orbell, J., & Dawes, R. (1981). Social dilemmas. In G. Stephenson & J. H. Davis (Eds.), *Progress in applied social psychology* (Vol. 1, pp. 37–66). Chichester, UK: Wiley.

Orcutt, J. D. (1973). Societal reaction and the response to deviation in small groups. *Social Forces*, *52*, 259–267.

Ouwerkerk, J. W., Kerr, N. L., Gallucci, M., & van Lange, P. A. M. (2005). Avoiding the social death penalty: Ostracism and cooperation in social dilemmas. In K. D. Williams, J. P. Forgas, J. P., & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 321–332). New York: Psychology Press.

Paichler, G. (1976). Norms and attitude change I: Polarization and styles of behaviour. *European Journal of Social Psychology*, *6*, 405–427.

Panksepp, J., Siviy, S., & Normansell, L. A. (1985). Brain opiods and social emotion. In M. Reite & T. Field (Eds.), *The psychobiology of attachment and separation* (pp. 3–49). New York: Academic Press.

Parks, C. D., & Nelson, N. L. (1999). Discussion and decision: The interrelationship between initial preference distribution and group discussion content. *Organizational Behavior and Human Decision Processes*, *80*, 87–101.

Pickett, C. L., & Gardner, W. L. (2005). The Social Monitoring System: Enhanced sensitivity to social cues as an adaptive response to social exclusion. In K. D. Williams, J. P. Forgas, J. P., & W. von

Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 213–226). New York: Psychology Press.

Postmes, T., & Spears, R. (1998). Deindividuation and anti-normative behavior: A meta-analysis. *Psychological Bulletin*, *123*, 238–259.

Price, M. E., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior*, *23*, 203–231.

Prislin, R., Levine, J. M., & Christensen, P. N. (2006). When reasons matter: Quality of support affects reactions to increasing and consistent agreement. *Journal of Experimental Social Psychology*, *42*, 593–601.

Pruitt, D. G. (1971). Choice shifts in group discussion: An introductory review. *Journal of Personality and Social Psychology*, *20*, 339–360.

Raven, B. H., & Kruglanski, A. (1970). Conflict and power. In P. Swingle (Ed.), *The structure of conflict* (pp. 69–109). New York: Academic Press.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.

Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Romero-Canyas, R., & Downey, G. (2005). Rejection sensitivity as a predictor of affective and behavioral responses to interpersonal stress: A defensive motivational system. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 131–154). New York: Psychology Press.

Sakurai, M. M. (1975). Small group cohesiveness and detrimental conformity. *Sociometry*, *38*, 340–357.

Sani, F., & Reicher, S. (1998). When consensus fails: An analysis of the schism within the Italian Communist Party. *European Journal of Social Psychology*, *28*, 623–645.

Sattler, D. N. (1998). The need principle in social dilemmas. *Journal of Social Behavior and Personality*, *13*, 667–678.

Schachter, S. (1951). Deviation, rejection, and communication. *Journal of Abnormal and Social Psychology*, *46*, 190–207.

Schachter, S., Ellertson, N., McBride, D., & Gregory, D. (1951). An experimental study of cohesiveness and productivity. *Human Relations*, *4*, 229–238.

Scheepers, D., Branscombe, N. R., Spears, R., & Doosje, B. (2002). The emergence and effects of deviants in low and high status groups. *Journal of Experimental Social Psychology*, *38*, 611–617.

Scheff, T. (1966). *Being mentally ill: A sociological theory*. Chicago: Aldine.

Schlesinger, A. M., Jr. (1965). *A thousand days*. Boston: Houghton-Mifflin.

Schmitt, M. T., & Branscombe, N. R. (2001). The good, the bad, and the manly: Threats to one's prototypicality and evaluations of fellow in-group members. *Journal of Experimental Social Psychology*, *37*, 510–517.

Schmitt, M. T., Silvia, P. J., & Branscombe, N. R. (2000). The intersection of self-evaluation maintenance and social identity theories: Intragroup judgment in interpersonal and intergroup contexts. *Personality and Social Psychology Bulletin*, *26*, 1598–1606.

Schulz-Hardt, S., Frey, D., Luethgens, C., & Moscovici, S. (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology*, *78*, 655–669.

Seashore, S. E. (1954). *Group cohesiveness in the industrial work group*. Ann Arbor, MI: Institute for Social Research.

Simpson, J. A., & Rholes, W. S. (Eds.). (1998). *Attachment theory and close relationships*. New York: Guilford Press.

Smith, E. R., Murphy, J., & Coats, S. (1999). Attachment to groups: Theory and measurement. *Journal of Personality and Social Psychology*, *77*, 94–110.

Stangor, C., & Crandall, C. S. (2000). Threat and social construction of stigma. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma* (pp. 62–87). New York: Guilford Press.

Stasser, G. (1999). A primer of social decision scheme theory:

Models of group influence, competitive model-testing, and prospective modeling. *Organizational Behavior and Human Decision Processes*, *80*, 3–20.

Stasser, G., Kerr, N. L., & Davis, J. H. (1989). Influence processes and consensus models in decision-making groups. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 279–326). Hillsdale, NJ: Erlbaum.

Stouffer, S. A. (1949). *The American soldier.* Princeton, NJ: Princeton University Press.

Suchner, R. W., & Jackson, D. (1976). Responsibility and status: A causal or only a spurious relationship? *Sociometry*, *39*, 243–256.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Chicago: Nelson Hall.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). San Diego, CA: Academic Press.

Tesser, A., Campbell, J., & Mickler, S. (1983). The role of social pressure, attention to the stimulus, and self-doubt in conformity. *European Journal of Social Psychology*, *13*, 217–234.

Thibaut, J. W., & Strickland, L. H. (1956). Psychological set and social conformity. *Journal of Personality*, *25*, 115–129.

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature*, *4*, 237–269.

Thornton, H., Menendez, E., Wrightsman, L. S., Posey, A., & Scheflin, A. W. (1995). *Hung jury: The diary of a Menendez juror.* Philadelphia: Temple University Press.

Timothy, M. (1975). *Jury woman.* Volcano, CA: Volcano Press.

Tindale, R. S., Meisenhelder, H. M., Dykema-Engblade, A. A., & Hogg, M. A. (2001). Shared cognition in small groups. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 1–30). Malden, MA: Blackwell.

Torrance, E. P. (1954). The behavior of small groups under the stress of conditions of survival. *American Sociological Review*, *19*, 751–755.

Turner, J. C. (1991). *Social influence.* Pacific Grove, CA: Brooks/Cole.

Turner, J. C., & Oakes, P. J. (1989). Self-categorization theory and social influence. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 233–275). Hillsdale, NJ: Erlbaum.

Twenge, J. M., & Baumeister, R. F. (2005). Social exclusion increases aggression and self-defeating behavior while reducing intelligent thought and prosocial behavior. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 27–46). New York: Psychology Press.

Twenge, J. M., Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Bartels, J. M. (in press). Social exclusion reduces prosocial behavior. *Journal of Personality and Social Psychology*.

Twenge, J. M., Catanese, K. R., & Baumeister, R. F. (2003). Social exclusion and the deconstructed state: Time perception, meaninglessness, lethargy, lack of emotion, and self awareness. *Journal of Personality and Social Psychology*, *85*, 409–423.

Tyler, T. R., & Degoey, P. (1995). Collective restraint in social dilemmas: Procedural justice and social identification effects on supports of authorities. *Journal of Personality and Social Psychology*, *69*, 482–497.

van Dijk, E., & Vermunt, R. (2000). Strategy and fairness in social decision making: Sometimes it pays to be powerless. *Journal of Experimental Social Psychology*, *36*, 1–25.

Van Vugt, M., & Hart, C. M (2004). Social identity as social glue: The origins of group loyalty. *Journal of Personality and Social Psychology*, *86*, 585–598.

Van Vugt, M., Jepson, S. F., Hart, C. M., & De Cremer, D. (2004). Autocratic leadership in social dilemmas: A threat to group stability. *Journal of Experimental Social Psychology*, *40*, 1–13.

Van Vugt, M., & Van Lange, P. A. M. (2006). The altruism puzzle: Psychological adaptations for prosocial behavior. In M. Schaller, J. A. Simpson, & D. T. Kenrick (Eds.), *Evolution and social psychology* (pp. 237–261). New York: Psychology Press.

Wahrman, R. (1970). High status, deviance and sanctions. *Sociometry*, *33*, 485–504.

Walker, E. L., & Heyns, R. W. (1962). *An anatomy for conformity.* Englewood Cliffs, NJ: Prentice-Hall.

Wheeler, L. (1966). Toward a theory of behavioral contagion. *Psychological Review*, *73*, 179–192.

Wheeler, L., & Caggiula, A. R. (1966). The contagion of aggression. *Journal of Experimental Social Psychology*, *2*, 1–10.

Wiggins, J. A., Dill, F., & Schwartz, R. D. (1965). On "status liability." *Sociometry*, *28*, 197–209.

Williams, K. D. (2001). *Ostracism: The power of silence.* New York: Guilford Press.

Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, *58*, 15.1–15.28.

Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*, 748–762.

Williams, K. D., Forgas, J. P., & von Hippel, W. (Eds.). (2005). *The social outcast: Ostracism, social exclusion, rejection, and bullying.* New York: Psychology Press.

Williams, K. D., & Govan, C. L. (2005). Reacting to ostracism: Retaliation or reconciliation? In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (pp. 47–62). New York: Psychology Press.

Williams, K. D., & Zadro, L. (2005). Ostracism: The indiscriminate early detection system. In K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), *The social outcast: Ostracism, social exclusion, rejection, and bullying* (pp. 19–34). New York: Psychology Press.

Willis, R. H. (1965). Conformity, independence, and anticonformity. *Human Relations*, *18*, 373–388.

Wills, T. A. (1991). Similarity and self-esteem in downward comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 51–78). Hillsdale, NJ: Erlbaum.

Wit, A. P., Wilke, H. A. M., & Oppewal, H. (1992). Fairness in asymmetric social dilemmas. In W. Liebrand, D. M. Messick, & H. A. M. Wilke (Eds.), *Social dilemmas, Theoretical issues and research findings* (pp. 183–197). New York: Pergamon Press.

Wood, W. (1999). Motives and modes of processing in the social influence of groups. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 547–570). New York: Guilford Press.

Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, *115*, 323–345.

Yamagishi, T., Hayashi, N., & Jin, N. (1994). Prisoner's dilemma networks: Selection strategy versus action strategy. In U. Schulz, W. Albers, & U. Mueller (Eds.), *Social dilemmas and cooperation* (pp. 233–250). Berlin: Springer-Verlag.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274.

Zdaniuk, B., & Levine, J. M. (2001). Group loyalty: Impact of members' identification and contributions. *Journal of Experimental Social Psychology*, *37*, 502–509.

CHAPTER 34

Cultural Processes

Basic Principles

CHI-YUE CHIU
YING-YI HONG

Culture exhibits the way people interpret their biology and environment. In public discourse, "culture" has also been given a tremendous amount of authority over people's act and thought (Rothstein, 1999). In anthropology, the concept of "culture" has been compared to the concepts of gravity in physics, and disease in medicine (Kroeber, Kluckholm, & Untereiner, 2001). Recently, psychologists have invoked the concept of "culture" to understand national and regional differences in a wide range of phenomena, from attention (Masuda & Nisbett, 2001) to categorization (Ji, Zhang, & Nisbett, 2004), thinking style (Peng & Nisbett, 1999), attributions (Menon, Morris, Chiu, & Hong, 1999; Morris & Peng, 1994), self-construals (Heine, Lehman, Markus, & Kitayama, 1999), self-regulatory focus (Lee, Aaker, & Gardner, 2000), prediction of future events (Ji, Nisbett, & Su, 2001), choice and motivation (Iyengar & Lepper, 1999), emotion (Cohen & Gunz, 2002), and life satisfaction (Oishi, Wyer, & Colcombe, 2000; Suh, Diener, Oishi, & Triandis, 1998). Contemporary reviews of the extant literature (Fiske, Kitayama, Markus, & Nisbett, 1995; Lehman, Chiu, & Schaller, 2004; Markus, Kitayama, & Heiman, 1996) have produced an impressive catalog of cultural differences in psychological processes and their outcomes.

However, a social psychology of culture is at best incomplete if it deals with cultural differences only. Studying cultural differences in psychological processes and their outcomes provides a litmus test of the generality of the basic principles in social psychology—basic principles should be able to accommodate and explain cul-

tural variations in behavior (Norenzayan & Heine, 2005). Nonetheless, culture does not explain cultural differences (see Bickhard & Campbell, 2000); what needs to be explained is culture itself. Principles of cultural processes are needed to account for the production, reproduction, and evolution of culture.

In this chapter, instead of providing a catalog of cultural differences, we attempt to construct a social psychology of cultural processes. We take a knowledge perspective to culture, and define culture as a network of distributed knowledge that is produced and reproduced among a collection of interconnected individuals. Our goal is to explicate the basic principles underlying how cultural knowledge is represented in the mind and applied in concrete situations, and how it emerges, spreads, and evolves. We begin with an overview of the key questions and basic premises in the knowledge perspective and proceed to expound on the specific principles.

KEY QUESTIONS AND BASIC PREMISES

By taking a knowledge perspective to culture, we seek to answer four interrelated questions. First, what is the nature of culture? As noted, we define culture as a network of distributed knowledge. However, how does cultural knowledge differ from personal knowledge? Is culture a coherent knowledge system? What forms can cultural knowledge take? Second, what are the social cognitive principles that govern the use of cultural knowledge in concrete situations? Third, what is culture for? What are

the societal and psychological functions of culture? Finally, how does culture spread and change? In this chapter, we address each question in turn.

We acknowledge that there are profound cultural differences in human behaviors. However, our goal is to explain cultural variations in terms of universal cultural processes. In psychology, it is customary to treat culture as contents (e.g., different kinds of self-construal, values, and beliefs). Researchers generally agree that individuals adhering to different cultural knowledge traditions may exhibit markedly different behaviors. However, there is disagreement over whether the same psychological principles can account for such cultural variations (see Hong & Chiu, 2001). Cultural relativists hold that because culture shapes the fundamental schemas used for grasping experiences, qualitatively different systems of psychological knowledge are needed to understand behaviors in specific cultures. We accept the premise that behaviors vary across cultures but reject the notion that culture creates incommensurate psychological systems. Most social psychologists believe in the power of situation. However, few see the need to develop different systems of social psychology for different kinds of situations. Instead, a common goal in social psychology is to uncover basic principles that explain situational variations in behaviors. Following the intellectual tradition in social psychology, we seek to identify the basic principles that underlie cultural variations in psychological phenomena.

We submit that this course of inquiry can contribute to the construction of a more complete body of social psychological knowledge. On the one hand, this course of inquiry may extend the generality of *some* existing psychological principles (e.g., principles of knowledge activation; Higgins, 1996) by illustrating how these principles also account for cultural variations in certain social psychological phenomena. On the other hand, this course of inquiry may also identify the boundary conditions of and suggest modifications to other principles (e.g., principles of self-enhancement). Furthermore, individuals have at their disposal both personal and cultural knowledge. These two kinds of knowledge differ in many respects (see below). Studying how cultural (vs. personal) knowledge is called out in concrete situations may inspire construction of new basic principles.

In short, our primary goal is not to advocate a distinct body of knowledge to account for behaviors in a particular cultural group but to use cultural variations as a source of inspirations for uncovering and revising basic principles, and as a litmus test for the generality of existing principles. Table 34.1 provides a summary of what we consider to be the basic principles of cultural processes.

WHAT IS CULTURE?

A Knowledge Perspective to Culture

We take a knowledge perspective to culture, and define culture as *a network of knowledge shared (albeit incompletely) among a collection of interconnected individuals*. Knowledge refers to all the ways of understanding that we use to make up our experienced, grasped reality. It consists of a

TABLE 34.1. DEFINITIONAL ISSUES AND BASIC PRINCIPLES

Definition: *Culture consists of a network of knowledge shared (albeit incompletely) among a collection of interconnected individuals.*

D.1 Cultural knowledge is distributed in a group of interconnected individuals, and is instantiated or communicated in one or several media, and distributed or communicated within a series of instituted social relations.

D.2 Items of cultural knowledge are domain-specific with loose interconnectedness.

D.3 Two major types of cultural knowledge are procedural knowledge and declarative knowledge; declarative knowledge can be further subdivided into person representations, event representations, and norm representations.

D.3 Cultural knowledge is conceptually different from cultural identification.

Describing cultural differences

P.1 Cultural differences can be understood in terms of the differing contents of the widely circulated and highly accessible procedural and declarative knowledge in different cultures.

P.1a Procedural knowledge: A learned response sequence (procedural knowledge) becomes automated through frequent practices; once a production is automated, its performance requires little cognitive deliberation.

P.1b Declarative knowledge: Chronic declarative cultural knowledge provides premises for subsequent judgments and behaviors.

P.2 Cultures also differ in how basic psychogenic needs, cultural knowledge, and personal experiences work together. Some cultures emphasize personal experiences, whereas others emphasize conformance to cultural norms. A behavior mediated by personal attitudes in one culture may be mediated by cultural norms in another culture. Self-motives and cultural norms may act in concert to produce the same behavior in some cultures, and push behaviors into opposite directions in others.

Cultural knowledge activation

P.3 People in every culture have constructed different cognitive representations of the people, events and norms. Cultural experiences determine the relative accessibility of these representations. Prevalence of cultural practices supporting certain representation renders it highly accessible.

P.4 Cultural knowledge items that have been frequently used in a group are usually widely shared, more frequently reproduced in communication, widely represented in external or public carriers of culture, and cognitively accessible to members of the group.

P.5 Contextual cues may increase the temporary accessibility of a body of knowledge, and momentarily raise the probability that this body of knowledge will be applied.

P.6 The probability that a knowledge item will be applied also depends on its applicability in the immediate context.

P.7 People are likely to use cultural knowledge when the situation calls for a readily available or conventionalized solution, or when the problem solver lacks the capability, motivation, or resource to consider alternative solutions.

P.8 People who are primed with a social category tend to assimilate their response into the primed category. However, if the individuals feel that they do not belong to the primed culture, culture priming may lead to contrastive responses.

TABLE 34.1. *(continued)*

Production, maintenance, and reproduction of culture

P.9   Cultural knowledge was originally developed to solve coordination problems in collective living.

P.10 The shared physical, social, and motivational environment in a culture affords opportunities to develop procedural knowledge through repeated practices of certain cognitive procedure.

P.11 Instantiation of cultural knowledge in external media enables ratcheting, and turns these media into effective means for the acquisition, transmission, activation, and maintenance of cultural knowledge.

P.12 Interpersonal communication may lead to spatially differentiation and clustering of attitudes and beliefs, and correlation of previously unconnected values and beliefs.

P.13 Through interpersonal communication, cultural knowledge that is widely shared among group members would likely be reproduced and consolidated in communication.

P.14 A cultural knowledge tradition is maintained partly because Individuals need culture to reduce uncertainty in social living, to manage existential terror, and fulfill the need for belongingness.

P.15 Culture change may occur because of endogenous and exogenous reasons. Endogenous culture change may take place through the processes of differentiation. Exogenous culture change often results from intercultural contacts.

set of learned routines of thinking, feeling, and interacting with other people, as well as a corpus of substantive assertions and ideas about aspects of the world (Barth, 2002).

For decades, anthropologists have debated what culture is. Some anthropologists (Braumann, 1999; Shore, 1996; Sperber, 1996) now believe that culture is a complex web of shifting meanings that connect people in a physical or virtual locale. Early 20th-century scholars understood culture to refer to a pattern of fixed traits (e.g., national or modal personalities) and assumed that such fixed traits had clear bounds. In the 1990s, some anthropologists became critical of the common practice of treating culture as a static monolith and using it synonymously with a demarcated population. In their opinion, through this practice, an inordinate degree of boundedness, homogeneity, coherence, and stability is attributed to the concept of culture (Appadurai, 1996; Friedman, 1994). In addition, the search for discrete cultural patterns may lead to reification of cultures and essentialism (Keesing, 1994). Similar objections were registered in psychology (Bandura, 2002). To escape from the predicament their discipline is facing, some anthropologists declared that "the culture concept has served its time" (Clifford, 1988, p. 274), while others even rallied their colleagues to write against culture (Abu-Lughod, 1991).

Unlike the fixed-trait approach, the knowledge perspective to culture highlights the dynamical, contextualized, and agentic aspects of culture. First, as Barth (2002) puts it, "knowledge is distributed in a population, while culture [traditionally defined] makes us think in terms of diffuse sharing" (p. 1). Because knowledge is distributed, the patterns of distributions and their determi-

nants cannot be taken for granted. Instead, they merit systematic investigations.

Second, items of cultural knowledge, like items of personal knowledge, are cognitive structures. Accordingly, basic principles of knowledge activation can be applied to understand when a cultural knowledge item is evoked to grasp experiences in concrete situations, hence permitting researchers to predict when response differences between two cultural groups may appear, disappear, or reverse (Hong & Chiu, 2001).

Third, cultural knowledge is *conceptually* different from cultural identification (Chiu & Chen, 2004; Kashima, 2000). Individuals who have inherited the knowledge tradition in their culture may identify with this tradition, be indifferent about it, or even reject it. Accordingly, it is important to examine the joint psychological effects of cultural knowledge and cultural identification (Hong, Wan, No, & Chiu, in press).

Finally, knowledge provides people with premises for judgment and action (see Kruglanski & Thompson, 1999). Thus, culture may also be compared to a toolkit that can be put to manifold uses (DiMaggio, 1997), and people in a cultural group can sample knowledge items from their cultural toolkit to fulfill their valued goals (Chiu & Chen, 2004; Chiu & Hong, 2005).

In short, the knowledge perspective offers some advantages over a fixed-trait approach. To exploit these advantages, it is important to clarify the nature of cultural knowledge. In the next two sections, we discuss (1) how cultural knowledge differs from personal knowledge, and (2) whether the items in a cultural knowledge tradition are organized around a small number of central themes or dimensions.

Cultural Knowledge and Personal Knowledge

Like items of personal knowledge, items of cultural knowledge are cognitive structures. However, unlike personal knowledge, cultural knowledge is not idiosyncratic knowledge that resides primarily in the head of the individual. Aside from its rich symbolic contents, culture also has material contents (e.g., strategy of subsistence; see Harris, 1964) and social contents (e.g., kinship terms; see Keesing, 1974) that are exhibited in various public forms. Oftentimes, the same cultural theme can be discerned in a social group's symbolic, material, and social culture. Indeed, the research literature is richly furnished with illustrations of how cultural knowledge is externalized in a variety of media and social relations, including childrearing practices (Miller, Wiley, Fung, & Liang, 1997), popular songs (Rothbaum & Tsang, 1998; Rothbaum & Xu, 1995), architecture (Miyamoto, Nisbett, & Masuda, 2006), language (Kashima & Kashima, 1998, 2003), news media (Cohen & Nisbett, 1997; Hallahan, Lee, & Herzog, 1997; Lee, Hallahan, & Herzog, 1996; Menon et al., 1999; Morris & Peng, 1994), proverbs (Ho & Chiu, 1994), advertisement (Aaker & Schmitt, 2001; Han & Shavitt, 1994; Kim & Markus, 1999), consumption symbols (Aaker, Benet-Martinez, & Garolera, 2001), cultural icons (Hong, Benet-Martinez, Chiu, & Morris, 2003; Hong, Chiu, & Kung, 1997; Hong, Morris, Chiu, &

Benet-Martinez, 2000; Verkuyten & Pouliasi, 2002; Wong & Hong, 2005), role models (Lockwood, Marshall, & Sadler, 2005), life practices (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Morling, Kitayama, & Miyamoto, 2002), and law and social policies (Cohen, 1996). In short, cultural knowledge is distributed in a group of interconnected individuals, instantiated or communicated in one or several media, and distributed or communicated within a series of instituted social relations (Barth, 2002).

In fact, instantiation of cultural knowledge in external media is what makes rapid accumulation of cultural knowledge possible. Unlike animal culture, human culture is cumulative: When humans discover a piece of knowledge, they start from there, and go on. Once the wheel was invented, the new generations do not need to invent it again. Instead, they build on it and invent the carriage, and then the automobile (Tomasello, 2001). This process is known as *ratcheting*. Because human memory has limited capacity and can be assessed by one individual only, if social transmission of knowledge relies exclusively on individual memory, the rate of ratcheting would have been slow. In contrast, many external memory devices (e.g., paintings, books, and the Internet) have virtually unlimited capacity and many retrieval paths and can be accessed by multiple users. Thus, they are more efficient tools than personal memory for storing and transmitting cultural knowledge (Donald, 1993).

Instantiation of cultural knowledge in external media turns these media into effective means for the acquisition, transmission, activation, and maintenance of cultural knowledge. For example, individuals can directly access cultural knowledge encoded in the external media (e.g., scriptures, fictions, and folktales). Cultural values (e.g., human rights) are transmitted through folklores (e.g., stories of the Civil War). The presence of some external carriers of culture (e.g., the picture of the Statue of Liberty) in the environment can call out the cultural knowledge encoded in the media. When a cultural idea is instantiated in many different external carriers, individuals in the culture are frequently exposed to the idea. As a consequence of frequent activation, this idea becomes chronically accessible in the cultural group.

Coherence of Cultural Knowledge

We posit that culture consists of a set of *loosely organized* distributed knowledge. This view stands in sharp contrast to the *systemic* view of culture, which maintains that culture is a *coherent* system of meanings with an identifiable central theme around which all cultural meanings are organized (see Shore, 2002). For example, in psychology, the dimension of individualism–collectivism has received most research attention. Individualism emphasizes individual achievement, individuality, and individual rights. Collectivism emphasizes collective achievement, close ties between individuals, and responsibility for fellow members of the group (Hofstede, 1980, 2001). Some writers (Greenfield, 2000) consider individualism–

collectivism as the "deep structure" of cultural differences, from which all other differences evolved.

However, there is only limited evidence for the view that individualism–collectivism is the central theme that underlies the deep structure of cultural differences. In a comprehensive meta-analytic review of the extant literature on country differences in individualism and collectivism, Oyserman, Coon, and Kemmelmeier (2002) found that contrary to popular assumptions in cross-cultural and cultural psychology, "European Americans were not more individualistic than African Americans, or Latinos, and not less collectivistic than Japanese or Koreans" (p. 3). In addition, there are remarkable interdomain variations in country differences in individualism and collectivism. For example, in the case of U.S.–Japanese differences, North Americans are *more collectivist* than Japanese in most domains, including accepting hierarchy, striving to maintain group harmony, and defining oneself contextually, as well as sense of belonging to groups. Japanese are more collectivist than North Americans only in the domain of preference for working in a group. Comparisons of North Americans with other national groups point to the same conclusion: The nature of the difference between North Americans and other regional groups depends on which aspect of individualism or collectivism is being assessed.

In another review, after surveying the literature in six major areas of cross-cultural research (perception, cognition, language, personality, emotions, and social behaviors), Poortinga (2003) reached a similar conclusion: "The evidence for a system of values or meanings in one culture that differs in essential ways from the systems of other cultures appears to be rather limited." He urged researchers to stop explaining cultural differences in behavior in terms of broad and inclusive concepts, and instead consider culture as "a loose set of mentifacts with incidental interconnectedness" (p. 275).

Furthermore, inconsistent and contrastive cultural ideas are often present in the same external carrier of cultural meanings. For example, popular sayings and idioms carry widely shared evaluative, prescriptive, or proscriptive beliefs and are embedded in many conversation scripts. In a content analysis of over 2,000 Chinese popular sayings and idioms, Ho and Chiu (1994) found that ideas supporting individualism and ideas supporting collectivism are both prevalent in Chinese sayings. Likewise, Zhang and Shavitt (2003) analyzed the values promoted in Chinese advertising and found that ideas pertinent to modernity and individualism predominate in current Chinese advertising. They also discovered that individualism and modern values are more pervasive in magazine advertisements than in television commercials, whereas collectivism and traditional values are more pervasive on television than in magazine advertisements.

These findings do not undermine the theoretical importance of individualism and collectivism. Compared to many other cultural ideas, these two ideas seem to have a broad range of convenience and mediate a wide range of cultural differences (Triandis, 1995). However, in light of these findings, it seems justified to try to understand cul-

tural differences in terms of specific cultural knowledge (Shore, 2002).

FORMS OF CULTURAL KNOWLEDGE

If culture consists of a coalescence of knowledge items, can these items be classified into theoretically meaningful categories? Borrowing insights from the knowledge representation literature, we classify cultural knowledge into two major categories: Procedural knowledge and declarative knowledge. We further subdivide declarative knowledge into person representations, event representations, and norm representations. Aside from providing a taxonomic classification of cultural knowledge, this categorization also sheds light on how a specific cultural knowledge item is cognitively represented in an individual's head, and when and how it will be activated.

Procedural Knowledge

Two major types of knowledge representations are *procedural* knowledge (knowing *how*) and *declarative* knowledge (knowing *that*; Wyer, 2004). Procedural knowledge consists of procedural representations of how to achieve a particular result. It consists of a learned sequence of responses to situational cues. Once the learned response sequence is automated through frequent practices, its performance requires little cognitive deliberation.

Cultural experiences can lead to the development of procedural knowledge in at least two ways. First, the physical and social environment in a culture affords opportunities to practice a certain response repeatedly. Second, a certain culture may emphasize a particular goal. As members of the culture pursue this goal routinely, the procedure used to attain the goal is automated and can be activated in response to the controlling stimuli without the individual's conscious awareness.

Procedural knowledge may underlie some universal cultural processes. As an illustration, consider the decoding of emotions from facial expressions. In a typical emotion recognition experiment, perceivers view emotion faces on slides or photos and identify the emotion each face is intended to portray by selecting an emotion term from several alternatives. In most studies, perceivers are more accurate in judging the emotions of expressors from their own culture than emotions of foreign expressors (Elfenbein & Ambady, 2002).

The seemingly universal ingroup advantage in recognizing facial expressions of emotions may arise from more frequent practices in decoding facial expressions of one's cultural members (vs. foreign cultural members). For example, Chinese in China have greater exposure to emotion faces of Chinese people than do Chinese Americans and non-Asian Americans. Conversely, non-Asian Americans and Chinese Americans have greater exposure to emotion faces of Americans than do Chinese in China. Consistent with the "practice makes perfect" idea, Chinese in China are faster and more accurate in recognizing emotions from facial expressions of Chinese pos-

ers than from those of American posers, and the reverse is true for non-Asian Americans and Chinese Americans. In addition, among Chinese Americans, the advantage for recognizing American over Chinese expressions is greater among second-generation Chinese Americans than among immigrant generation Chinese Americans. Finally, Tibetans residing in China are faster and more accurate when they decode emotions from Chinese faces than from American faces. Likewise, Africans residing in the United States are more proficient in recognizing emotions from American faces than from Chinese faces (Elfenbein & Ambady, 2003). In short, proficiency (as measured by accuracy and speed) in recognizing the facial expressions of emotions of a certain group increases with the amount of opportunities to practice the skill.

Procedural knowledge may also underlie many cross-cultural differences, ranging from visual scanning (Braine, 1968; Hoosain, 1986) to language comprehension (Ishii, Reyes, & Kitayama, 2003; Kitayama & Ishii, 2002), attention (Kitayama, Duffy, Kawamura, & Larsen, 2003; Masuda & Nisbett, 2001), categorization (Ji et al., 2004), reasoning (Norenzayan, Smith, Kim, & Nisbett, 2002), and problem solving (Kim, 2002). For example, American undergraduates often use internal speech to keep track of their thoughts during problem solving. Requiring them to think aloud while solving reasoning problems does not affect their performance on the problem-solving task. However, thinking aloud is not an automated cognitive strategy for most East Asian undergraduates. To these East Asian undergraduates, thinking aloud requires cognitive effort and may lead to impaired performance in problem solving (Kim, 2002).

As another example, for most people, the preferred direction of visual scanning is left to right. One exception is found among Israelis. Israelis' preferred visual scanning direction is right to left, and Hebrew is written from right to left (Braine, 1968). Likewise, Chinese characters are sometimes read from left to right and sometimes read from top to bottom. The habitual way Chinese characters are read influences Chinese readers' preferred visual scanning direction. Unlike Americans, Chinese do not display lower visual acuity in vertical scanning than in horizontal scanning (Freeman, 1980). Furthermore, Chinese American children who do not have any experience with written Chinese have poorer acuity in vertical scanning than in horizontal scanning, just as the American adults do (Hoosain, 1986).

Procedural knowledge may also underlie cultural differences in attention and perceptual strategy. Compared to each other, European Americans attend to focal objects more spontaneously, and East Asians attend to contextual information more spontaneously. For example, East Asians (compared to European Americans) spend more time looking at the background in a visual display (Chua, Boland, & Nisbett, 2005), mention contextual information more frequently when they describe an animated scene, have better performance in detecting changes in background information (Masuda & Nisbett, 2004), and recognize previously seen objects more accurately when the objects are seen in their original settings

than in novel settings (Masuda & Nisbett, 2001). By comparison, European Americans spend more time looking at the focal object in a visual display (Chua et al., 2005), and have better performance in detecting changes in focal objects (Masuda & Nisbett, 2004). East Asians are also more accurate in processing contextual information (the length of the line relative to the height of a squared frame) than absolute information (the length of a line), and the reverse is true for European Americans (Ji et al., 2001; Kitayama, Duffy, Kawamura, & Larsen, 2003). Finally, when responding to an oral presentation of an emotional word, Japanese respond automatically to its vocal tone (contextual information), whereas European Americans respond automatically to its verbal contents (Ishii et al., 2003; Kitayama & Ishii, 2002).

A set of recent studies links some of these cultural differences to the perceptual environment in East Asia and North America. Based on a detailed textural analysis of city scenes in Japanese and American cities, Miyamoto and colleagues (2006) found that in U.S. cities, objects are usually distinctive and stand out from the background. Thus the experience of living in the American environment tends to direct one's attention to the distinctive and focal objects rather than to the background. Conversely, in Japanese cities, objects are more ambiguous and difficult to distinguish from the background. The experience of living in the Japanese environment may direct one's attention to the relationship between the focal object and the background.

To test the connection between environmental affordances and attention strategies, Miyamoto and colleagues (2006) had Japanese and European American undergraduates view either Japanese or American scenes. Subsequent to this manipulation, both Japanese and American participants in the Japanese scene condition displayed increased sensitivity to changes in the background, whereas those in the American scene condition displayed increased sensitivity to changes in the focal objects.

Aside from environmental affordances, pursuit of culturally important goals also affords opportunities to practice the procedures for attaining these goals. Compared to Westerners, Easterners are more worried about being isolated. They also tend to pay more attention to contextual information. In a series of studies, Kim and Markman (2006) linked the East–West difference in attention strategy, described earlier, to cross-cultural variation in the fear of isolation. According to these researchers, when individuals fear being isolated, they attend to the social context. Because fear of isolation is a chronic concern in East Asian contexts, East Asians have plenty of opportunities to practice the attention strategy connected to this social concern. As expected, in these studies, cultural differences in fear of isolation mediated East–West difference in attention strategies. Furthermore, among European Americans, those primed with fear of isolation exhibited greater sensitivity to contextual information than did control participants.

Taken together, these findings suggest that the shared physical, social, and motivational environment in a culture affords opportunities to develop procedural knowledge through repeated practices of a certain cognitive procedure.

Declarative Knowledge

Declarative knowledge is knowledge that is either true or false. It describes objects and events by referring to the characteristic properties of the events but does not specify the actions needed to obtain a result (Turban & Aronson, 1988). When a piece of declarative representation is activated, it may constrain subsequent inferences and decisions. The three types of declarative knowledge that have received the most attention in culture and psychology research are representations of persons, events, and norms.

Person Representations

A person representation is a network of associations between a central concept and a number of individual features. The referent of the central concept can be the self, a person other than the self, a group, or a social category. The features that are associatively linked to the referent may include alternative labels of the referent, traits, prototypical behaviors, or physical characteristics. The associations between individual features and the central concept differ in associative strength and may be specific to a particular type of situation.

Several different representations may be constructed for the same referent, each with a different set of associated features. For example, a person can construct a personal self that is associatively linked to a set of personal attributes, a social self that is linked to a set of social roles and role expectations, and a collective self that is linked to a set of collective memberships (Triandis, 1989). Moreover, each person representation is a separate cognitive unit that can be independently retrieved from memory (Wyer, 2004).

Cultural differences in person representations can be seen from the way individuals describe the referent of a person representation. When individuals retrieve a person representation from memory and report the features that are associatively linked to the referent, although several different representations of the same referent are available in memory, description may depend on which representation is most accessible in memory and, therefore, is most likely to be retrieved and used. Accordingly, although people in every culture may have constructed several different cognitive presentations of a person, cultural experiences determine the relative accessibility of these representations.

Consistent with this idea, research has shown that for European Americans, the most cognitively accessible mental representation of other people consists of general traits, whereas the one for Asians includes social roles and situation-specific behaviors. For example, compared to Asians, European Americans are more likely to use abstract traits to describe others, and less likely to use social role descriptions and situation-specific behaviors (Miller, 1984). They are more likely to attribute situationally induced events to the actor's dispositions

(Choi & Nisbett, 1998; Masuda & Kitayama, 2004; Miyamoto & Kitayama, 2002) and attribute the causes of social events to the actor's internal factors (Miller, 1984; Morris & Peng, 1994). In addition, they are less aware of the influence of the situation on behavior (Morris & Peng, 1994), and they make stronger predictions of trait-relevant behavior based on previous trait-relevant behavior (Norenzayan, Choi, & Nisbett, 2002). By comparison, East Asians are more affected by information about situational constraints when predicting trait-relevant behavior in a particular situation (Norenzayan, Choi, & Nisbett, 2002).

The same principle can also be applied to understand cultural variations in the representation of the self. According to Triandis (1989), all individuals can include in their self-representations three kinds of features: the private self (knowledge about one's own traits, states, or behaviors), the public self (knowledge about the generalized other's view of the self), and the collective self (knowledge about some collective's view of the self). Individuals in different cultural groups sample these three kinds of self with different probabilities. How likely the private self and the collective self will be sampled depends in part on the relative emphasis on personal versus collective goals in the cultural context. Some cultures (e.g., Western cultures) value independence of the self from others (Triandis, Bontempo, Villareal, Asai, & Lucca, 1988). In these cultures, the private self (self-reliance, independence, self-esteem, self-concept clarity, and self-realization) is likely to be sampled. Other cultures (e.g., East Asian cultures) value interdependence with ingroup members. In these cultures, the widely accepted view is that people should avoid pitting their personal goals against the collective goals. If a conflict between personal and group goals is inevitable, people should subordinate their personal goal to the group goal (Triandis et al., 1988). In such cultural contexts, higher rates of sampling the collective self are expected.

Consistent with this idea, some studies have shown that Chinese Americans are more likely than European Americans to describe themselves spontaneously in terms of attributes that reference group memberships rather than personal attributes (Bond & Cheung, 1983; Cousins, 1989; Trafimow, Triandis, & Goto, 1991). Other studies have shown that East Asians use more interdependent or group-related statements to describe themselves, whereas European North Americans use more independent self-statements (Hong, Ip, Chiu, Morris, & Menon, 2001; Rhee, Uleman, Lee, & Roman, 1995; Triandis, 1989; Wang, 2001, 2003). In a similar vein, Markus and Kitayama (1991, 2003) reported that European Americans are more inclined to view the self as a self-contained entity, whereas Japanese are more disposed to view the self as being embedded in social relations. These cultural differences may reflect differences in the chronic accessibility of the independent and interdependent self-representations.

Chronic representations of the self provide premises for subsequent judgments and behaviors. First, there is evidence that when the self is seen as being interdependent of others, individuals seek closeness with ingroup members. For example, the interdependent self-representation has higher chronic accessibility among Turks than European Canadians, and Turks have a stronger desire for closeness with family members and acquaintances than do European Canadians (Uskul, Hynie, & Lalonde, 2004). Second, the interdependent self is more widely shared and accessible in Eastern cultures, whereas the independent self is more widely shared and accessible in Western cultures (Markus & Kitayama, 1991). Under the influence of the interdependent self, compared to Westerners, Easterners have a greater tendency to see the group as agentic (Kashima et al., 2005); relational information (the target's interpersonal network, or community memberships) versus individuating information as useful in making social predictions (Gelfand, Spurlock, Sniezek, & Shao, 2000); commercial advertisements that appeal to ingroup benefits, harmony, and family integrity as persuasive; and those that appeal to personal preferences and benefits as unconvincing (Han & Shavitt, 1994). Furthermore, relative to North Americans, East Asians have better memory of other people's perspective on the self (Cohen & Gunz, 2002) and of the information encoded in relation to the collective self (Wagar & Cohen, 2003), are more likely to take other people's perspective when reading other people's emotions (Cohen & Gunz, 2002), pay closer attention to the common ground in communication (Haberstroh, Oyserman, Schwarz, Kuehnen, & Ji, 2002), feel more obliged to justify their choice for others (Fein, Hoshino-Browne, Davies, & Spencer, 2003), and have a greater tendency to mimic other people's behaviors (van Baaren, Maddux, Chartrand, de Bouter, & van Knippenberg, 2003).

By comparison, Westerners who see the self as being independent of others are motivated to make the personal self the nucleus of their social cognitive activities. They rely more on individuating (vs. relational) information in making social predictions (Gelfand et al., 2000). In addition, compared to Easterners, Westerners have higher self-esteem (Heine et al., 1999; Hetts, Sakuma, & Pelham, 1999), place heavier emphasis on self-consistency (Choi & Choi, 2002; Kashima, Kashima, et al., 2004; Suh, 2002), have a greater tendency to project one's own emotions to others (Cohen & Gunz, 2002), have better memories of information encoded in relation to the personal self (Wagar & Cohen, 2003), are more likely to make choices to highlight personal uniqueness (Aaker & Schmitt, 2001; Iyengar & Lepper, 1999; Kim & Drolet, 2003; Kim & Markus, 1999), feel more obliged to justify their personal choices (Heine & Lehman, 1997b), and are more likely to display egocentric biases in self-appraisals such as self-enhancement on individualist traits (viewing themselves as being more intelligent, unique, and independent than they really are; Gelfand et al., 2002; Heine & Lehman, 1997a; Hetts et al., 1999) and unrealistic optimism (perceiving the self as more invulnerable and more likely to experience positive events than it really is; Chang, Asakawa, & Sanna, 2001; Heine & Lehman 1995).

Finally, pursuit and attainment of interdependent goals and quality of interpersonal relationships are

better predictors of life satisfaction for East Asians than for European Americans, whereas self-esteem, identity consistency, personal freedom, pursuit and attainment of individual goals, and personal affect are better predictors of life satisfaction for Westerners than for Easterners (Diener & Diener, 1995; Kwan, Bond, & Singelis, 1997; Oishi & Diener 2001; Schimmack Radhakrishnan, Oishi, Dzokoto, & Ahadi, 2002; Suh, 2002; Suh et al., 1998).

Unlike procedural knowledge, which consists of automated productions resulting from repeated performance, person representations are propositional representations of people (including the self) abstracted from cultural experiences (see Koo & Choi, 2005). Such experiences provide materials for reflections on the self, other people, and the society and its norms. Furthermore, prevalence of cultural practices supporting a certain representation renders it highly accessible. In line with this principle, there is evidence that cultural practices that promote personal agency and self-esteem are more prevalent in the United States than in Japan (Kitayama et al., 1997; Morling et al., 2002), whereas cultural practices that promote adjustment to social constraints are more prevalent in Japan than in the United States (Morling et al., 2002). Furthermore, after individuals of a particular cultural group have engaged in a new culture for a prolonged period of time, their new experiences may render a previously inaccessible person representation accessible (Hetts et al., 1999).

Indeed, when an individual has constructed for the same referent several representations, each with different implications for a judgment or decision, the response that is made will depend on which representation is retrieved and used (see Wyer, 2004). Priming a particular self-representation can increase the temporary accessibility and hence the probability of retrieving and using this representation. As mentioned, the representation of the self as an independent entity is more accessible to Westerners than to Easterners. However, subsequent to being primed with an interdependent self, Westerners use group memberships rather than personal attributes to describe oneself (Trafimow et al., 1991; Trafimow, Silverman, Fan, & Law, 1997), adhere more strongly to collectivist values than individualist values (Gardner, Gabriel, & Lee, 1999), exhibit a greater tendency to mimic the interaction partner's behaviors (van Baaren et al., 2003) and a lesser tendency to assert one's uniqueness (Kim & Drolet, 2003), and attend more closely to the common ground in communication (Haberstroh et al., 2002).

In summary, person representations are propositional knowledge of people abstracted from cultural experiences. Although people in every culture may have constructed several different cognitive presentations of a person, the prevalence of cultural practices supporting certain person representations renders these representations more accessible than others. Chronic representations of the self provide premises for subsequent judgments and behaviors. Priming a particular self-representation can also increase the temporary accessi-

bility and hence the probability of retrieving and using this representation.

Event Representations

People spontaneously construct mental representations of temporally and thematically related sequences of events. An event representation allows individuals to picture in their mind a state of affairs or an event in a specific situation and how it transforms into another state or event. To capture the gist of the event sequence, a caption or header (e.g., "visiting a restaurant") can be attached to the event representation (see Wyer, 2004).

When an event representation becomes widely shared in human group, it becomes a cultural script. One widely studied cultural script in the United States is the culture of honor that persists in the American South (Cohen, 1996, 1998; Cohen & Nisbett, 1997). In the United States, Caucasian Americans in the South are thrice as likely as Caucasians in the North to commit conflict-, argument-, or brawl-related homicides (Cohen, 1998; Nisbett, 1993; Nisbett & Cohen, 1996). In addition, compared to Caucasians in the North, those in the South are more accepting of the use of violence for self-protection, to answer an affront, or to socialize children (Cohen, 1996). In the Old South where the law was weak, citizens had to depend on themselves to protect their own life, families, and wealth. Tolerance of insults became a symbol of vulnerability to predation, whereas violence in response to aggression and insults was an indication of one's determination to get even with the aggressor. Although the need for the use of violence as a means of self-protection has diminished, pockets of the culture of honor still persist in the South.

A cultural script consists of sequentially connected frames. When a frame is activated, the subsequent frames are activated as well. The script for the culture of honor contains the following frames:

Frame 1: A man receives negative remarks on himself and his family.
Frame 2: He interprets the remarks as an insult and an attack on his honor.
Frame 3: This thought infuriates him, and he wants to retaliate.
Frame 4: He uses or threatens the use of violence to coerce submission from the perpetrator as a means to restore his honor.
Frame 5: Regardless of the outcome, he feels proud for standing up against an attack on his honor and sees defending one's honor with violent acts as justified and honorable.

In societies in which the culture of honor prevails (e.g., the American South, Brazil, and Chile), violence in response to an attack on one's honor is often seen as justified and not blameworthy (Cohen & Nisbett, 1997; Vandello & Cohen, 2003). In addition, men who feel that their honor has been damaged would feel angry, display dominant behaviors, and be physically prepared to use

aggression (as revealed in an increase in cortisol and testostereone level; Cohen, Nisbett, Bowdle, & Schwarz, 1996).

Prototypical event representations are representations of events that routinely occur in a certain type of situation. A prototypical event representation can function as an "implicit theory" about the events that occur in a particular type of situation (Wyer, 2004). Cross-cultural differences in implicit theories have been reported. Two types of implicit theories have received the most attention in cultural psychology.

A widely researched implicit theory is the belief in the fixedness or malleability of people and the world. Cultures differ in the nature of these representations. Su and colleagues (1999; Norenzayan, Choi, & Nisbett, 2002) found that East Asians are substantially more likely to believe that the world is relatively fixed and individual persons relatively malleable (Chiu, Dweck, Tong, & Fu, 1997). If these implicit theories are used as guides for relating the self to the social environment and as premises for making causal inferences, East Asians should be more inclined to adjust to their social environment than to influence it, whereas North Americans should have the reverse preference. In addition, East Asians should be less inclined than North Americans to explain social behavior with the actor's global traits but relatively more inclined to reference external forces (dispositions of the group and situational constraints) in their social explanations. Both predictions have received clear empirical support (Chiu, Morris, Hong, & Menon, 2000; Menon et al., 1999; Miller, 1984; Miyamoto & Kitayama, 2002; Morling et al., 2002; Morris & Peng, 1994; Norenzayan, Choi, & Nisbett, 2002; Peng & Knowles, 2003).

People in different cultures also construct different implicit theories of change. For example, a widely held belief in Confucian societies (e.g., China and South Korea) is that two opposing forces, manifest in various forms in nature (the weak vs. the strong; evil vs. divine, illness vs. health, coldness vs. warmth, darkness vs. light, bad vs. good fortune), push themselves into the place of the other, resulting in changes. According to this view, similar to how the sun and moon continually emit their light, good fortune and ill are continually prevailing each against the other by an exact rule. This exact rule mandates a cyclical trajectory of changes. The continual prevailing of opposing forces gives rise to a stable reality, because all transient changes engendered by one force will be canceled out by the changes instigated by its opposing force. Thus, these beliefs about the trajectory of change may reinforce the belief in a fixed reality (Chan et al., 2001; Peng & Nisbett, 1999).

In contrast, since the beginning of the industrial revolution, optimism and the belief in incremental change and progress have dominated the social philosophy in Western Europe and the New World. The intellectuals witnessed how their world was being transformed into one of iron, coal, and steam; of machinery and engines; and of railroads, steamships, and telegraph wires. They were optimistic about the future; they believed that a better world lay just around the corner, and the making

of it was in people's own hand (Burchell, 1966). Theories of biological evolution and economic development that surfaced at that time forecast the extinction of unfit species, the decline of maladaptive social systems, the domination of the superior species, and the rise of more advanced social systems. These beliefs about the trajectory of change support a malleable view of the world and its institutions.

These culture-characteristic beliefs about the trajectory of change continue to have authority over Chinese and North American students. When things are moving in a particular direction, compared to their American peers, Chinese undergraduates are more likely to anticipate change in direction of movement. For example, they believe more strongly that a couple who have been dating each other for 2 years will break up, someone who has been a chess champion for 3 years will lose in the next game, a student from a poor family will become rich one day, two kindergarten children who have been fighting will become friends one day, and a trend in the growth rates of the world economy or the worldwide death rate for cancer will reverse in the future (Ji et al., 2001). Furthermore, individuals who expect the development of events to change course should be less surprised by unexpected events than those who expect events to develop following a linear trend. Consistent with the idea that East Asians subscribe to a cyclical theory of change and Americans a linear theory, Korean undergraduates exhibit a stronger hindsight bias than do their U.S. counterparts (Choi & Nisbett, 2000).

The belief that opposing forces operate together also increases the sensitivity to competing concerns in conflict situations and the motivation to reconcile them. In one study, Peng and Nisbett (1999; Cheung et al., 2003) had Chinese and American students analyze everyday life situations that involve intrapersonal conflicts (e.g., a conflict between having fun and going to school) or interpersonal conflicts (e.g., a conflict between mothers and their daughters). Consistent with the idea that the Confucian theory of change is more widely distributed in Asia than in the United States, Chinese responses tended to focus on the reconciliation of contradiction by considering merit and fault on both sides ("both the mothers and the daughters have failed to understand each other"). In contrast, American responses tended to come down in favor of one side or the other ("mothers should respect daughters' independence").

Norm Representations

A norm representation is an implicative proposition consisting of three elements: the antecedent circumstances, the norm, and the consequent conditions (Lindahl & Odelstad, 2000). The antecedent conditions specify the activation circumstances of the representation. They include the range of concrete situations where the norm is applicable. The norm refers to the state of affairs that is generally believed to be the case (e.g., the shared belief that the needy will receive help). In constructing a norm representation, individuals need to have access to the dis-

tribution of social knowledge in the society (e.g., they need to know the extent of agreement in the group with the idea of helping the needy—Ho & Chiu, 1998; Lau, Chiu, & Lee, 2001; Wan et al., in press; Wan, Chiu, Peng, & Tam, in press). Finally, the consequent conditions specify the behavioral implications of the norm. When the antecedent circumstances are present in a given situation, a certain state of affairs is designated as the norm, and the individual is expected to see to it that this state of affairs will take place. Because norm representations are implicative propositions, they have direct authority over behavior.

Cultural norms are widely distributed norm representations. For example, a highly uniform norm in the United States is one for experiencing emotions—Americans agree that they *should* feel happy (Eid & Diener, 2001). Americans who can live up to cultural expectations for experiencing positive emotions have higher levels of life satisfaction (Schimmack et al., 2002).

Different cultures have different norms. For example, in resolving conflicts, the prevailing norms in East Asian societies prescribe the use of mediational and accommodating strategies to minimize interpersonal animosity. In contrast, the prevailing norms in Western countries prescribe the use of more direct, confrontational strategies to win the negotiation game (Adair, Okumura, & Brett, 2001; Derlega, Cukur, Kuang, & Forsyth, 2002; Gelfand et al., 2001; Ohbuschi, Fukushima, & Tedeschi, 1999).

In summary, cultural differences can be understood in terms of the differing contents of the widely circulated and highly accessible knowledge structures in different cultures. Defining a culture tradition in terms of its component knowledge items (procedural knowledge, person representations, event representations, and norm representations) invites researchers to articulate clearly the type and nature of each knowledge item, as well as its range of applicability, activation circumstances, and inferential and behavioral implications. For example, environmental affordances support the development of procedural knowledge, which is activated automatically in the presence of the controlling situational cues. Activation of a particular self-construal (e.g., interdependent self) will call out its associated behaviors (e.g., behavioral mimicry). The applicability of a norm is augmented in situations in which cultural identities are salient (Jetten, Postmes, & McAuliffe, 2002), and when people are accountable to their cultural group for their behaviors (Briley, Morris, & Simonson, 2000; Gelfand & Realo, 1999).

THE INDIVIDUAL AND THE CULTURAL

Cultural knowledge, once activated, constrains subsequent behaviors. However, not all behaviors are expressions of cultural knowledge. Aside from cultural knowledge, every individual has some idiosyncratic personal experiences. In addition, regardless of their cultural background, all individuals have some basic psychogenic needs (such as the need for positive self-regard). Behaviors are the result of the dynamic interaction of basic psychogenic needs, cultural knowledge, and personal experiences.

Cultures also differ in how these three elements work together. First, some cultures emphasize personal experiences, whereas others emphasize conformance to cultural norms. For example, conformity is more emphasized in Asia than in North America. Indian adolescents value conformity more than do U.S. adolescents (Sundberg, Rohla, & Tyler, 1970). Hong Kong Chinese students are more likely to shift their opinions in the direction of the majority than are U.S. undergraduates (Meade & Barnard, 1973). In a quiz, compared to their U.S. peers, Taiwan Chinese are more likely to conform to their coparticipant's answer, particularly when the coparticipant seems to be knowledgeable about the topic (Huang & Harris, 1973). In the United States, people who have higher satisfaction have more intense positive affect and less intense negative affect. By comparison, in collectivist countries (e.g., China, Indonesia, and Colombia), both emotional experiences and norms regarding life satisfaction (how satisfied the ideal person would feel about his or her life) predict life satisfaction. In these countries, people who have high life satisfaction have pleasant emotional experiences, *and* expect the ideal person to have high life satisfaction (i.e., they view having a high level of life satisfaction to be a socially desirable characteristic cherished by the ideal person in their culture; Suh et al., 1998).

Second, a behavior mediated by personal attitudes in one culture may be mediated by cultural norms in another culture. Joan Miller and her colleagues (Baron & Miller, 2000; Miller & Bersoff, 1992, 1994, 1998; see also Janoff-Bulman & Leggett, 2002) reported that in North America, individual rights and freedom form the ultimate criteria for judging the moral correctness of social actions. Reciprocal helping is more often categorized as a personal choice rather than as a moral duty, and whether one feels the obligation to help another person is often affected by how likable the target is. According to Miller and Bersoff (1998), the American's tendency to consider liking for the target in assessing moral responsibilities to help may reflect a voluntaristic view of social relationship in U.S. culture. In contrast, in India, duties form the ultimate criterion for evaluating the moral rightness of social action. Helping is morally required, and reciprocal prosocial behavior is seen as a moral obligation instead of a social exchange.

Finally, cultures may also differ in the relationship between self-motives and cultural norms. Self-motives and cultural norms may act in concert to produce the same behavior in some cultures and push behaviors into opposite directions in others. In the United States, both the need for self-esteem and the dominant cultural norm lead individuals into self-enhancement (Bond, Leung, & Wan, 1978), and research has found consistent self-enhancement biases among European Americans on a wide variety of measures (Davis & Stephan, 1980; Miller & Ross, 1975). In contrast, in Asian cultures, unbridled expression of the positive self goes against the norm of mod-

esty (Bond, Leung, & Wan, 1978; Kim, Kim, Kam, & Shin, 2003), and overt expressions of the positive self are permissible only under certain circumstances. As a consequence, in these cultures, self-enhancement is relatively infrequent (Akimoto & Sanbonmatsu 1999; Gelfand et al., 2002; Heine & Renshaw, 2002; Oishi et al., 2000; Ross, Heine, Wilson, & Sugimori, 2005), particularly in public situations (Kanagawa, Cross, & Markus, 2001). The seeming lack of self-enhancement in public arenas among Japanese has led some researchers (Heine et al., 1999; Kitayama, Takagi, & Matsumoto, 1995) to conclude that the Japanese may not have the need for positive self-regard.

However, a closer examination of the empirical evidence reveals that like Americans, Asians also have the need for positive self-regard; they do not self-enhance as much as their American peers do because of the relatively heavy emphasis on conformance to the modesty norm in Asian contexts (Kurman, 2001, 2003; Kurman & Sriram, 1997). In situations in which the norm of modesty is not enforceable, as when all responses are completely anonymous, Japanese also self-enhance (Kudo & Numazaki, 2003). In addition, Japanese display self-enhancing responses when self-enhancement is measured via implicit or indirect measures (Kitayama & Karasawa, 1997; Kobayashi & Greenwald, 2003; Muramoto, 2003). Japanese also exhibit self-enhancement biases (feel confident in their performance, and review the performance feedback when they outperform their competitor) in a competitive situation, where the goal of winning is more important than the goal of being modest (Takata, 2003). Finally, although Asians feel compelled to display humility when presenting their agentic qualities, they feel comfortable in exaggerating their positive communal qualities (e.g., cooperation). Several studies have shown that while European American undergraduates self-enhance more on agentic traits than do their Asian peers, Asian undergraduates self-enhance *more* on communal traits (Kobayashi & Brown, 2003; Kurman, 2001; Lalwani, Shavitt, & Johnson, 2006; Sedikides, Gaertner, & Toguchi, 2003).

In short, to understand behaviors in cultural contexts, it is not enough to just describe the relative distributions of individual knowledge items. It is also important to consider the dynamic interactions of universal motives, cultural knowledge, and personal experiences. In some cultures, these elements may be unrelated. In other cultures, they may reinforce each other, or compete with each other for authority over behavior.

PRINCIPLES OF CULTURAL KNOWLEDGE APPLICATION

Taking the knowledge perspective to culture allows researchers to apply basic principles of knowledge application to predict how culture works in concrete situations. Application of knowledge activation principles to cultural knowledge has shed light on how cultural knowledge impacts behaviors in concrete situations.

Chronic Accessibility

A body of cultural knowledge gains chronic accessibility when it has been used frequently. Cultural knowledge items that have been frequently used in a group are usually widely shared (Lau et al., 2001; Sechrist & Stangor, 2001), more frequently reproduced in communication (Lyons & Kashima, 2001, 2003), widely represented in external or public carriers of culture (Menon & Morris, 2001), and cognitively accessible to members of the group (Hong et al., 2000). We have illustrated with many examples how chronic accessibility in procedural and declarative knowledge may mediate a wide range of cultural differences.

Temporary Accessibility

Contextual cues may increase the temporary accessibility of a body of knowledge and momentarily raise the probability that this body of knowledge will be applied. We have described how priming a specific representation of the self calls out its associated cognitive and behavioral responses. Recently, researchers have applied the principle of temporary accessibility to illuminate the process of cultural frame switching.

Flexible switching of cultural frames is an experience familiar to people with multicultural background. When individuals who have engaged in both Chinese and American cultures (e.g., Chinese Americans and Westernized Hong Kong Chinese undergraduates) are primed with either Chinese cultural icons (e.g., the Chinese dragon) or American cultural icons (Mickey Mouse), they assimilate their responses into the primed culture (Hong et al., 1997, 2000). When primed with Chinese (vs. American) cultural icons, these bicultural individuals make more group attributions and fewer individual attributions. Analogous culture priming effects have been found on spontaneous self-construal (Ross, Xun, & Wilson, 2002), perception of physical events (Peng & Knowles, 2003), encoding and memory of person information (Sui, Zhu, & Chiu, in press), spontaneous inference of cultural values (Fu, Chiu, Morris, & Young, in press), and cooperative behavior (Wong & Hong, 2005). In addition, the culture priming effect has been reported in studies that used different bicultural samples (Chinese–Canadian, Dutch–Greek bicultural children), and a variety of cultural primes (e.g., language and experimenter's cultural identity; Ross et al., 2002; Verkuyten & Pouliasi, 2002).

Applicability

The probability that a knowledge item will be applied also depends on its applicability in the immediate context. Knowledge applicability is defined in terms of the extent of mapping between "the features of a stored construct and the attended features of a stimulus" (Higgins & Brendl, 1995, p. 220). In a series of culture priming experiments, Hong and colleagues (2003) found that among Chinese American bicultural individuals, culture priming affects the likelihood of making group attribu-

tion or individual attribution only when the tension between group agency and individual agency in the stimulus event is highlighted, rendering the cultural theory of group versus individual agency applicable in the judgment task. Furthermore, previous research has shown that in Chinese societies, the norm of cooperation applies to friendship but not to interactions with strangers (Ho & Chiu, 1994). Consistent with the principle of applicability, Hong Kong Chinese primed with Chinese (vs. American) culture icons make more cooperative choices when they play a prisoner's dilemma game with friends, but not when they play it with strangers (Wong & Hong, 2005).

Motivation and Cognitive Load

Culture is like a collection of chronically accessible cognitive tools (DiMaggio, 1997). Consistent with this analogy, research has shown that people are likely to use these tools when the situation calls for a readily available or conventionalized solution, or when the problem solver lacks the capability, motivation, or resource to consider alternative solutions (Knowles, Morris, Chiu, & Hong, 2001).

Cultural differences also tend to accentuate when people have high chronic need for cognitive closure (the need for a certain answer). For the Chinese, high need for cognitive closure is positively related to the tendency to make group attributions, whereas for European Americans, high need for cognitive closure is positively related to the tendency to make individual attributions. Situational inducement of the need for cognitive closure by manipulating time pressure produces a similar effect. Putting the Chinese under time pressure and thus increasing their need for cognitive closure increases their tendency to make group attribution. Conversely, putting European Americans under time pressure increases their tendency to make individual attribution (Chiu et al., 2000; Fu et al., in press).

This finding has important implications for understanding the acculturation pattern of new immigrants. Individuals with a high need for cognitive closure rely on readily available knowledge to resolve interpretive ambiguity. However, when the readily available knowledge is no longer seen as applicable, they will eagerly seize new knowledge from their surrounding to reduce interpretive ambiguity. Consistent with this idea, Kosic, Kruglanski, Pierro, and Mannetti (2004) found that immigrants who are surrounded by other members of their ethnocultural group perceive the knowledge of their home culture as applicable. For them, a higher need for cognitive closure is associated with a stronger motivation to adhere to the culture of origin. Conversely, immigrants who are surrounded by members of the host country no longer perceive the knowledge of their home culture as applicable. For them, a higher need for cognitive closure is associated with a stronger motivation to assimilate into the host culture.

People rely on readily accessible knowledge when they lack cognitive resources. If cultural knowledge is chronically accessible knowledge, people will use it particularly when they are cognitively busy. Consistent with this hypothesis, European Americans have a greater tendency to make individual attribution than do Chinese, but only when they are under cognitive load. When individuals are not cognitively busy, this cultural difference is significantly attenuated (Knowles et al., 2001; Lieberman, Jarcho, & Obayashi, 2005).

Self and Contrast Effect

People who are primed with a social category tend to assimilate their response into the primed category. For example, university students walk more slowly than they normally do after they have been subliminally primed with the "elderly" stereotype (Bargh, Chen, & Burrows, 1996). However, activation of social stereotypes can also result in automatic behavioral contrast if a comparison of the self to the stereotyped group is provoked. For example, when university undergraduates are primed with the self and the elderly stereotype, a comparison intention is activated. They may feel that they do not belong to the elderly category and display a behavioral contrast (see Dijksterhuis et al., 1998; Dijksterhuis, Spears, & Lepinasse, 2001; Schubert & Hafner, 2003; Spears, Gordijn, Dijksterhuis, & Stapel, 2004).

As noted, priming a culture often leads to assimilative responses. However, if the participants feel that they do not belong to the primed culture, culture priming may lead to contrastive responses. Bond and his colleagues have reported contrast effects in a study that used languages as culture primes. In this study (Yang & Bond, 1980), Chinese–English bilingual participants responded to a Chinese value survey. They showed stronger adherence to Chinese values when the survey was in English than in Chinese. In this case, the presence of an outgroup language reminded the participants that they did not belong to the primed cultural group. As a consequence, a contrast effect was found. In another study (Bond & Cheung, 1984), Cantonese-speaking Hong Kong Chinese undergraduates filled out a survey of traditional Chinese beliefs. Mainland Chinese are generally seen as more traditional than Hong Kong Chinese. Participants who received oral instructions in Putonghua (the spoken language in Mainland China) responded more like Westerners, compared to those who received instructions in Cantonese (a dialect used in Hong Kong).

Some bicultural individuals view their dual cultural identities as oppositional (I cannot be both a Chinese and an American at the same time), while others see them as independent or complementary (e.g., I am both a Chinese and an American; Benet-Martinez, Leu, Lee, & Morris, 2002; Tsai, Ying, & Lee, 2000). Variations in how dual identities are managed are related to bicultural individuals' responses to cultural priming. Those who view their dual identities as independent or complementary tend to assimilate their responses to primed culture. Those who view their dual identities as oppositional may feel ambivalent about either cultural identity and thus respond reactively to the cultural primes, displaying contrastive responses (Benet-Martinez et al., 2002). In

summary, activation of cultural knowledge follows the basic principles of knowledge activation, which govern the dynamic interactions between cultural knowledge, the situation, the individuals' current cognitive and motivational states, and cultural identities.

WHAT IS CULTURE FOR?

Cultural knowledge is a collective resource accumulated over time. This resource is used to attain collective and individual goals. A major research objective in the social psychology of culture is to identify the societal and psychological functions of culture and to explain how culture evolves.

Societal Functions of Culture

It is widely accepted that culture is a collectively constructed device to solve coordination problems (Cohen, 2001; Fiske, 2000; Heylighen & Campbell, 1995; Kashima, 1999). To ensure that individuals facing coordination problems will make joint decisions that produce optimal outcomes to all parties involved, several major types of social control mechanisms have evolved in human societies. The most rudimentary type is mutual monitoring. Through informal face-to-face transmissions, the involved parties agree on some common beliefs and practices, and those with deviating beliefs or patterns of behaviors are held in check by different forms of ostracism. Individuals may also internalize and identify with these beliefs and practices. Indeed, according to Heylighen and Campbell (1995), in human societies, beliefs spread through conformist transmission and internalization are often the predominant mechanisms of social control. These beliefs are then crystallized into cultural knowledge.

Successful solution of complex coordination problems also requires coordination of plans and actions. In addition to being a social control mechanism, culture also provides the shared, standard operating procedures, unstated assumptions, tools, norms, and values for grasping experiences (Triandis, 1996). As such, by virtue of culture, individuals are capable of constructing shared representations of their experiences, and coordinating their plans and actions (Kashima, 1999).

Furthermore, culture provides conventionalized solutions to recurrent coordination problems (Kashima, 1999). Culture is to society as what memory to individuals (Kluckhohn, 1954). Culture encodes collective memories of schematized approaches to solving coordination problems (Kashima, 1999). These conventionalized solutions are widely accepted solutions in the community. Competent members of the culture can retrieve such conventionalized solutions to solve emergent coordination problems.

In short, coordination problems arise in all human communities in which people need to make joint decisions on problems with more than one possible optimal solution. Culture provides different means for solving such coordination problems.

Psychological Functions of Culture

Culture also serves several psychological functions. First, it serves to reduce uncertainty in social living. Because cultural knowledge is widely distributed in a human group, it serves an epistemic function by offering individuals a socially validated perspective for grasping experiences (Kosic et al., 2004). Knowing that others share one's own knowledge increases confidence in the knowledge. For example, individuals who learn that others share their attitudes increase their confidence in the attitudes and the likelihood of using these attitudes to guide their behavior (e.g., Sechrist & Stangor, 2001). Accordingly, cultural knowledge, as a kind of socially validated knowledge, has greater authority over behaviors, compared to personal knowledge.

Because cultural knowledge affords a sense of epistemic security, or the feeling of having certain answers to questions in life, when people need certain answers, they may tend to rely on cultural knowledge. For this reason, cultural knowledge is particularly likely to be used when people are under cognitive load or time pressure (Chiu et al., 2000; Knowles et al., 2001).

Second, cultural knowledge may provide a buffer against the anxiety arising from the vulnerability of the individual. According to the terror management theory (Greenberg, Solomon, & Pyszczynzki, 1997), the thought of death is anxiety provoking. Because cultural knowledge is reproduced across generations, it outlasts the individual. By identifying with their cultural tradition, individuals may feel that they are a part of the imperishable cultural tradition and hence feel not as terrified by the thought of death. The idea that culture may offer a psychological buffer against existential terror has received some support. Making salient the threat of mortality strengthens adherence to individualist values among Australians and weakens it among the Japanese (Kashima, Halloran, Yuki, & Kashima, 2004). In addition, among Aborginal Australian bicultural individuals, mortality salience increases adherence to Aborginal values when the Aborginal identity is primed, and adherence to Australian values when the Australian identity is primed (Halloran & Kashima, 2004).

Finally, individuals need culture when they need to feel that they belong to a group. By identifying with the group and its culture, individuals fulfill their need for belongingness. The knowledge tradition of a group is a defining feature of the group (D'Andrade, 1987). When the need to belong to a group is activated, people who identify strongly with their ingroup will be motivated to adhere to the group's knowledge tradition (Heaven, 1999; Heaven, Stones, Simbayi, & Le Roux, 2000; Hong, Chiu, Yeung, & Tong, 1999; Lam, Lau, Chiu, Hong, & Peng, 1999) and will defend their cultural identity by affirming their cultural knowledge tradition (Jetten et al., 2002).

The need to belong may change across situations. The need to belong to the ingroup will be stronger either when people realize that they are different from other members of the ingroup or when they realize that they are not different from members of an outgroup (Brewer,

1991; Pickett & Brewer, 2001; Picket, Bonner, & Coleman, 2002). In addition, when people see visual icons of their group, they tend to see themselves as members of a group rather than as individuals (Briley & Wyer, 2002). When the need to belong is induced, people tend to adhere to the group's cultural tradition. In summary, people need culture to satisfy basic social and psychological needs, including the needs to reduce uncertainty in social living, to manage existential terror, and to fulfill the need for belongingness.

HOW DOES CULTURE SPREAD?

Culture is a shared phenomenon. Through communication, people make the contents of their mind accessible to others. Not surprisingly, dynamical systems theorists have attempted to explain how culture spreads through interpersonal communication. For example, Kashima, Woolcock, and Kashima (2000) have developed a connectionist model to account for the reproduction and stabilization of cultural meanings in a group. The model assumes that cultural ideas are not evenly distributed across members of a group. Individuals in a group resemble simple processing units in a network. These units receive information from each other and reproduce the information through the network. The network of connections is updated as serial reproductions proceed. Due to memory decay and schema-driven distortions, errors are introduced in the reproduced messages. In addition, to overcome the cognitive limits of the individual processing units, cognitive division of labor and externalization of memory are implemented. Through self-organization processes, cultural knowledge that is widely shared among group members (e.g., stereotypical knowledge) will have a high likelihood of being reproduced and consolidated in communication.

This kind of connectionist models has several conceptual advantages. First, it explains how cultural knowledge and meanings are reproduced and maintained in everyday interpersonal communication. Second, it relies on a set of self-organizational processes to produce and maintain shared meanings spontaneously. Thus, it does not require a homunculus-like agent or a collective will to oversee or manage the reproduction of cultural meanings. Third, the model is specific enough to permit precise simulation of the postulated cultural meaning reproduction processes in controlled experiments. Findings from such experiments have lent support to the proposed model (Lyons & Kashima, 2001, 2003).

Communication theories have also been recruited to explain spatial distribution of shared attitudes and beliefs. According to dynamic social impact theory (Latane, 1996), people influence and are influenced by the proximal people with whom they communicate. Through interpersonal communication, a set of values and beliefs become spatially differentiated (or clustered), and previously unconnected values and beliefs become correlated (see also Brauer, Judd, & Jacquelin, 2001). These self-organization processes have been demonstrated in computer simulations (Latane & Bourgeois, 1996; Latane & L'Herrou, 1996), in computer-mediated communication games, and in a longitudinal study of college students' political socialization in their academic discipline (Guimond & Palmer, 1996).

HOW DOES CULTURE CHANGE?: A FUTURE RESEARCH AGENDA

Culture is constantly being reproduced, but culture also changes. In a recent review of the extant literature in culture and psychology, Lehman and colleagues (2004) noted a lopsided emphasis on cultural differences in the psychological research and concluded that consideration of how societies adapt to globalization and culture change will become more and more important in the coming years.

Culture change may occur because of endogenous and exogenous reasons. Endogenous culture change may take place through the processes of differentiation. According to Triandis (2004), a cultural idea that works well may be replaced by an idea that works slightly better, but frequently both tools are retained in the culture. Accordingly, random variation and selective retention result in different dimensions of cultural variation.

Exogenous culture change often results from intercultural contacts, through which people from a culture are incidentally exposed to knowledge from another culture. With rapid development of electronic communication, people find themselves increasingly enmeshed in global symbolic environments (Bandura, 2002). As cultural boundaries become increasingly permeable and fuzzy, intercultural contacts could take place without face-to-face contacts with members of other cultures.

Culture change often involves complex transactions between old and new cultural knowledge in the society. In many developing countries, the arrival of the global culture has also brought forward cultural identity issues that are intertwined with complicated socioeconomic and political issues (Hong et al., in press). Some of these issues have caught psychologists' attention. For example, Arnett (2002) noted that with globalization, people worldwide now face the issue of relating to their local culture and the global culture. How would individuals react and relate to the confluence of the local and global cultures? What would the resulting bicultural identity be like (Orr, Mana, & Mana, 2003; Tafarodi, Kang, & Milne, 2002)? How can people navigate multiple cultures competently (Chiu & Hong, 2005)? How do people manage knowledge from different cultural traditions (Arends-Toth & Van De Vijver, 2003)? What are the implications of individuals' responses to multiculturalism for psychological well-being (Downie, Koestner, ElGeledi, & Cree, 2004)? The field of culture change is one with many questions and few answers. Yet, answers to these questions are of foremost importance as they will connect cultural psychology to the psychology of social identification, political psychology, and other social science disciplines.

CONCLUDING REMARKS

> Every science passes through a phase in which it considered
> its basic subject matter to be some sort of substance or
> structure. Every science has passed beyond that phase,
> recognizing its subject matter as being some sort of process.
> —Bickhard (2003, p. 122)

Similar historical shift is beginning to take place in the social psychology of culture. The current focus in the social psychology of culture is still on some sort of substance or structure: How do cultures differ? What are the deep structures of cultures? However, there is increasing recognition that to explain culture, we need to understand cultural processes. Bickhard (2003) submits that causality resides in process, not substances. A paradigmatic shift in physics occurred when physicists recognized that there are no particles. What appear as particle interactions are actually quantized oscillatory processes. Unlike particles, which are supposed to take on some irreducible forms, quantum fields are organized processes. In quantum physics, all causal power is resident in process organizations, and new organizations of process give rise to higher-level phenomena (Bickhard, 2003). Extending this view to the social psychology of culture, explaining culture involves identification of basic principles that give rise to organized cultural processes.

In this chapter, we adopted a knowledge perspective to culture, which places a strong emphasis on organized cultural processes. From the knowledge perspective, culture consists of a collection of loosely organized shared knowledge constantly being created, maintained, reproduced, and transformed by a collection of interconnected individuals. Cultural differences in behavior are understood not with reference to differing deep structures in different cultures but in terms of specific knowledge structures. Each knowledge structure has a specified range of applicability. Thus, instead of focusing on global cultural differences, our analysis focused on the social psychological consequences when specific knowledge structures are activated in concrete situations. In addition, cultural meanings are created, maintained, reproduced, and transformed for adaptive purposes. Thus, explaining cultural processes also demands attention to culture's social and psychological functions.

REFERENCES

Aaker, J., Benet-Martinez, V., & Garolera, J. (2001). Consumption symbols as carriers of culture: A study of Japanese and Spanish brand personality. *Journal of Personality and Social Psychology*, *81*, 492–508.

Aaker, J., & Schmitt, B. (2001). Culture-dependent assimilation and differentiation of the self: Preferences for consumption symbols in the United States and China. *Journal of Cross-Cultural Psychology*, *32*, 561–576.

Abu-Lughod, L. (1991). Writing against culture. In R. G. Fox (Ed.), *Recapturing anthropology: Working in the present* (pp. 137–162). Santa Fe, NM: School of American Research Press.

Adair, W. L., Okumura, T., & Brett, J. M. (2001). Negotiation behavior when cultures collide: The United States and Japan. *Journal of Applied Psychology*, *86*, 371–385.

Akimoto, S. A., & Sanbonmatsu, D. M. (1999). Differences in self-effacing behavior between European and Japanese Americans: Effect on competence evaluation. *Journal of Cross-Cultural Psychology*, *30*, 159–177.

Appadurai, A. (1996). *Modernity at large: Cultural dimensions of globalization*. Minneapolis: University of Minnesota Press.

Arends-Toth, J., & Van de Vijver, F. J. R. (2003). Multiculturalism and acculturation: Views of Dutch and Turkish–Dutch. *European Journal of Social Psychology*, *33*, 249–266.

Arnett, J. J. (2002). The psychology of globalization. *American Psychologist*, *57*, 774–783.

Bandura, A. (2002). Social cognitive theory in cultural context. *Applied Psychology: An International Review*, *51*, 269–290.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230–244.

Baron, J., & Miller, J. G. (2000). Limiting the scope of moral obligations to help: A cross-cultural investigation. *Journal of Cross-Cultural Psychology*, *31*, 703–725.

Barth, F. (2002). An anthropology of knowledge. *Current Anthropology*, *43*, 1–18.

Benet-Martinez, V., Leu, J., Lee, F., & Morris, M. W. (2002). Negotiating biculturalism: Cultural frame switching in biculturals with oppositional versus compatible cultural identities. *Journal of Cross-Cultural Psychology*, *33*, 492–516.

Bickhard, M. H. (2003). Part II: Applications of process-based theories: Process and emergence: Normative function and representation. *Axiomathes*, *14*, 121–155.

Bickhard, M. H., & Campbell, D. T. (2000). Emergence. In P. B. Andersen, C. Emmeche, N. O. Finnemann, & P. V. Christiansen (Eds.), *Downward causation* (pp. 322–348). Aarhus, Denmark: University of Aarhus Press.

Bond, M. H., & Cheung, M-k. (1984). Experimenter language choice and ethnic affirmation by Chinese trilinguals in Hong Kong. *International Journal of Intercultural Relations*, *8*, 347–356.

Bond, M. H., & Cheung, T-S. (1983). College students' spontaneous self-concept: The effect of culture among respondents in Hong Kong, Japan, and the United States. *Journal of Cross-Cultural Psychology*, *14*, 153–171.

Bond, M. H., Leung, K., & Wan, K-c. (1978). The social impact of self-effacing attributions: The Chinese case. *Journal of Social Psychology*, *118*, 157–166.

Braine, L. G. (1968). Asymmetries of pattern perception observed in Israelis. *Neuropsychologia*, *6*, 73–88.

Brauer, M., Judd, C. M., & Jacquelin, V. (2001). The communication of social stereotypes: The effects of group discussion and information distribution on stereotypic appraisals. *Journal of Personality and Social Psychology*, *81*, 463–475.

Braumann, C. (1999). Writing for culture: Why a successful concept should not be discarded. *Current Anthropology*, *40*, S1–S27.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin*, *17*, 475–482.

Briley, D. A., Morris, M. W., & Simonson, I. (2000). Reasons as carriers of culture: Dynamic versus dispositional models of cultural influence on decision-making. *Journal of Consumer Research*, *27*, 157–178

Briley, D. A., & Wyer, R. S., Jr. (2002). The effect of group membership salience on the avoidance of negative outcomes: Implications for social and consumer decisions. *Journal of Consumer Research*, *29*, 400–415.

Burchell, S. C. (1966). *Age of progress*. New York: Time-Life Books.

Chan, H. M., Chan, K. M., Cheung, T. S., King, A. Y. C., Chiu, C-y., & Yang, C. F. (2001). How Confucian are Chinese today?: Construction of an ideal type and its application to three Chinese communities. *Sociological Study*, *92*, 33–48.

Chang, E. C., Asakawa, K., & Sanna, L. J. (2001). Cultural variations on optimistic and pessimistic bias: Do Easterners really expect

the worst and Westerners really expect the best when predicting future life events? *Journal of Personality and Social Psychology*, *81*, 476–491.

Cheung, T-S., Chan, H-M., Chan, K-M., King, A. Y-K., Chiu, C., & Yang, C-F. (2003). On Zhongyong rationality: The Confucian doctrine of the mean as a missing link between instrumental rationality and communicative rationality. *Asian Journal of Social Science*, *31*, 107–127.

Chiu, C-y., & Chen, J. (2004). Symbols and interactions: Application of the CCC model to culture, language, and social identity. In S-h. Ng, C. Candlin, & C-y. Chiu (Eds.), *Language matters: Communication, culture, and social identity*. Hong Kong: City University of Hong Kong Press.

Chiu. C-y., Dweck, C. S., Tong, Y-y., & Fu, H-y. (1997). Implicit theories and conceptions of morality. *Journal of Personality and Social Psychology*, *73*, 923–940.

Chiu, C-y., & Hong, Y. (2005). Cultural competence: Dynamic processes. In A. Elliot & C. S. Dweck (Eds.), *Handbook of motivation and competence* (pp. 489–505). New York: Guilford Press.

Chiu, C-y., Morris, M. W., Hong, Y-y., & Menon, T. (2000). Motivated cultural cognition: The impact of implicit cultural theories on dispositional attribution varies as a function of need for closure. *Journal of Personality and Social Psychology*, *78*, 247–259.

Choi, I., & Choi, Y. (2002). Culture and self-concept flexibility. *Personality and Social Psychology Bulletin*, *28*, 1508–1517.

Choi, I., & Nisbett, R. E. (1998). Situational salience and cultural differences in the correspondence bias and actor-observer bias. *Personality and Social Psychology Bulletin*, *24*, 949–960.

Choi, I., & Nisbett, R. E. (2000). Cultural psychology of surprise: Holistic theories and recognition of contradictions. *Journal of Personality and Social Psychology*, *79*, 890–905.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences USA*, *102*, 12629–12633.

Clifford, J. (1988). *The predicament of culture: Twentieth-century ethnography, literature, and art.* Cambridge, MA: Harvard University Press.

Cohen, D. (1996). Law, social policy, and violence: The impact of regional cultures. *Journal of Personality and Social Psychology*, *70*, 961–978.

Cohen, D. (1998). Culture, self-organization, and patterns of violence. *Journal of Personality and Social Psychology*, *75*, 408–419.

Cohen, D. (2001). Cultural variation: Considerations and implications. *Psychological Bulletin*, *127*, 451–471.

Cohen, D., & Gunz, A. (2002). As seen by the other . . . Perspectives on the self in the memories and emotional perceptions of Easterners and Westerners. *Psychological Science*, *13*, 55–59.

Cohen, D., & Nisbett, R. E. (1997). Field experiments examining the culture of honor: The role of institutions in perpetuating norms about violence. *Personality and Social Psychology Bulletin*, *23*, 1188–1199.

Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the Southern culture of honor: An "experimental ethnography." *Journal of Personality and Social Psychology*, *70*, 945–960.

Cousins, S. D. (1989). Culture and self-perception in Japan and the United States. *Journal of Personality and Social Psychology*, *56*, 124–131.

D'Andrade, R. (1987). A folk model of the mind. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 112–148). Cambridge, UK: Cambridge University Press.

Davis, M. H., & Stephan, W. G. (1980). Attributions for exam performance. *Journal of Applied Psychology*, *10*, 235–248.

Derlega, V. J., Cukur, C. S., Kuang, J. C. Y., & Forsyth, D. R. (2002). Interdependent construal of self and the endorsement of conflict resolution strategies in interpersonal, intergroup, and international disputes. *Journal of Cross-Cultural Psychology*, *33*, 610–625.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology*, *68*, 653–663.

Dijksterhuis, A., Spears, R., & Lepinasse, V. (2001). Reflecting and deflecting stereotypes: Assimilation and contrast in impression formation and automatic behavior. *Journal of Experimental Social Psychology*, *37*, 286–299.

Dijksterhuis, A., Spears, R., Postmes, T., Stapel, D., Koomen, W., van Knippenberg, A., et al. (1998). Seeing one thing and doing another: Contrast effects in automatic behavior. *Journal of Personality and Social Psychology*, *75*, 862–871.

DiMaggio, D. (1997). Culture and cognition. *Annual Review of Sociology*, *23*, 263–287.

Donald, M. (1993). Precis of "Origins of the modern mind: Three stages in the evolution of culture and cognition." *Behavioral and Brain Sciences*, *16*, 737–791.

Downie, M., Koestner, R., ElGeledi, S., & Cree, K. (2004). The impact of cultural internalization on well-being among tricultural individuals. *Personality and Social Psychology Bulletin*, *30*, 305–314.

Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intranational differences. *Journal of Personality and Social Psychology*, *81*, 869–885.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*, 203–235.

Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, *85*, 276–290.

Fein, S., Hoshino-Browne, E., Davies, P. G., & Spencer, S. J. (2003). Self-image maintenance goals and sociocultural norms in motivated social perception. In S. J. Spencer, S. Fein, M. P. Zanna, & J. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 21–44). Mahwah, NJ: Erlbaum.

Fiske, A. P. (2000). Complementarity theory: Why human social capacities evolved to require cultural complements? *Personality and Social Psychology Review*, *4*, 76–94.

Fiske, A. P., Kitayama, S., Markus, H. R., & Nisbett, R. E. (1995). The cultural matrix of social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 868–981). New York: McGraw-Hill.

Freeman, R. D. (1980). Visual acuity is better for letters in rows than in columns. *Nature*, *286*, 62–64.

Friedman, J. (1994). *Cultural identity and global process*. London: Sage.

Fu, H.-Y., Chiu, C-y., Morris, M. W., & Young, M. (in press). Spontaneous inferences from cultural cues: Varying responses of cultural insiders, outsiders, and sojourners. *Journal of Cross-Cultural Psychology*.

Fu, H.-Y., Morris, M. W., Lee, S.-L., Chao, M.-C., Chiu, C-y., & Hong, Y-y. (in press). Epistemic motives and cultural conformity: Need for closure, culture, and context as determinants of conflict judgments. *Journal of Personality and Social Psychology*.

Gardner, W. L., Gabriel, S., & Lee, A. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, *10*, 321–326.

Gelfand, M. J., Higgins, M., Nishii, L. H., Rava, J. L., Dominquez, A., Murakami, F., et al. (2002). Culture and egocentric perceptions of fairness in conflict and negotiation. *Journal of Applied Psychology*, *87*, 833–845.

Gelfand, M. J., Nishii, L. H., Holcombe, K. M., Dyer, M., Ohbuchi, K-l., & Fukuno, M. (2001). Cultural influences on cognitive representations of conflict: Interpretations of conflict episodes in the United States and Japan. *Journal of Applied Psychology*, *86*, 1059–1074.

Gelfand, M. J., & Realo, A. (1999). Individualism-collectivism and accountability in intergroup negotiations. *Journal of Applied Psychology*, *84*, 721–736.

Gelfand, M. J., Spurlock, D., Sniezek, J. A., & Shao, L. (2000). Culture and social prediction: The role of information in enhancing confidence in social predictions in the United States and China. *Journal of Cross-Cultural Psychology*, *31*, 498–516.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empiri-

cal assessments and conceptual refinements. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 61–139). New York: Academic Press.

Greenfield, P. M. (2000). Three approaches to the psychology of culture: Where do they come from? Where can they go? *Asian Journal of Social Psychology*, *3*, 223–240.

Guimond, S., & Palmer, D. L. (1986). The political socialization of commerce and social science students: Epistemic authority and attitude change. *Journal of Applied Social Psychology*, *26*, 1985–2013.

Haberstroh, S., Oyserman, D., Schwarz, N., Kuehnen, U., & Ji, L. (2002). Is the interdependent self more sensitive to question context than the independent self? Self-construal and the observation of conversational norms. *Journal of Experimental Social Psychology*, *38*, 323–329.

Hallahan, M., Lee, F., & Herzog, T. (1997). It's not just whether you win or lose, it's also where you play the game: A naturalistic, cross-cultural examination of the positivity bias. *Journal of Cross-Cultural Psychology*, *28*, 768–778.

Halloran, M. J., & Kashima, E. S. (2004). Social identity and worldview validation: The effects of ingroup identity primes and mortality salience on value endorsement. *Personality and Social Psychology Bulletin*, *30*, 915–925.

Han, S-p., & Shavitt, S. (1994). Persuasion and culture: Advertising appeals in individualistic and collectivistic societies. *Journal of Experimental Social Psychology*, *30*, 326–350.

Harris, M. (1964). *The nature of cultural things*. New York: Random House.

Heaven, P. C. L. (1999). Group identities and human values. *Journal of Social Psychology*, *139*, 590–595.

Heaven, P. C. L., Stones, C., Simbayi, L., & Le Roux, A. (2000). Human values and social identities among samples of white and black South Africans. *International Journal of Psychology*, *35*, 67–72.

Heine, S. J., & Lehman, D. R. (1995). Cultural variation in unrealistic optimism: Does the West feel more invulnerable than the East? *Journal of Personality and Social Psychology*, *68*, 595–607.

Heine, S. J., & Lehman, D. R. (1997a). The cultural construction of self-enhancement: An examination of group-serving bias. *Journal of Personality and Social Psychology*, *72*, 1268–1283.

Heine, S. J., & Lehman, D. R. (1997b). Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin*, *23*, 389–400.

Heine, S. J., Lehman, D. R., Markus, H, R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*, 766–794.

Heine, S. J., & Renshaw, K. (2002). Interjudge agreement, self-enhancement, and liking: Cross-cultural divergences. *Personality and Social Psychological Bulletin*, *28*, 442–451.

Hetts, J. J., Sakuma, M., & Pelham, B. W. (1999). Two roads to positive self-regard: Implicit and explicit self-evaluation and culture. *Journal of Experimental Social Psychology*, *35*, 512–559.

Heylighen, F., & Campbell, D. T. (1995). Selection of organization at the social level: Obstacles and facilitators of metasystem transitions. *World Futures: The Journal of General Evolution*, *45*, 181–212.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. E. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T., & Brendl, C. M. (1995). Accessibility and applicability: Some "activation rules" influencing judgment. *Journal of Experimental Social Psychology*, *31*, 218–243.

Ho, D. Y. F., & Chiu, C-y. (1994). Component ideas of individualism, collectivism, and social organization: An application in the study of Chinese culture. In U. Kim, H. C. Triandis, C. Kagitcibasi, G. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method and applications* (pp. 137–156). Thousand Oaks, CA: Sage.

Ho, D. Y. F., & Chiu, C-y. (1998). Collective representations

as a metaconstruct: An analysis based on methodological relationalism. *Culture and Psychology*, *4*, 349–369.

Hofstede, G. (1980). *Culture's consequences*. Beverly Hill, CA: Sage.

Hofstede, G. (2001). *Culture's consequences* (2nd ed.). Thousand Oaks, CA: Sage.

Hong, Y-y, Benet-Martinez, V., Chiu, C-y., & Morris, M. W. (2003). Boundaries of cultural influence: Construct activation as a mechanism for cultural differences in social perception. *Journal of Cross-Cultural Psychology*, *34*, 453–464.

Hong, Y-y, & Chiu, C-y. (2001). Toward a paradigm shift: From cultural differences in social cognition to social cognitive mediation of cultural differences. *Social Cognition*, *19*, 118–196.

Hong, Y-y., Chiu, C-y., & Kung, T. M. (1997). Bringing culture out in front: Effects of cultural meaning system activation on social cognition. In K. Leung, Y. Kashima, U. Kim, & S. Yamaguchi (Eds.), *Progress in Asian social psychology* (Vol. 1, pp. 135–146). Singapore: Wiley.

Hong, Y-y., Chiu, C-y., Yeung, G., & Tong, Y. (1999). Effects of self-categorization on intergroup perceptions: The case of Hongkong facing 1997. *International Journal of Intercultural Relations*, *23*, 257–279.

Hong, Y-y., Ip, G., Chiu, C., Morris, M. W., & Menon, T. (2001). Cultural identity and the dynamic construction of the self: Collective duties and individual rights in Chinese and American cultures. *Social Cognition*, *19*, 251–268.

Hong, Y-y., Morris, M. W., Chiu, C-y., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*, 709–720.

Hong, Y-y., Wan, C., No. S., & Chiu, C-y. (in press). Multicultural identities. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology*. New York: Guilford Press.

Hoosain, R. (1986). Language, orthography and cognitive processes: Chinese perspectives for the Sapir–Whorf hypothesis. *International Journal of Behavioral Development*, *9*, 507–525.

Huang, L., & Harris, M. B. (1973). Conformity in Chinese and Americans: A field experiment. *Journal of Cross-Cultural Psychology*, *4*, 427–434.

Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological Science*, *14*, 39–46.

Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology*, *76*, 349–366.

Janoff-Bulman, R., & Leggett, H. K. (2002). Culture and social obligation: When "shoulds" are perceived as "wants." *Journal of Research in Personality*, *36*, 260–270.

Jetten, J., Postmes, T., & McAuliffe, B. (2002). "We're all individuals': Group norms of individualism and collectivism, levels of identification and identity threat. *European Journal of Social Psychology*, *32*, 189–207.

Ji, L-j., Nisbett, R. E., & Su, Y. (2001). Culture, change, and prediction. *Psychological Science*, *12*, 450–456.

Ji, L-j., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, *87*, 57–65.

Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). "Who am I?": The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, *27*, 90–103.

Kashima, E. S., Halloran, M. J., Yuki, M., & Kashima, Y. (2004). The effects of personal and collective mortality salience on individualism: Comparing Australians and Japanese with higher and lower self-esteem. *Journal of Experimental Social Psychology*, *40*, 384–392.

Kashima, E. S., & Kashima, Y. (1998). Culture and language: The case of cultural dimensions and personal pronoun use. *Journal of Cross-Cultural Psychology*, *29*, 461–486.

Kashima, Y. (1999). Culture, groups, and coordination problems. *Psychologische Beitrage*, *41*, 237–251.

Kashima, Y. (2000). Conceptions of culture and person for psychology. *Journal of Cross-Cultural Psychology*, *31*, 14–32.

Kashima, Y., & Kashima, E. (2003). Individualism, GNP, climate, and pronoun drop: Is individualism determined by affluence and climate, or does language use play a role. *Journal of Cross-Cultural Psychology*, *34*, 125–134.

Kashima, Y., Kashima, E., Chiu, C-y., Farsides, T., Gelfand, M., Hong, Y-y., et al. (2005). Culture, essentialism, and agency: Are individuals universally believed to be more real entities than groups? *European Journal of Social Psychology*, *35*, 147–169.

Kashima, Y., Kashima, E., Farsides, T., Kim, U., Strack, F., Werth, L., et al. (2004). Culture and context-sensitive self: The amount and meaning of context-sensitivity of phenomenal self differ across cultures. *Self and Identity*, *3*, 125–141.

Kashima, Y., Woolcock, J., & Kashima, E. (2000). Group impressions as dynamic configurations: The tensor product model of group impression formation and change. *Psychological Review*, *107*, 914–942.

Keesing, R. M. (1974). Theories of culture. *Annual Review of Anthropology*, *3*, 73–97.

Keesing, R. M. (1994). Theories of culture revisited. In R. Borofsky (Ed.), *Assessing cultural anthropology* (pp. 301–10). New York: McGraw-Hill.

Kim, H. S. (2002). We talk, therefore we think?: A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology*, *83*, 828–842.

Kim, H. S., & Drolet, A. (2003). Choice and self-expression: A cultural analysis of variety-seeking. *Journal of Personality and Social Psychology*, *85*, 373–382.

Kim, H. S., & Markus, H. R. (1999). Deviance or uniqueness, Harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, *77*, 785–800.

Kim, J., Kim, M-s., Kam, K. Y., & Shin, H-c. (2003). Influence of self-construals on the perception of different self-presentation styles in Korea. *Asian Journal of Social Psychology*, *6*, 89–101.

Kim, K., & Markman, A. B. (2006). Differences in fear of isolation as an explanation of cultural differences: Evidence from memory and reasoning. *Journal of Experimental Social Psychology*, *42*, 350–364.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at New Look. *Psychological Science*, *14*, 201–206.

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition and Emotion*, *16*, 29–59.

Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin*, *23*, 736–742.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, *72*, 1245–1267.

Kitayama, S., Takagi, H., & Matsumoto, H. (1995). Causal attribution of success and failure: Cultural psychology of the Japanese self. *Japanese Psychological Review*, *38*, 247–280.

Kluckhohn, K. (1954). Culture and behavior. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 921–976). Cambridge, MA: Addison-Wesley.

Knowles, E. D., Morris, M. W., Chiu, C-y., & Hong, Y-y. (2001). Culture and process of person perception: Evidence for automaticity among East Asians in correcting for situational influences on behavior. *Personality and Social Psychology Bulletin*, *27*, 1344–1356.

Kobayashi, C., & Brown, J. D. (2003). Self-esteem and self-enhancement in Japan and America. *Journal of Cross-Cultural Psychology*, *34*, 567–580.

Kobayashi, C., & Greenwald, A. G. (2003). Implicit–explicit differences in self-enhancement for Americans and Japanese. *Journal of Cross-Cultural Psychology*, *34*, 522–541.

Koo, M., & Choi, I. (2005). Becoming a holistic thinker: Training effect of oriental medicine n reasoning. *Personality and Social Psychology Bulletin*, *31*, 1264–1272.

Kosic, A., Kruglanski, A. W., Pierro, A., & Mannetti, L. (2004). The social cognition of immigrants' acculturation: Effects of the need for closure and the reference group at entry. *Journal of Personality and Social Psychology*, *86*, 796–813.

Kroeber, A. L., Kluckholm, C., & Untereiner, W. (2001). *Culture: A critical review of concepts and definitions*. New York: Greenwood Press.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, *10*, 83–109.

Kudo, E., & Numazaki, M. (2003). Explicit and direct self-serving bias in Japan: Reexamination of self-serving bias for success and failure. *Journal of Cross-Cultural Psychology*, *34*, 511–521.

Kurman, J. (2001). Self-enhancement: Is it restricted to individualistic cultures? *Personality and Social Psychology Bulletin*, *27*, 1705–1716.

Kurman, J. (2003). Why is self-enhancement low in certain collectivist cultures? An investigation of two competing explanations. *Journal of Cross-Cultural Psychology*, *34*, 496–510.

Kurman, J., & Sriram, N. (1997). Self-enhancement, generality of self-evaluation, and affectivity in Israel and Singapore. *Journal of Cross-Cultural Psychology*, *28*, 421–441.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations for life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality and Social Psychology*, *73*, 1038–1051.

Lalwani, A., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, *90*, 165–178.

Lam, S-F., Lau, I. Y., Chiu, C-y., Hong, Y-y., & Peng, S.-Q. (1999). Differential emphases on modernity and Confucian values in social categorization: The case of Hong Kong adolescents in political transition. *International Journal of Intercultural Relations*, *23*, 237–256.

Latane, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication*, *46*, 13–25.

Latane, B., & Bourgeois, M. J. (1996). Experimental evidence for dynamic social impact: The emergence of subcultures in electronic groups. *Journal of Communication*, *46*, 35–47.

Latane, B., & L'Herrou, T. (1996). Spatial clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of Personality and Social Psychology*, *70*, 1218–1230.

Lau, I. Y-m., Chiu, C-y., & Lee, S-l. (2001). Communication and shared reality: Implications for the psychological foundations of culture. *Social Cognition*, *19*, 350–371.

Lee, A. Y., Aaker, J. L., & Gardner, W. L. (2000). The pleasure and pains of distinct self-construals: The role of interdependence in regulatory focus. *Journal of Personality and Social Psychology*, *78*, 1122–1134.

Lee, F., Hallahan, M., & Herzog, T. (1996). Explaining real-life events: How culture and domain shape attributions. *Personality and Social Psychology Bulletin*, *22*, 732–741.

Lehman, D., Chiu, C-y., & Schaller, M. (2004). Culture and psychology. *Annual Review of Psychology*, *55*, 689–714.

Lieberman, M. D., Jarcho, J. M., & Obayashi, J. (2005). Attributional inference across cultures: Similar automatic attributions and different controlled corrections. *Personality and Social Psychology Bulletin*, *31*, 889–901.

Lindahl, L., & Odelstad, J. (2000). An algebraic analysis of normative systems. *Ratio Juris*, *13*, 261–278.

Lockwood, P., Marshall, T. C., & Sadler, P. (2005). Promoting success or preventing failure: Cultural differences in motivation by positive and negative role models. *Personality and Social Psychology Bulletin*, *31*, 379–392.

Lyons, A., & Kashima, Y. (2001). The reproduction of culture: Communication processes tend to maintain cultural stereotypes. *Social Cognition*, *19*, 372–394.

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication?: The influence of stereotype sharedness. *Personality and Social Psychology Bulletin*, *85*, 989–1005.

Markus, H. R., & Kitayama, S. (1991). Culture and self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Markus, H. R., & Kitayama, S. (2003). Culture, self, and the reality of the social. *Psychological Inquiry*, *14*, 277–283.

Markus, H. R., Kitayama, S., & Heiman, R. J. (1996). Culture and basic psychological principles. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 857–914). New York: Guilford Press.

Masuda, T., & Kitayama, S. (2004). Perceiver-induced constraint and attitude attribution in Japan and in the U.S.: A case for cultural dependence of the correspondence bias. *Journal of Experimental Social Psychology*, *40*, 409–416.

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*, 922–934.

Masuda, T., & Nisbett, R. E. (2004). *Culture and change blindness.* Manuscript submitted for publication.

Meade, R. D., & Barnard, W. A. (1973). Conformity and anti-conformity among American and Chinese. *Journal of Social Psychology*, *89*, 15–24.

Menon, T., & Morris, M. W. (2001). Social structure in North American and Chinese cultures: Reciprocal influence between objective and subjective structures. *Journal of Psychology in Chinese Societies*, *2*, 27–50.

Menon, T., Morris, M. W., Chiu, C-y., & Hong, Y-y. (1999). Culture and the construal of agency: Attribution to individual versus group dispositions. *Journal of Personality and Social Psychology*, *76*, 701–717.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213–225.

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*, 961–978.

Miller, J. G., & Bersoff, D. M. (1992). Culture and moral judgment: How are conflicts between justice and interpersonal responsibilities resolved? *Journal of Personality and Social Psychology*, *62*, 541–554.

Miller, J. G., & Bersoff, D. M. (1994). Cultural influences on the moral status of reciprocity and the discounting of endogenous motivation. *Personality and Social Psychology Bulletin*, *20*, 592–602.

Miller, J. G., & Bersoff, D. M. (1998). The role of liking in perceptions of moral responsibility to help: A cultural perspective. *Journal of Experimental Social Psychology*, *34*, 443–469.

Miller, P., Wiley, A. R., Fung, H., & Liang, C-h. (1997). Personal storytelling as a medium of socialization in Chinese and American families. *Child Development*, *68*, 557–568.

Miyamoto, Y., & Kitayama, S. (2002). Cultural variation in correspondence bias: The critical role of attitude diagnosticity of socially constrained behavior. *Journal of Personality and Social Psychology*, *83*, 1239–1248.

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and physical environment: Holistic versus analytical perceptual affordances. *Psychological Science*, *17*, 113–119.

Morling, B., Kitayama, S., & Miyamoto, Y. (2002). Cultural practices emphasize influence in the United States and adjustment in Japan. *Personality and Social Psychology Bulletin*, *28*, 311–323.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949–971.

Muramoto, Y. (2003). An indirect self-enhancement in relationship among Japanese. *Journal of Cross-Cultural Psychology*, *34*, 552–566.

Nisbett, R. E. (1993). Violence and U.S. regional culture. *American Psychologist*, *48*, 441–449.

Norenzayan, A., Choi, I., & Nisbett, R. E. (2002). Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Personality and Social Psychology Bulletin*, *28*, 109–120.

Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin*, *131*, 763–784.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*, 653–684.

Ohbuschi, K-i., Fukushima, O., & Tedeschi, J. T. (1999). Cultural values in conflict management: Goal orientation, goal attainment, and tactical decision. *Journal of Cross-Cultural Psychology*, *30*, 51–71.

Oishi, S., & Diener, E. (2001). Goals, culture, and subjective well-being. *Personality and Social Psychology Bulletin*, *27*, 1674–1682.

Oishi, S., Wyer, R. S., Jr., & Colcombe, S. J. (2000). Cultural variation in the use of current life satisfaction to predict the future. *Journal of Personality and Social Psychology*, *78*, 434–445.

Orr, E., Mana, A., & Mana, Y. (2003). Immigrant identity of Israel adolescents from Ethiopia and the former USSR: Culture-specific principles of organization. *European Journal of Social Psychology*, *33*, 71–92.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3–72.

Peng, K., & Knowles, E. D. (2003). Culture, education, and the attribution of physical causality. *Personality and Social Psychology Bulletin*, *29*, 1272–1284.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741–754.

Pickett, C. L., Bonner, B. L., & Coleman, J. M. (2002). Motivated self-stereotyping: Heightened assimilation and differentiation needs result in increased levels of positive and negative self-stereotyping. *Journal of Personality and Social Psychology*, *82*, 543–562.

Pickett, C. L., & Brewer, M. B. (2001). Assimilation and differentiation needs as motivational determinants of perceived ingroup and outgroup homogeneity. *Journal of Experimental Social Psychology*, *37*, 341–348.

Poortinga, Y. H. (2003). Coherence of culture and generalizability of data: Two questionable assumptions in cross-cultural psychology. In V. Murphy-Berman & J. J. Berman (Eds.), *Cross-cultural differences in perspectives on the self: Nebraska Symposium on Motivation* (Vol. 49, pp. 257–305). Lincoln: University of Nebraska Press.

Rhee, E., Uleman, J. S., Lee, H. K., & Roman, R. J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. *Journal of Personality and Social Psychology*, *69*, 142–152.

Ross, M., Heine, S. J., Wilson, A. E., & Sugimori, S. (2005). Cross-cultural discrepancies in self-appraisals. *Personality and Social Psychology Bulletin*, *31*, 1175–1188.

Ross, M., Xun, W. Q. E., & Wilson, A. E. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin*, *28*, 1040–1050.

Rothbaum, F., & Tsang, B. Y-P. (1998). Lovesongs in the United States and China on the nature of romantic love. *Journal of Cross-Cultural Psychology*, *29*, 306–319.

Rothbaum, F., & Xu, X. (1995). The theme of giving back to parents in Chinese and American songs. *Journal of Cross-Cultural Psychology*, *26*, 698–713.

Rothstein, E. (1999, June 12). In a word, culture means anything, bad as well as good. *New York Times*, p. B1.

Schimmack, U., Radhakrishnan, P., Oishi, S., Dzokoto, V., & Ahadi, S. (2002). Culture, personality, and subjective well-being: Integrating process models of life satisfaction. *Journal of Personality and Social Psychology*, *82*, 582–593.

Schubert, T. W., & Hafner, M. (2003). Contrast from social stereotypes in automatic behavior. *Journal of Experimental Social Psychology*, *39*, 577–584.

Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology*, *80*, 645–654.

Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, *84*, 60–79.

Shore, B. (1996). *Culture in mind: Cognition, culture, and the problem of meaning*. New York: Oxford University Press.

Shore, B. (2002). Taking culture seriously. *Human Development*, *45*, 226–228.

Spears, R., Gordijn, E., Dijksterhuis, A., & Stapel, D. A. (2004). Reaction in action: Intergroup contrast in automatic behavior. *Personality and Social Psychology Bulletin*, *30*, 605–616.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Malden, MA: Blackwell.

Su, S. K., Chiu, C-y., Hong, Y-y., Leung, K., Peng, K., & Morris, M. W. (1999). Self organization and social organization: American and Chinese constructions. In T. R. Tyler, R. Kramer, & O. John (Eds.), *The psychology of the social self* (pp. 193–222). Mahway, NJ: Erlbaum.

Suh, E. M. (2002). Culture, identity consistency, and subjective well-being. *Journal of Personality and Social Psychology*, *83*, 1378–1391.

Suh, E. M., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology*, *74*, 482–493.

Sui, J., Zhu, Y., & Chiu, C-y. (in press). Bicultural mind, self-construal, and recognition memory: Cultural priming effects on self- and mother-reference effect. *Journal of Experimental Social Psychology*.

Sundberg, N. D., Rohila, P. K., & Tyler, L. E. (1970). Values of Indian and American adolescents. *Journal of Personality and Social Psychology*, *16*, 374–397.

Tafarodi, R. W., Wang, S-j., & Milne, A. B. (2002). When different becomes similar: Compensatory conformity in bicultural visible minorities. *Personality and Social Psychology Bulletin*, *28*, 1131–1142.

Takata, T. (2003). Self-enhancement and self-criticism in Japanese culture: An experimental analysis. *Journal of Cross-Cultural Psychology*, *34*, 542–551.

Tomasello, M. (2001). Cultural transmission: A view from chimpanzees and human infants. *Journal of Cross-Cultural Psychology*, *32*, 135–146.

Trafimow, D., Silverman, E. S., Fan, R. M-t., & Law, J. S. F. (1997). The effects of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology*, *28*, 107–123.

Trafimow, D., Triandis, H. C., & Goto, S. G. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology*, *60*, 649–655.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*, 506–520.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, *51*, 407–415.

Triandis, H. C. (2004). Dimensions of culture beyond Hofstede. In H. Vinken, J. Soeters, & P. Ester (Eds.), *Comparing cultures: Dimensions of culture in a comparative perspective*. Leiden, Germany: Brill.

Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspective on self-group relationships. *Journal of Personality and Social Psychology*, *54*, 323–338.

Tsai, J. L., Ying, Y-w., & Lee, P. A. (2000). The meaning of being Chinese and being American: Variation among Chinese American young adults. *Journal of Cross-Cultural Psychology*, *31*, 302–332.

Turban, E., & Aronson, J. (1988). *Decision support systems and intelligent systems*. Upper Saddle River, NJ: Prentice-Hall.

Uskul, A. K., Hynie, M., & Lalonde, R. N. (2004). Interdependence as a mediator between culture and interpersonal closeness for Euro-Canadians and Turks. *Journal of Cross-cultural Psychology*, *35*, 174–191.

van Baaren, R. B., Maddux, W. W., Chartrand, T., de Bouter, C., & van Knippenberg, A. (2003). It takes two to mimic: Behavioral consequences of self-construals. *Journal of Personality and Social Psychology*, *84*, 1093–1102.

Vandello, J. A., & Cohen, D. (2003). Male honor and female infidelity: Implicit cultural scripts that perpetuate domestic violence. *Journal of Personality and Social Psychology*, *84*, 997–1010.

Verkuyten, M., & Pouliasi, K. (2002). Biculturalism among older children: Cultural frame switching, attributions, self-identification, and attitudes. *Journal of Cross-Cultural Psychology*, *33*, 596–609.

Wagar, B., & Cohen, D. (2003). Culture, memory, and the self: An analysis of the personal and collective self in long-term memory. *Journal of Experimental Social Psychology*, *39*, 468–475.

Wan, C., Chiu, C-y., Peng, S., & Tam, K.-P. (in press). Measuring cultures through intersubjective norms: Implications for predicting relative identification with two or more cultures. *Journal of Cross-Cultural Psychology*.

Wan, C., Chiu, C-y., Tam, K.-P., Lee, S.-L., Lau, I. Y.-M., & Peng, S.-Q. (in press). Perceived cultural importance and actual self-importance of values in cultural identification. *Journal of Personality and Social Psychology*.

Wang, Q. (2001). Culture effects on adults' earliest childhood recollection and self-description: Implications for the relation between memory and the self. *Journal of Personality and Social Psychology*, *81*, 220–233.

Wang, Q. (2003). The emergence of cultural self-constructs: Autobiographical memory and self-description in European American and Chinese children. *Developmental Psychology*, *40*, 3–15.

Wong, R. Y., & Hong, Y-y. (2005). Dynamic influences of culture on cooperation in the prisoner's dilemma. *Psychological Science*, *16*, 429–434.

Wyer, R. S., Jr. (2004). *Social comprehension and judgment: The role of situated models, narratives, and implicit theories*. Mahwah, NJ: Erlbaum.

Yang, K. S., & Bond, M. H. (1980). Ethnic identification by Chinese bilinguals. *Journal of Cross-Cultural Psychology*, *11*, 411–425.

Zhang, J., & Shavitt, S. (2003). Cultural values in advertisements to the Chinese X-generation. *Journal Advertising*, *32*, 23–33.

PART VI

APPLICATIONS OF SOCIAL PSYCHOLOGY

Psychology and the Law

Reconciling Normative and Descriptive Accounts
of Social Justice and System Legitimacy

Tom R. Tyler

John T. Jost

> The fundamental problem of a correct psychological doctrine is not why a hungry man steals
> but the exact opposite: Why doesn't he steal?
>
> —Reich (1934/1989, p. 158)

The legal system is a codified set of rules developed to regulate interactions and exchanges among people. The authorities who make and implement legal rules draw on a fairly wide range of sources in seeking to understand, predict, and control human behavior. Some of these sources are purely cultural, in the sense that they reflect received wisdom that is passed down within specific traditions, including philosophical, civil, and political traditions. These traditions contain commonsense explanatory frameworks for understanding and regulating human behavior—"folk psychologies" in which behavior is caused by individual beliefs, desires, and intentions (e.g., Dennett & Haugeland, 1987). The extent to which folk psychology is accurate is a matter of persistent debate among philosophers, cognitive scientists, and others (Greenwood, 1991; Wegner, 2002), but there can be little doubt that it is pervasive and influential in legal contexts and countless other social situations (e.g., Malle, 2004).

In addition to cultural sources of knowledge concerning the putative causes of human behavior, there are also scientific sources that may be considered (to varying degrees) by legislators, judges, juries, and other legal decision makers when they are creating and implementing laws. The possibility exists, therefore, that current scientific opinion about the "facts" of human nature can influence the legal system (see Blasi & Jost, 2006). Research on core topics in law and psychology—including the accuracy of eyewitness testimony, the fairness of police lineups, jury decision making, and attitudes toward the death penalty—has affected legal practice, at least to some degree (e.g., Ellsworth & Gross, 1994; Hastie, Penrod, & Pennington, 1983; Heuer & Penrod, 1989; Köhnken,

1996; Loftus, 1979; Monahan & Loftus, 1982; Wells, Leippe, & Ostrom, 1979). There is also the possibility that psychology can motivate attempts to reform the legal system (i.e., to make laws more just), at least in part because it can provide a more accurate set of facts about the causes and consequences of human behavior (e.g., Carson, 2003; Haney, 1993; Kang, 2005; Krieger, 1995). In this way, psychology can contribute to the actual (as well as perceived) legitimacy of the legal system, although this process may be very slow indeed because of obstacles to interdisciplinary collaboration in general as well as the institutional conservatism inherent in legal and other social and political systems (e.g., Blasi & Jost, 2006; Meyer & Rowan, 1977; Zucker, 1977).

THE AIMS AND ACHIEVEMENTS
OF PSYCHOLOGICAL JURISPRUDENCE

Psychology can contribute to the field of law by providing legal scholars with the most accurate and up-to-date understanding of scientific knowledge possible in those areas of psychology that are most relevant for the legal system. We refer to this endeavor as *psychological jurisprudence*—the effort to shape law through knowledge of human psychology (Darley, Fulero, Haney, & Tyler, 2002). As has already been suggested, psychology as a discipline is always centrally relevant to the law, whether or not legal authorities recognize it. This is because, as Haney (1982) put it, "Laws embody theories of behavior. Legal rules, doctrines, and procedures necessarily reflect basic assumptions about human nature" (p. 191).

Psychological jurisprudence thus reflects a distinctly empirical perspective on the problems presented by the law. From this perspective, the legal conception of human action and responsibility should be based on scientific research addressing cognition, motivation, and behavior (see also Darley et al., 2002; Kang, 2005; Krieger, 1995). Much like proponents of the *legal realism* movement, we argue that the roots of effective legal doctrine must lie in an accurate understanding of the nature of the social world. Psychological jurisprudence carries this basic premise further by exploiting the methodological tools of psychology, including experimentation, in order to achieve *behavioral realism* (see also Blasi & Jost, 2006; Kang & Banaji, 2006; Krieger & Fiske, 2006). From this perspective, legal doctrines should have a strong empirical basis and should be linked explicitly to scientific evidence concerning the principles of human nature.

The Scope of This Chapter

The goal of psychological jurisprudence, as we have suggested, is to make legal assumptions about human nature as consistent with contemporary psychological knowledge as possible, that is, to close the gap between folk and scientific theories of the person. At the same time, it must be acknowledged that psychology is an evolving discipline, and scientific consensus about specific determinants of human behavior will change as new research results emerge. It is quite possible that a chapter on psychology and law 25 years from now will endorse different psychological principles than the ones that we emphasize today. What presumably will not change is the underlying value of shaping law with some attention to empirical conclusions drawn from contemporary psychological research.

In this chapter, we make no attempt to cover all the areas in which psychology is relevant to the field of law. Topics such as violence and the media, mental illness and criminal responsibility, the battered woman syndrome, eyewitness memory, effectiveness of polygraph tests, jury instructions and decision making, group profiling, and attitudes toward sentencing and capital punishment are reviewed comprehensively by Ellsworth and Mauro (1998) and are not revisited here. Instead, we focus on the implications of basic social psychological research for understanding human motivation in general and the specific types of motivation that lead people to obey the law and comply with legal authorities. Legal scholars and decision makers, we submit, would benefit considerably by reconceptualizing their approaches to regulation and taking into account a wider variety of instrumental and symbolic motivations, including the genuine desire to adhere to moral principles that are espoused by legitimate authorities on behalf of a just social system.

LAW AND THE PSYCHOLOGY OF HUMAN MOTIVATION

One of the main functions of the law is to regulate the behavior of the citizenry by maximizing the likelihood that people will comply with normative standards of conduct—that is, with socially shared definitions of acceptable conduct as they are enshrined in rules, norms, and laws. If the law is to be effective in fulfilling its regulatory role, most citizens must obey most laws most of the time (Easton, 1975; Tyler, 2006b). The need for legal authorities to secure compliance has been widely noted by legal scholars and social scientists. Fuller (1969), for instance, observed that "the lawgiver must be able to anticipate that the citizenry as a whole will . . . generally observe the body of rules he has promulgated" (p. 201). Decisions made by police officers and judges mean little if people fail to take them seriously, and laws lack importance if they do not affect public behavior (Tyler, 2006b; Tyler & Huo, 2002).

Although widespread behavioral compliance is necessary to the effective functioning of society, obtaining compliance cannot be taken for granted (e.g., Sherman, 1993). Because many laws restrict the ability of individual citizens to behave as they wish, people sometimes resist them. Mastrofski, Snipes, and Supina (1996) note that "although deference to legal authorities is the norm, disobedience occurs with sufficient frequency that skill in handling the rebellious, the disgruntled, and the hard to manage—or those potentially so—has become the street officer's performance litmus test" (p. 272). Their observational study of police encounters with the public in Richmond, Virginia, provides useful evidence concerning the frequency of noncompliance. Mastrofski and colleagues estimated an overall noncompliance rate of 22%. Citizens failed to comply with police requests 18% of the time when they were told to cease illegal behavior, 19% of the time when they were told to leave another person alone, and 33% of the time when they were instructed to stop some form of disorder. Similar results were obtained in a replication by McCluskey, Mastrofski, and Parks (1998) in Indianapolis and St. Petersburg. The overall noncompliance rate was 20%. Citizens resisted 14% of the time when asked to leave another person alone, 21% of the time when ordered to cease illegal behavior, and 25% of the time when told to stop engaging in disorder.

These studies investigated short-term compliance—that is, whether the person behaved as instructed—and not whether people willingly accepted the decisions made by the authorities, bought into their resolution of a problem, or believed that the restrictions on their behavior were reasonable or appropriate. However, as Mastrofski and colleagues (1996) note, "citizens who acquiesce at the scene can renege" (p. 283). In other words, if citizens fail to internalize legal restrictions, further police intervention will eventually be required. An important contribution of psychological jurisprudence, therefore, is to shed light on the question of how to best motivate public rule-following behavior, both in general and with respect to specific encounters with authorities (Darley, Tyler, & Bilz, 2003; Tyler, 2006b). The legal system depends on obedience, and the law and police officers and judges generally expect public deference. The power of the legal system to gain cooperation is substantially increased when strategies for obtaining deference

are based on an accurate and relatively complete model of human motivation.

PERSPECTIVES ON SOCIAL CONTROL

An underlying assumption of the legal system is that public deference is a desirable attribute. This view fits well with a "consensus" view of society (Dahl, 1956, Easton, 1965; Lipset, 1959; Parsons, 1967; Sears, 2003; Tyler, 2006a). From this perspective, all members of society benefit from stability and social order of the type produced by an effectively administered rule of law. Hence, it is beneficial when people obey the law, and facilitating such an outcome through legal authorities and institutions is an important goal.

In contrast to this perspective, "conflict"-based views of society, such as those associated with Marxism, realistic group conflict theory, social identity theory, and social dominance theory, look at society as composed of competing groups. Group distinctions may correspond to economic classes (e.g., rich vs. poor), or they could be based on ethnic or national group memberships (White vs. minority; native born vs. immigrant). However a given society is constituted, those groups that dominate through their control of cultural and economic institutions tend to establish rules that are designed to benefit their own group and recruit authorities (e.g., police) to enforce those rules. From this perspective, the rules are created and administered to benefit particular groups and will be only marginally likely to produce justice for the members of disenfranchised groups (in the best case scenario). In legal scholarship, the conflict view is reflected in the writing of Edelman and Scheingold, among others. To the degree that such a view of society is adopted, it is not unambiguously good to maintain social order. In fact, it would be better if people were to become more sensitive to injustice in the system and more resistant to the status quo; they would then be more mobilized to create social change.

The ways in which these perspectives diverge can be demonstrated by considering one legal procedure in particular—dispute mediation. It is often difficult within the legal system to produce dispute resolution decisions that are mutually acceptable to the parties who come to court. Mediation has become popular among legal authorities because it is found to produce more acceptable decisions, as reflected in both expressions of satisfaction on the part of disputants and long-term adherence to those agreements. Hence, the use of mediation is viewed by consensus theorists as a desirable way to manage everyday disputes involving landlords and tenants, merchants and customers, and many other parties.

However, from a community-organizing perspective that is sensitive to the presence of conflict in society, mediation is not so favorably evaluated. Mediation provides a mechanism by which the most highly motivated complainants can receive some measure of justice and satisfaction for their disputes. At the same time, it "cools out" those individuals who are most likely to be socially aware and committed members of the community. Once their individual level grievance has been settled, they have less motivation to join with others to work at a collective level to rectify injustices. So, for example, a merchant may engage in unfair pricing practices but resolve grievances individually only with those who are most angered by their actions. This saps energy from politically motivated collective action aimed at establishing regulations that would constrain the merchant to act more ethically. While we recognize that the consensus and conflict perspectives provide very different normative perspectives on issues of compliance and settlement, our focus is on the effectiveness of different approaches for motivating compliance, setting aside the ultimate desirability of being able to obtain widespread compliance.

THE DETERRENCE MODEL

In recent decades the exercise of legal authority has become increasingly associated with the use of threat and punishment aimed at deterring people from engaging in criminal behavior (e.g., Kahan, 1999; Nagin, 1998). From this perspective, the focus is (and should be) on the power of legal authorities and institutions to shape behavior by threatening to deliver (or by actually delivering) negative sanctions for rule breaking. Within legal circles, this way of viewing the relationship between legal authorities and citizens is referred to as the deterrence or social control model, and it is this model of human behavior that—for better or worse—currently dominates law and public policy.

Motivational Assumptions of the Deterrence Model

To implement deterrence strategies police officers carry guns and clubs, and they are empowered to threaten citizens with physical injury and incapacitation, among other penalties. The goal is to establish legal authority and, as Reiss (1971) points out, "The uniform, badge, truncheon, and arms all may play a role in asserting authority" in the effort to "gain control of the situation" (p. 46). The police thereby seek to gain control over the individual's behavior "by manipulating an individual's calculus regarding whether 'crime pays' in any particular instance" (Meares, 2000, p. 396). More generally, agents of the legal system who are charged with producing compliant behavior concern themselves with shaping environmental contingencies in such a way that citizens will be faced with the prospect of heavy losses (e.g., incarceration) that are intended to outweigh the anticipated gains of engaging in criminal behavior. Judges, for example, attempt to influence people's acceptance of their decisions by threatening fines or jail time for failure to comply. The deterrence model dictates that the responsibility of lawmakers is to decide which acts should be prevented and then to specify sufficiently strict penalties—generally fines or prison terms—so that the prohibited behavior is rarely enacted.

The notion that people's behavior with respect to the law is shaped by calculations of expected gains and losses is a core premise of rational choice theory, as derived

from neoclassical economics (Blumstein, Cohen, & Nagin, 1978; Nagin, 1998). It is assumed that most people will calculate expected utilities by multiplying the probability of an outcome (e.g., getting caught for armed robbery or drunk driving) by its valence (very, very bad). If the laws are well calibrated, people will arrive at the desired conclusion that they should follow the law. Thus, rational self-interest is the motivational engine of the deterrence/social control model. To regulate behavior, this model suggests that decision makers should adjust criminal sanctions to the needed level so that the expected losses associated with lawbreaking will minimize the likelihood that people will break the law.

Research supports the notion that variations in the perceived certainty and severity of punishment do affect people's compliance with the law, at least to some degree. In particular, people's behavior is often, though not always, shaped by their estimate of the likelihood that if they disobey the law, they will be caught and punished (see Nagin & Paternoster, 1991; Paternoster, 1987, 1989; Paternoster & Iovanni, 1986; Paternoster, Saltzman, Waldo, & Chiricos, 1983). At the same time, however, perceptions of the likelihood of being caught and punished seem to have a relatively minor influence on people's behavior, as we will see (MacCoun, 1993; Robinson & Darley, 1995, 1997; Ross, 1982). Consequently, social control strategies based exclusively on a deterrence model of human behavior have had at best limited success (see also Tyler, 1997a, 1997b, 1997c, 1998).

Problems with the Deterrence Model

The deterrence model, with all of its motivational assumptions, has had dramatic effects on the nature of American society. Consider the case of the U.S. prison population (Haney & Zimbardo, 1998). Because of the widespread belief that crime is deterred by the threat of punishment and/or the experience of punishment, a massive number of citizens have been convicted and sentenced to serve time in prison. Today, the United States is a world leader in the proportion of citizens it holds in prison. In the year 2000 there were over 2 million Americans in jail or prison, a large percentage of the adult population (U.S. Department of Justice, 2001), far surpassing incarceration rates in Europe and elsewhere (Garland, 2001). Given the heavy costs of imprisonment to individuals and communities (especially members of racial and ethnic minority groups, which are overrepresented in the prison system), it is crucial to ask whether the deterrence model is based on sound psychological principles.

Costs of Surveillance

Assuming unlimited resources on the part of law enforcement officials, there is probably nothing inherently untenable (from a purely pragmatic point of view) about controlling people's behavior through threats of punishment. One of the key problems with sanctioning systems, however, is that they require near-constant surveillance of individual behavior. For obvious reasons, people are strongly motivated to hide their behavior from authori-

ties to avoid punishment; authorities must therefore develop surveillance systems for detecting rule-breaking behavior. Sometimes surveillance is easy, because the structure of the situation makes it easy. For example, wage earners' incomes are easy for the government to monitor, because businesses withhold percentages from each paycheck and send the withheld amount to the government. This makes tax violations among this group relatively easy to prosecute. In other cases, however, surveillance can be quite difficult. The police, for example, have tremendous difficulty monitoring public behavior in order to identify people who are using illegal drugs, just as tax authorities have trouble monitoring the incomes of street vendors, waiters and waitresses, and small business owners.

For all these reasons, as Meares (2000) notes, the effectiveness of "instrumental means of producing compliance always depend[s] on resource limits" (p. 401). The relevant questions are how much in terms of financial and other resources are authorities willing to expend in order to control crime, and how much power to intrude into citizens' lives are people willing to allow the authorities to have?

Furthermore, resources must be deployed in strategic and cost-effective ways. Sherman (1998), for example, notes that within the United States, police resources are typically used more in response to political pressures than to actual crime threat levels. As a result, police officers do not most heavily patrol the highest crime areas, so the ability of the police to deter crime is typically suboptimal. Sherman suggests that a greater effort is needed to put surveillance where the crime problem lies. Current deployments reflect the reality that public resources are allocated in ways that respond to political pressures, and the effective implementation of social control strategies often conflicts with those pressures.

The deterrence model probably works best in the case of crimes that are committed for instrumental reasons. For example, car theft, burglary, and crimes of this type are at least to some extent motivated by calculations about the costs and benefits expected from lawbreaking behavior. Thus, deterrence approaches work best in affecting the occurrence of instrumentally motivated crimes. They are significantly less effective in controlling criminal behavior that is motivated by factors other than economic gain.

Ross (1982) focuses on the problem of drunk driving to outline some of the problems associated with using deterrence to shape law-related behavior. He suggests that raising risk estimates to a level that is high enough to lower the rate of lawbreaking behavior, while not necessarily impossible, involves prohibitively high costs in terms of police manpower and citizen willingness to accept state intrusions into their personal lives. Interestingly, Ross finds that changes in laws can lead to short-term declines in lawbreaking because the high level of media exposure to police activities leads people to temporarily overestimate the risks of being caught and punished for lawbreaking behavior. Ross further points out that even the intensive efforts of Scandinavian authorities to create high estimates of risk using random road-

blocks and other similarly expensive and intrusive law enforcement measures are insufficient to create and maintain subjective risk estimates that are high enough to deter drunk driving over the long term.

As we have noted, many of the problems associated with deterrence-based strategies identified by Ross (1982) and others are structural in nature and involve variations in the degree to which the police are able to monitor citizens' behavior. This suggests that there should be situations in which deterrence strategies will be more or less effective in deterring lawbreaking behavior. From a deterrence approach, the two key variables determining the extent to which such strategies will be effective are (1) the ease of behavioral surveillance and (2) the level of resources that society is willing to devote to the task of surveillance.

Insensitivity to the Magnitude of Punishment

Deterrence works reasonably well in at least some cases of murder, mainly because society has devoted considerable resources to preventing murder and enforcing penalties for it. The objective risk of being caught and punished for murder is relatively high: approximately 45% (Robinson & Darley, 1997). The likelihood of being caught for committing a murder is high enough for deterrence to be effective in lowering the murder rate. Even in this case, however, criminals are not as sensitive to the magnitude of the penalty as they are to the estimated probability of being apprehended. As a result, capital punishment does not serve to deter murder more effectively than does life imprisonment (Ellsworth & Mauro, 1998).

Studies consistently find that the most important issue to people who are deciding whether or not to break the law is their estimate of the likelihood of being punished for their actions and not the expected severity of their punishment (e.g., Nagin & Paternoster, 1991). Consequently, societies cannot enforce rules simply by developing more and more draconian punishments. To be effective, authorities must engage in extensive surveillance strategies that increase the likelihood of detection. In other words, they must increase the number of supervisors who are watching employees or increase the size of the police force that is watching citizens. Effective strategies are, therefore, inevitably costly. There are no deterrence-based quick fixes that can be gained cheaply—despite the frequent suggestion among policymakers that a few instances of dramatically severe punishment will depress crime.

With regard to less serious crimes, the deterrence model is even less effective. Despite the expenditure of significant societal resources in the form of increased police efforts in the "War on Drugs," deterrence strategies have failed to improve rates of public compliance with drug laws. MacCoun (1993) estimates that only about 5% of the variance in people's use of illegal drugs is attributable to their perceptions of the likelihood of being caught and punished for rule breaking. For most people, and especially for addicts, the decision to obtain or con-sume controlled substances is simply not based on a rational calculation of prospective legal costs.

Subjectivity and Bias in Risk Estimates

There are many other crimes that are motivated not by instrumental concerns but by temporary emotional states—what Loewenstein, Prelec, and Shatto (1998) refer to as a "hot state." For example, crimes such as rape, assault, and many murders occur on the "spur of the moment" and in the "heat of passion." In such cases, the assumption that rational calculations of costs and benefits enter into the perpetrator's "decision" about whether or not to commit such crimes is naive in the extreme. In part, this is because of intrapersonal "empathy gaps": The person who is in a "hot state" cannot access or identify with how he or she will think or feel later upon returning to a "cold state" (Loewenstein et al., 1998). Thus, crimes of passion as well as crimes committed under states of intoxication are relatively unaffected by deterrence strategies, regardless of the actual or even perceived likelihood of being caught and punished for wrongdoing.

The lack of a direct correspondence between objective and subjective risks leads to another problem with the psychology of the deterrence model, namely, its failure to take into account "threshold effects." That is, to influence people's behavior at all, risk estimates need to be high enough to exceed some threshold of being psychologically meaningful (Ross, 1982; Teevan, 1975). In most situations the objective risk of being caught and punished is quite low. For example, the approximate objective risk of being caught, convicted, and imprisoned for rape is 12%; for robbery it is 4%; and for assault, burglary, larceny, and motor vehicle theft the incarceration rate is approximately 1% (Robinson & Darley, 1997). Of course, psychologists know that subjective estimates of risk are stronger determinants of people's behavior than are objective risks. However, research suggests that subjective risk estimates for infrequent events are, if anything, even lower than objective risks (e.g., Bazerman, 1990). Furthermore, it is reasonable to assume that people's estimates of whether or not they will be apprehended by the police are subject to egocentric biases and the "illusion of invulnerability" (Dunning, 1999; Taylor & Brown, 1988).

Limits of Extrinsic Motivation

There are still other reasons to think that the deterrence model is based on flawed motivational assumptions. For most crimes, the resources devoted to law enforcement are low and the opportunities for cheating are high. This is as true of white-collar crime and corporate wrongdoing as it is of other crimes (e.g., Moore & Loewenstein, 2004). In many cases, imposing monetary penalties can paradoxically *decrease* rather than increase rates of compliance, because of the fact that people construe the fine as a price and determine that they are willing to pay the cost in order to benefit from rule-breaking behavior (Gneezy & Rustichini, 2000; Tenbrunsel & Messick, 1999).

An analysis in terms of general principles of human motivation further suggests that if people comply with the law only in response to coercive power, they will be less likely to obey the law in the future because acting in response to external pressures diminishes internal motivations to engage in a behavior (e.g., Brehm, 1966; French & Raven, 1959; Tyler & Blader, 2000). This follows from the well-known distinction in social psychology between intrinsic and extrinsic motivation. Research on intrinsic motivation shows that when people are motivated solely by the prospect of obtaining external rewards and punishments they become less likely to perform the desired behavior in the absence of such environmental reinforcements (e.g., Deci, 1975). On the other hand, if people are motivated by intrinsic reasons for behaving in a certain way, their compliance becomes much more reliable and less context dependent.

Studies of regulatory authorities indeed demonstrate that seeking to regulate behavior through the use of threat serves to undermine people's commitment to rules and authorities, compared to other methods of regulation (Frey, 1994; Frey & Oberholzer-Gee, 1997). From a motivational perspective, instrumental approaches are not self-sustaining and require the maintenance of institutions and authorities that can keep the probability of detection for wrongdoing at a sufficiently high level to constantly motivate the public through external means (i.e., the threat of punishment). Over time it becomes more and more important to have external constraints in place, for whatever intrinsic motivation people originally had is gradually "crowded out" by extrinsic concerns.

Distrust and Suspicion

The use of surveillance systems also has deleterious effects on the social climate of groups. The use of surveillance implies distrust, which decreases people's ability to feel positively about themselves, their groups, and the system itself (e.g., Kramer & Jost, 2002; Kramer & Tyler, 1996). Furthermore, people may experience intrusions into their lives as procedurally unfair, leading to anger and other negative emotions often associated with perceptions of injustice (e.g., Gurr, 1970; Tyler & Smith, 1998). Whether surveillance works or not, it is often demotivating and introduces new costs in terms of distrust and perhaps even paranoia in subsequent social interaction. Such costs are borne by groups, organizations, and societies to which people belong, as they lose the gains that occur when people are willing to cooperate with each other. Research suggests that the increasing use of deterrence strategies and social control has exerted precisely this type of negative influence on the U.S. social climate. It has created an adversarial relationship between legal authorities and members of the communities they serve, especially with respect to racial and ethnic minority group members (Tyler & Huo, 2002), leading the public to grow less compliant with the law and less willing to help the police to fight crime (Sunshine & Tyler, 2003b).

Achieving a better understanding of the psychology of human motivation should be of paramount interest to le-

gal authorities, to members of the legal profession, and to those working within legal institutions such as the courts, the police, and prisons. The problems inherent in implementing many laws—and the inadequacy of the deterrence model in general—have led to widespread calls from legal authorities and scholars for social science to help in understanding how to secure the effective rule of law. Their concerns suggest that current models of the determinants of human behavior are not providing legal authorities with an adequate basis for effective social regulation. This presents an important opportunity for psychologists to put forward a new and more empirically grounded perspective on the relationship between the individual and society and the following of social rules. Taking psychological jurisprudence seriously means linking our understanding of motivation and social influence in legal contexts to a broader psychological understanding of the person (e.g., see Cohn & White, 1990; Krislov, Boyum, Clark, Shaefer, & White, 1966; Melton, 1985; Tapp & Levine, 1977).

AN ALTERNATIVE MODEL BASED ON SOCIAL PSYCHOLOGICAL PRINCIPLES

Most legal scholarship, as we have seen, addresses issues of motivation and rule following in the context of deterrence and social control (i.e., by imposing external constraints on people largely through the threat of punishment). Social psychologists can contribute to a more sophisticated understanding of compliance. Specifically, a good deal of research indicates that self-regulatory motivations are activated when people believe that the law reflects their views about right and wrong and that it is therefore a moral responsibility and even an obligation to conform to the law. Consequently, people who identify with legal authorities and imbue the legal system with legitimacy will voluntarily abide by laws and defer to authorities (Darley et al., 2003; Jost & Major, 2001; Tyler, 2006a; Tyler & Blader, 2000). We develop this position in more detail in the remainder of this chapter.

Commitment Based on Legitimacy and Moral Values

Ever since Kurt Lewin's (1936) field theory, social psychologists have assumed that behavior is determined by two main forces. The first is the pressure of the situation or the environment, and the second includes the motives and perceptions that the person brings to the situation. In Lewin's famous equation, behavior is understood to be a function of the person and the environment: $B = f(P, E)$. An expanded conception of the person term includes the set of social and moral values that shape the individual's thoughts and feelings about what is ethical or normatively appropriate to do. We focus on two such values: (1) the conviction that following the rules of the community is (in most cases) the morally appropriate thing to do, and (2) commitment to the notion that if the rules are fair and legitimate, they ought to be obeyed.

From a social psychological perspective, the first step is to recognize that the legal system depends at least in part

on the willingness of citizens to consent to the operation of legal authorities and to actively cooperate with them. Second, willing acceptance comes most quickly and completely to the extent that people view the law as (1) determined and implemented through procedurally fair means, and (2) consistent with cherished moral values.

These notions are consistent with the proposals of Ayres and Braithwaite (1992), who suggest that legal authorities should approach citizens by appealing to their moral values. The idea is that if authorities are successful in joining legal and moral concerns, it will be much easier to isolate the relatively small number of citizens who do not share the values of the community. Third, if the overwhelming majority accepts that the law is based on sound moral principles carried out by legitimate authorities, the majority is likely to exert strong social pressure on deviants to conform to normative standards. This is a kind of social self-regulation that enhances the effectiveness of legal authorities by freeing them to pay attention to those problems or people who, for various reasons, are not amenable to self-regulation (Ayres & Braithwaite, 1992).

The Concept of Legitimacy

Modern discussions of legitimacy are usually traced to the writings of Weber (1968) on authority and the social dynamics of authority (e.g., Zelditch, 2001). Weber, like Machiavelli and others before him, argued that successful leaders and institutions use more than brute force to execute their will. More specifically, they strive to win the consent of the governed so that their commands will be voluntarily obeyed (Tyler, 2006a). As Kelman (1969) put it: "It is essential to the effective functioning of the nation-state that the basic tenets of its ideology be widely accepted within the population. . . . This means that the average citizen is prepared to meet the expectations of the citizen role and to comply with the demands that the state makes upon him, even when this requires considerable personal sacrifice" (p. 278). Widespread voluntary cooperation with the state and the social system allows authorities to concentrate their resources most effectively on pursuing the long-term goals of society. The authorities do not need to provide incentives or sanctions to all citizens to get them to support every rule or policy they enact.

Legitimacy, according to this general view, is a quality that is possessed by an authority, a law, or an institution that leads others to feel obligated to accept its directives. It is, in other words, "a quality attributed to a regime by a population" (Merelman, 1966, p. 548). When people ascribe legitimacy to the system that governs them, they become willing subjects whose behavior is strongly influenced by official (and unofficial) doctrine. They also internalize a set of moral values that is consonant with the aims of the system, and—for better or for worse—they take on the ideological task of justifying the system and its particulars (see also Jost & Major, 2001).

Although the concept of legitimacy has not featured prominently in recent discussions of social regulation with respect to law-abiding behavior, there is a strong intellectual tradition that emphasizes the significance of developing and maintaining positive social values toward cultural, political, and legal authorities (Easton, 1965, 1975; Krislov et al., 1966; Melton, 1985; Parsons, 1967; Tapp & Levine, 1977). This work builds on the sociological tradition associated with Weber and others. According to "consensus" theories such as these, the smooth functioning of society depends on the existence of supportive attitudes and values among members of the population in general. Presumably, attitudes and values that support the social system begin to develop during childhood and adolescence as part of the process of political (and legal) socialization (e.g., Cohn & White, 1990; Niemi, 1973).

The value of cultivating system legitimacy consists in its enabling the effective (and, ideally, fair) exercise of social authority. While authorities can exercise power directly through the promise of rewards or the threat of punishment, such approaches to deterrence are expensive, inefficient, and psychologically naive. They may be especially problematic during times of instability or crisis, when authorities need the support of the people at a time in which they lack control over resources. An organization or society whose governance is motivated only by incentives and sanctions is at risk of disintegrating during times of trouble or change. In contrast, if a system enjoys widespread legitimacy, authorities can appeal to members based on their shared purposes and values, providing the system with much-needed stability. From this perspective, legitimacy is a highly desirable feature of social systems (see also Tyler, 2006b; Tyler & Huo, 2002).

Underlying this generally positive view of the role of legitimacy and social values in motivating cooperation with the social system is the tenet of consensus theories that there is a mutual benefit that comes from voluntarily accepting societal norms. According to this view, the rulers and the ruled alike gain from having "a stable social and political order" that is helped by widespread shared beliefs that the system is legitimate and consistent with people's moral values (Sears, 2003, p. 322). Clearly, legitimacy and stability facilitate regulation—the process whereby authorities seek to bring the behavior of individuals into line with system rules. The police and courts depend very heavily on the widespread voluntary compliance of most of the citizens most of the time (Tyler, 2006b). This compliance presumably allows authorities to focus their attention on those individuals and groups whose behavior seems to be responsive only to threats of punishment. The legal system would be overwhelmed immediately if it were required to regulate the behavior of the majority citizens solely through sanctioning or the threat of sanctioning.

Legitimacy, as we have suggested, has many appealing features as a possible basis for the rule of law. On its face it appears to be an all-purpose mechanism of social coordination, insofar as people feel obligated to obey whatever laws or decisions authorities make, within some realm of legitimacy. Much as studies of confidence and trust in government focus on people's overall evaluations of the government, its institutions, and its authorities (Citrin & Muste, 1999), studies of "legal consciousness" focus on whether people have "trust and confidence" in

the legal system, whether they think that the law works to help everyone, and how and when people have duties and obligations to legal institutions and authorities (Ewick & Silbey, 1988; Finkel, 1995; Flanagan & Longmire, 1996; Hamilton & Sanders, 1992; Merry, 1990).

Research by Tyler (2006b) demonstrates that perceptions of system legitimacy do shape everyday compliance with the law, which is a conclusion that is also supported by other studies (Sunshine & Tyler, 2003b; Tyler & Huo, 2002). Furthermore, perceived legitimacy seems to have more influence on compliance than do subjective assessments of the likely risk of punishment. When people perceive the system as legitimate, they feel an intrinsic moral obligation to comply with its demands.

Internalization of Moral Values

Moral values are influential because they are based on internalized feelings of responsibility to follow certain principles (see Robinson & Darley, 1995; Tyler & Darley, 2000). A key feature of moral values is that people feel personally obligated to adhere to them, and they feel guilty when they fail to do so. Hence, moral values—once they exist—are self-regulatory in nature; people who possess them are strongly motivated to bring their conduct into line with normative standards. The internalized sense of morality is central to the work of, among others, Freud, Weber, and Durkheim (e.g., see Sunshine & Tyler, 2003a). Hoffman (1977) writes:

> The legacy of both Sigmund Freud and Emile Durkheim is the agreement among social scientists that most people do not go through life viewing society's moral norms as external, coercively imposed pressures to which they must submit. Though the norms are initially external to the individual and often in conflict with [a person's] desires, the norms eventually become part of [a person's] internal motive system and guide [a person's] behavior even in the absence of external authority. Control by others is thus replaced by self control [through a process labeled internalization]. (p. 85)

The idea is that internalized values become self-regulating, so that people accept and act on the basis of values that produce respect for societal institutions, authorities, and rules. Public standards are taken on as private values that are associated with a moral responsibility to act in accordance with ethical judgments about what is right and wrong. Presumably, this occurs during childhood as part of the socialization process (Cohn & White, 1990; Greenstein, 1965; Hess & Torney, 1967; Hyman, 1959; Merelman, 1966; Niemi, 1973). Robinson and Darley (1995) conclude that people's moral values form during childhood socialization and are not easy to change later in their lives.

The significance of morality is illustrated by research on punishment. Studies demonstrate that people's views about appropriate sentencing decisions in criminal cases are driven by moral judgments about deservingness rather than by instrumental judgments concerning how to deter future criminal conduct (Carlsmith, Darley, & Robinson, 2002; Darley, Carlsmith, & Robinson, 2000).

People accept that a punishment is appropriate when it accords with their moral sense of what is appropriate given the level and type of wrong committed. More generally, research shows that people are more willing to comply with the law to the extent that they view it as consistent with their moral values (e.g., Robinson & Darley, 1995; Tyler, 2006b). As a consequence, an important question for the law is the degree to which it is congruent with public moral values. If people correctly understand the law, and if the law truly reflects moral standards of the community, then the internalized sense of morality acts as a force for law-abidingness.

Justice as Intrinsically Motivating

Our value-based model of motivation to obey the law is also consistent with two additional principles of human behavior that should be made more explicit. One is that people, as a general rule, desire justice; that is, they are motivated by fairness concerns. In summarizing decades of research, Lerner (2003) refers to this as the "justice motive," and he suggests that it cannot be reduced to considerations of self-interest. There is indeed evidence that people experience unfair advantages (as well as disadvantages) to be psychologically aversive and that they are happier when justice is served (e.g., Boll, Ferring, & Filipp, 2005; Loewenstein, Thompson, & Bazerman, 1989; Walster, Walster, & Berscheid, 1978). This work provides a reasonably strong basis for assuming that people *want* to follow the law and that they want others to follow the law—*provided that they also believe that the law is fair and just*. To draw on the justice motive as a source of legal obedience, therefore, authorities must pursue policies that are generally consistent with people's sense of right and wrong (Sunshine & Tyler, 2003a).

A second and related fact about human behavior, to which we have alluded already, is that people are more consistently motivated by intrinsic than extrinsic considerations (e.g., Deci, 1975; Frey & Oberholzer-Gee, 1997). Thus, to the extent that social values become internalized, they become a part of the person and lead him or her to exercise self-regulation so that their behavior is consistent with the principles and values that define their sense of themselves. Under such circumstances, people may be expected to follow the rule of law not out of a temporary fear of surveillance but because they have internalized the values that are codified by the legal system. Deference, then, will not be experienced as a cost but as the kind of benefit that comes from doing what is defined as morally right.

A value-based perspective on human motivation suggests the importance of developing and sustaining a civic culture in which people abide by the law because they feel that it is morally required. For this model to work, society must create and maintain public values that are conducive to following justice norms. Political scientists refer to this set of values as a "reservoir of support" for government and society (Dahl, 1956). Although it may not always be easy for authorities to maintain high reservoir levels, a value-based model is consistent with a social psychological understanding of how authorities can ef-

fectively regulate citizen behavior, maintain social order, and promote an effective, well-functioning society by developing and maintaining a culture of supportive social values that will be internalized by the citizenry.

The value-based model we have outlined in this section avoids many of the pitfalls of the deterrence model. Specifically, it does not require extensive surveillance efforts, is more sophisticated concerning the genuine causes of human behavior, engages intrinsic (and not just extrinsic) motivation, and fosters a positive social climate based on a shared commitment to moral values rather than a negative social climate based on suspicion and distrust. But there is yet another important advantage of our value-based model to which we have only alluded thus far. To the extent that people are in fact internalizing appropriate moral values, deferring to legal authorities who implement fair procedures, and obeying laws that are truly just, then the model of human behavior we have sketched will lead not only to an efficient and well-ordered society but also to one that has a profoundly legitimate basis for regulating the behavior of its citizenry.

SYSTEM JUSTIFICATION AND THE PROBLEM OF FALSE CONSCIOUSNESS

Up to this point, we have assumed that fostering perceptions of system legitimacy is uniformly desirable, especially from the standpoint of legal authorities. Astute readers will have realized, however, that the power of legitimacy to motivate compliance can be abused. That is, Machiavellian leaders can use their influence to lend legitimacy to immoral as well as moral causes and to maintain consensual support for oppressive as well as beneficent forms of social control (e.g., Jost & Major, 2001). And, as we have noted, conflict models of society suggest that elite individuals and groups are highly motivated to wield their power and the legitimacy they possess in order to achieve ends that suit their own interests. The social psychological processes that we describe, therefore, are morally ambiguous. We return later to the question of whether it is possible to reconcile descriptive accounts of how and why people actually perceive systems and authorities as legitimate and normative accounts of whether people *ought* to perceive them as legitimate and therefore worthy of compliance.

Legitimacy, Authorization, and Obedience to Authority

Legitimacy, we have suggested, is the property of a rule or authority that leads people to feel morally obligated to voluntarily defer to that rule or authority. From a social influence perspective, a legitimate authority is a person or group that is regarded by others as being entitled to have its rules accepted and followed (e.g., French & Raven, 1959). In this way, ascriptions of legitimacy play a key role in persuasion and motivation, especially to the extent that legitimate authority figures can lead people to suspend considerations of self-interest and to ignore personal goals and values that are at odds with officially sanctioned values. This capacity for systems and authorities that are perceived as legitimate to influence and motivate others is a double-edged sword. It means that people will be loyal and committed to following the norms and rules of the system, but it also means that they can be made to obey unjust laws and pursue cruel objectives in the name of the system.

In their study of obedience to authority, Kelman and Hamilton (1989) use the term "authorization" to capture the process by which followers license their leaders to define what counts as appropriate behavior in a given situation. In this way, citizens, employees, soldiers, and others abdicate moral responsibility and follow orders conscientiously. Kelman and Hamilton note "Behaviorally, authorization obviates the necessity of making judgments or choices. Not only do normal moral principles become inoperative, but—particularly when the actions are explicitly ordered—a different type of morality, linked to the duty to obey superior orders, tends to take over" (p. 16).

Imbuing institutions and authorities with legitimacy has the consequence of "authorizing" certain behavior that might not otherwise be seen as normatively appropriate. In fact, people often view their own moral values as irrelevant when a legitimate authority figure is present. An example is provided by Milgram's (1974) famous experimental studies of obedience to authorities. As Kelman and Hamilton (1989) note:

> It is interesting that, in the postexperimental debriefing, he [a participant who shocked the confederate at the maximum level] seems unable to understand a question about whether there are any conditions under which he might have stopped administering shocks to [the confederate]. As far as he is concerned, he did stop—and he seems dismayed to learn that some participants did not stop. He stopped; it was the [authority] who continued. His hand may have been on the switch, but the decision and the responsibility were clearly the experimenter's. (p. 155)

Thus, under some (perhaps exceptional) circumstances, granting legitimacy to authorities and internalizing the moral values that are encouraged by any given system can lead good people to engage in horrific activities (see also Darley, 2004). The problem of obedience to a legitimized authority figure, however, is only one of several related problems that can result from idealizing the system and its authorities.

False Consciousness and the Legal System

As we have noted, consensus theorists influenced by Weber and his intellectual heirs assume that virtually everyone benefits from a stable and effective social system. This assumption is not shared by conflict theorists inspired by Marx and his followers (see Zelditch, 2001). Rather, a Marxian perspective suggests that members of elite or dominant groups have an ideological interest in maintaining the existing social system, whereas the objective interests of subordinated groups lie in opposing the status quo (e.g., Parkin, 1971). However, due to socialization pressures, members of subordinated groups often

develop ideological beliefs that are supportive of the system rather than of their own needs and interests.

The dominant class, in other words, uses its control over cultural institutions—including churches, schools, and the mass media—to promulgate an image of society that obscures the oppression of the subordinate class and legitimates the status quo by portraying existing institutions as fair, appropriate, and just. From a Marxian perspective, then, internalized social values may reflect distortions of the true social situation. They may reflect, in other words, "false consciousness" (e.g., Jost, 1995; Tyler & McGraw, 1986). Members of disadvantaged groups, according to this view, are often made worse off by acquiescing to dominant ideologies and accepting existing authorities and institutions as legitimate.

It has been argued that citizens' relatively widespread acceptance of the legitimacy of legal authorities and institutions reflects some degree of false consciousness (e.g., Fox, 1999). Scheingold (1974), for instance, proposed that at a system level of analysis, the function of the courts is to provide disadvantaged people with symbolic satisfaction rather than to make rulings that create substantive justice. And in an article that was subtitled "Let Them Eat Due Process," Haney (1991) further suggested that many litigants are manipulated into believing that they are receiving procedural justice from the court system when in fact they are being hoodwinked into accepting poor outcomes.

An even stronger indictment of the legal system comes from the work of Sidanius and Pratto (1999), who argue that the law is in fact an instrument of oppression that is wielded by elite members of society against those who are disadvantaged, especially members of racial and ethnic minority groups. From these perspectives, legal institutions do not deserve the degree of legitimacy that they currently enjoy.

System justification theory, which builds on Marxian analyses of ideology and false consciousness and Lerner's (1980) work on the belief in a just world, among other influences, focuses on the social psychological tendency to rationalize the status quo and to see "what is" as "what ought to be" (Jost & Banaji, 1994). The theory suggests that people are apt to exaggerate the legitimacy of existing institutions, in part because they have little choice but to live with and adapt to them (Jost & Hunyady, 2002). Because it is psychologically aversive to feel as if one is subject to unjust treatment, capricious authorities, and oppressive institutions, people tend to believe that the status quo is fair and just. Thus, they develop stereotypes and ideologies—some of which may be operating at a nonconscious level of awareness—that bolster existing institutions and narrow the range of conceivable and acceptable alternatives to the status quo (e.g., see Jost, Banaji, & Nosek, 2004).

System justification—defined as the tendency to justify the status quo—is often observed in the economic realm. People tend to uncritically accept meritocratic explanations for inequality (Jost, Pelham, Sheldon, & Sullivan, 2003; Tyler & McGraw, 1986) and to blame individuals rather than systems for the existence of poverty (Hochschild, 1981; Kluegel & Smith, 1986). Further-

more, research by Jost, Blount, Pfeffer, and Hunyady (2003) suggests that people who live under capitalist systems are motivated to believe in "fair market ideology," according to which market-based outcomes and exchanges are not only efficient but also inherently fair and legitimate. Consistent with this general theoretical perspective, Tyler (2004) showed that people who endorsed free-market procedures as fair were more likely to resist governmental intervention in the economic sphere and less likely to support social change aimed at increasing economic justice.

Why would people engage in system justification, especially when it is contrary to their own individual and collective interests? Jost and Hunyady (2002) proposed that system-justifying ideologies serve a palliative function in that they decrease anxiety, uncertainty, guilt, frustration, and dissonance and increase people's satisfaction with their own situations and with the status quo in general. Research by Wakslak, Jost, Tyler, and Chen (in press) provides support for these affective benefits of system justification. Specifically, they demonstrated that increased system justification—either measured in terms of strength of ideological endorsement or experimentally manipulated by activating a "Horatio Alger" rags-to-riches mindset—produces a significant reduction in emotional distress, which in turn leads to a significant reduction in support for social change and the redistribution of resources.

Although system justification theory has only recently been applied to the law (Blasi & Jost, 2006; Greenwald & Krieger, 2006; Hanson & Yosifon, 2004; Kang, 2005; Travis, 2005; see also Jolls & Sunstein, 2006), there are numerous consequences of the theory for the legal system and its legitimacy. For one thing, the fact that people come to believe that the system is fair and legitimate does not in any way mean that it actually is. Researchers have shown that are a number of cognitive and motivational biases that lead people to rationalize outcomes and procedures that are neither favorable nor inherently just (see Jost et al., 2004, for a review). This work integrates and builds on earlier evidence that people evince a remarkable capacity to acquiesce in the face of injustice (e.g., Crosby, 1982; Major, 1994; Moore, 1978; Tyler & McGraw, 1986). Thus, system justification theory suggests that, for a wide range of psychological reasons, false consciousness about legal and other institutions may be relatively widespread (see also Jost, 1995; Jost & Banaji, 1994).

Is there a way out of this predicament? How can citizens and their leaders increase the likelihood that system legitimacy is genuine rather than illusory? There may be no way of guaranteeing that legal institutions will deliver substantive justice, but it seems to us that adhering to normatively defensible standards of procedural fairness maximizes the likelihood of obtaining outcomes that are objectively (and not just subjectively) fair (see also Jost & Ross, 1999). In other words, efforts to attain system legitimacy should be based not merely on bringing about consensus *per se* but on establishing and protecting standards that are likely to guarantee genuine or substantive fairness.

NATURALIZING JUSTICE: RECONCILING NORMATIVE AND DESCRIPTIVE ACCOUNTS OF LEGITIMACY

The social psychological principles we have described throughout this chapter are morally ambiguous. That is, they can be used equally to win compliance for just and unjust systems and authorities. However, for our model to contribute to genuine (rather than merely apparent) system legitimacy, there must be congruence between subjective and objective justice considerations. That is, the system that seems morally justified must in fact *be* morally justified, according to normative theories of justice as developed, for example, in the philosophical traditions of Kant, Mill, Marx, Rawls, and others. The issue we are raising here is how normative and descriptive accounts of social justice and system legitimacy can be reconciled.

To some readers, it may seem like an impossible task to distinguish between "true" and "false" consciousness or, relatedly, between social systems that provide genuine, substantive justice rather than merely the subjective appearance of justice. We readily acknowledge that this is a challenging problem that may require the active collaboration of philosophers, legal scholars, sociologists, psychologists, and others in order to solve it. At the same time, the problem is similar to the one facing students of *rationality*. By juxtaposing descriptive psychological data concerning the ways in which people actually make decisions under ambiguity and uncertainty with normative accounts of rationality derived from philosophical and economic theories, scholarship on judgment and decision making has advanced considerably (e.g., Kahneman, Slovic, & Tversky, 1982; Thaler, 1994). Just as we can ask whether a person is acting rationally, we can ask whether a system is just. In either case, the key issue is to create a normative model against which psychological findings can be compared.

Within legal research, the attempt to reconcile normative and descriptive approaches has spawned the influential and productive "behavioral law and economics" movement (e.g., Sunstein, 2000). In general, such efforts to "naturalize" the study of traditionally philosophical questions by bringing scientific evidence to bear on them have been quite successful.

A similar opportunity exists for behavioral realists who are interested in reforming legal code to make it more just in light of social psychological principles involving stereotyping, bias, and discrimination (e.g., Blasi & Jost, 2006; Kang & Banaji, 2006; Krieger & Fiske, 2006). For example, current discrimination law falls short of protecting members of disadvantaged groups against automatic, unintentional biases that likely harm their chances in a variety of employment and educational contexts. The question is how the law can and should be changed to prevent disparate treatment of minorities and women in light of scientific evidence demonstrating that implicit stereotyping and prejudice can produce discriminatory effects (e.g., Ayres, 2001; Fazio & Olson, 2003; Greenwald & Krieger, 2006; Kang, 2005; Krieger, 1995).

On an even grander scale, it may be possible to bring descriptive sociological and psychological research to bear on normative questions of what constitutes a truly just social system. Probably the most progress has occurred with respect to empirical assessments of John Rawls's (1971) theory of justice as fairness. In the Kantian philosophical tradition, Rawls argued that the most just social system would be the one chosen by rational decision makers under a "veil of ignorance," that is, without knowing anything about their own status or position within the resulting social system. Drawing at least in part on psychological theory and research, Rawls reasoned that under such circumstances people would choose a social system that (1) minimized the degree of inequality in social and economic outcomes, and (2) maximized the social and economic outcomes of those who would occupy the worst position in the new system.

Although there is a limited sense in which empirical research can be said to approximate what Rawls's hypothetical rational decision makers would choose under the "veil of ignorance" as he described it, studies have shown that actual decision makers do tend to favor at least some of the justice principles Rawls advocated (e.g., Bond & Park, 1991; Frohlich & Oppenheimer, 1992; Jost & Ross, 1999; Mitchell, Tetlock, Mellers, & Ordonez, 1993). That is, in evaluating the fairness of various distributional schemes, people who lack information that would enable them to act on self-interest motives tend to behave in a manner that is at least somewhat related to how Rawls theorized they *should* behave. We suggest that similarly useful points of contact exist between normative (i.e., legal and philosophical) and descriptive (sociological and psychological) approaches to issues of *procedural* fairness. The important point is that it is—at least in principle—possible to investigate whether certain normative conceptions of justice and morality fit with empirical data in psychology and other social and behavioral sciences (see also Doris, 2002; Flanagan, 1991; Harman, 1999).

Procedural Fairness as the Basis of Legitimacy

For system legitimacy to be genuine, effective, and long lasting, it needs to be based on more than consensus alone. Aristotle linked the stability of the state to the voluntary acceptance that occurs when resources are actually being fairly distributed across society, that is, to objective distributive justice (Zelditch, 2001).

Authorities also gain a great deal in terms of legitimacy when they follow clear norms of procedural justice, including impartiality, transparency, and respect for human dignity (e.g., Miller, 2001; Tyler, 2001a). Thus, implementing fair procedures as well as outcomes can provide a solid basis for establishing system legitimacy. And, as with distributive justice, procedural justice can be viewed as an objective property of a procedure. For example, in their classic work on procedural justice Thibaut and Walker (1975) compared the procedural justice of inquisitorial and adversary procedural justice using objective criteria. One such criterion was the ability to combat prior bias. They experimentally induced prior bias for or against the defendant and then had cases tried

using one of the two procedures. They reasoned that a just procedure should wash away the effects of prior bias, leading the proportion of guilty verdicts to be the same irrespective of the direction of prior bias. When evaluated against this objective criterion of justice, the adversarial procedure was more just than the inquisitorial system. Such evaluations had nothing to do with how the participants evaluated the justice of the procedures. Rather, they were linked to objective criteria defined by the experimenters.

We want to argue for the particular benefits that come from basing legitimacy on specific mechanisms that communicate and deliver substantive procedural fairness as defined by legal philosophers and scholars who are concerned not only with what people are likely to perceive as just (the descriptive question) but also with what *is* just (the normative question). In other words, we suggest the value of using procedures that both *are* just when evaluated in relationship to objective criteria and *are experienced as being just* by those involved in them.

The legitimacy of authorities is an especially promising basis for the rule of law because research suggests that it is not yoked to agreement with the decisions rendered by legal authorities. If people view as legitimate only those authorities who make decisions with which they agree, it would be difficult for legal authorities to maintain their legitimacy insofar as they are required to make unpopular decisions and to deliver unfavorable outcomes. Fortunately, from the perspective of legal authorities, studies suggest that ascriptions of legitimacy are tied to the perceived fairness of the procedures used by authorities to make decisions rather than to the favorability of outcomes per se (Lind & Tyler, 1988; Thibaut & Walker, 1975; Tyler, 1990; Tyler, Boeckmann, Smith, & Huo, 1997; Tyler & Smith, 1998). These results suggest that legal authorities may be able to maintain their legitimacy in the long term by making decisions in an ethically defensible manner.

People are more likely to regard the police as legitimate if they believe that the police exercise their authority through fair and impartial means (Sunshine & Tyler, 2003b; Tyler, 2001b). Indeed, the evidence suggests that procedural justice judgments are more central to judgments of legitimacy than are such factors as the perceived effectiveness of the police in combating crime. To the extent that people perceive law enforcement officials as legitimate, they are significantly more willing to defer to individual authorities (Tyler & Huo, 2002), and they are also more likely to be in compliance with the law in general (Sunshine & Tyler, 2003b; Tyler, 2006b). These observations hold true when people are reporting on their own personal experiences with legal authorities (Tyler, 2006b; Tyler, Casper, & Fisher, 1989) and also when they are evaluating system-level authorities such as the Supreme Court (Tyler, 1994; Tyler & Mitchell, 1994).

Perhaps most important, from the perspective of the legal system, a number of recent studies link judgments about procedural fairness to the willingness to accept particular legal decisions (Kitzman & Emery, 1993; Lind, Kulik, Ambrose, & de Vera Park, 1993; Wissler, 1995) and to generally follow laws and legal rules (Kim &

Mauborgne, 1993; Sparks, Bottoms, & Hay, 1996; Tyler, 2006b). Procedural justice cues play an especially important rule in securing compliance over time (Paternoster, Brame, Bachman, & Sherman, 1997; Pruitt, Peirce, McGillicuddy, Welton, & Castrianno, 1993). It is by now very clear that people's reactions to law and legal authorities are heavily influenced by their assessments of the fairness of legal procedures.

The procedural basis of legitimacy is especially strong with respect to public opinion concerning political and legal institutions. Studies of the presidency (Tyler, Rasinski, & McGraw, 1985), the legislature (Hibbing & Theiss-Morse, 1995, 2002), and the Supreme Court (Tyler & Mitchell, 1994) all suggest that when citizens are evaluating government institutions, they focus primarily on the fairness of the procedures by which the institutions make policies and implement their decisions. Research on work organizations also suggests that perceived legitimacy has a strong procedural basis (Elsbach, 2001; Tyler & Blader, 2000, 2005).

The procedural basis of legitimacy on an institutional level is consistent with the argument that support for the rules of governance (i.e., procedures and institutions) is theoretically and empirically distinguishable from support for particular individuals or their policies. Studies suggest that reactions to individual leaders and policies are more strongly linked to outcome desirability than are reactions to institutions (Rasinski, Tyler, & Fridkin, 1985). In general, however, it is support for the rules of governmental operation—what Easton (1965) referred to as "diffuse system support" (p. 444)—that is seen as crucial to long-term governmental stability.

Our thesis is that the most reliable way of attaining real as well as ostensible legitimacy and maintaining diffuse system support for legal institutions and authorities is by establishing and protecting procedural safeguards. Indeed, the need for procedural safeguards is one of the strongest arguments for the constitutional separation of executive, representative, and judicial branches of government. To the extent that procedures for ensuring genuine fairness are compromised, the system will begin to lose legitimacy and—over time—fail to inspire the kind of cooperation and deference that is often taken for granted during periods of stability.

The fact that, as system justification theory holds, people tend to perceive the status quo as relatively fair and legitimate means that perceptions of *injustice* should be taken very seriously by anyone who is committed to maintaining genuine system legitimacy. In the case of law, issues of distributive injustice are linked to questions about the appropriate or fair response to violations of the law ("just deserts"). To the extent that a wide rift develops between public opinion and legal doctrines about appropriate punishments in response to wrongdoing (e.g., concerning the severity of drug laws or the desirability of capital punishment), this can be an important sign that the legal system may be lacking in substantive justice and, over time, that it may eventually suffer in terms of perceived legitimacy and capacity for motivating widespread compliance with the law (Robinson & Darley, 1995, 1997).

IMPLICATIONS OF THE VALUE-BASED MODEL

Now that we have laid out the basic tenets of our value-based model, which emphasizes (1) system legitimacy as the basis of social influence and motivation to comply with the law and (2) genuine procedural fairness as the basis of system legitimacy, we are in a position to consider some further legal implications of the model. We focus on three areas of application in particular: reactions to rule breaking, restorative justice, and corporate wrongdoing. In each of these cases, our model makes different recommendations from those that are commonly derived from the deterrence model, which currently dominates scholarly and professional thinking about both criminal justice and business ethics.

Reactions to Rule Breaking

In this section, we return to the central question posed by the deterrence model, namely, what to do with people who break the rules so that neither they nor others will repeat the offense. A frequent justification of incarceration as a common method of deterrence is the notion that once in prison, people who have broken the law can be rehabilitated into citizens who will comply with the law in the future. Unfortunately, there is little evidence that the legal system in its present state effectively changes the future behavior of those who come before the law because of past illegal actions (i.e., to "reform" criminals; see Ellsworth & Mauro, 1998). Much of the fault, no doubt, lies with prison conditions, which are extremely unlikely to create the social and psychological circumstances necessary for learning and self-improvement to occur (e.g., Haney & Zimbardo, 1998).

The work of Darley and colleagues (2000) on rule breaking suggests that the same social values that can help to motivate rule following are also useful for understanding reactions to legal transgressions (see also Carlsmith et al., 2002). They find that people respond to rule breaking by seeking to restore a moral balance between the rule breaker and society. The focus of this work, however, is on how other members of society react to wrongdoing, and not on changes in the rule breakers themselves. If, as we have suggested, legitimacy and moral values are central to rule following, then one guiding concern with regard to rehabilitation should be with how to reconnect the offender to their original moral values.

Consider an example from a recent study of citizen–police encounters (Paternoster et al., 1997). In this study researchers examined what transpired when the police were called to homes to deal with issues of domestic violence. The concern of the study was with subsequent compliance to the law on the part of the abusive men whose behavior led to the initial call. The deterrence model would predict that compliance would be increased by threats and/or punishments meted out by the police. From a social psychological perspective, however, we would argue for the value and effectiveness of more constructive police efforts designed to create and maintain respect for the law on the part of the abuser.

The results of Paternoster and colleagues' (1997) study were consistent with our model of legitimacy and motivation. When the police treated abusers in a procedurally fair manner during their encounter, those abusers were more likely to comply with the law afterward. Fair treatment increased feelings of respect for the law and led abusers to be more willing to obey the law in the future. The influence of procedural fairness was greater than that of threatened or enacted punishments.

This study illustrates one of the core premises of psychological jurisprudence, as we understand it—that legal authorities should be concerned with fostering the shared commitment of citizens to overarching moral principles. To do this, authorities must be aware of citizens' values, their experiences with the legal system, and their judgments about the practices and policies of legal authorities. From this perspective, the key to successfully implementing the rule of law lies in an understanding of procedural fairness as the basis of legitimacy rather than in efforts to more effectively deploy coercive tactics.

Restorative Justice and Reintegrative Shaming

The most comprehensive attempt to reconnect offenders to societal norms and to their own sense of moral values is through *restorative justice* (Braithwaite, 1989; Roberts & Stalans, 2004). During a restorative justice conference the rule-breaking behavior is recognized and punished, but an effort is also made to encourage the rule breaker to recognize that his or her behavior violates social and moral codes that partially constitute the perpetrator's own self-image. As a consequence, his or her criminal behavior should be personally upsetting. Thus, an effort is made to use the incident of rule breaking as a way of encouraging the perpetrator to redouble his commitment to obeying the rules in the future.

To achieve these goals, the restorative justice movement advocates sentences such as formal acknowledgement of wrongdoing, public apologies, and acts of restitution that connect people with the wrongfulness of their actions. From a restorative justice perspective, transgressions and disputes arising from them should be resolved through "reintegrative shaming" techniques (Braithwaite, 2002). Reintegrative shaming combines strong moral disapproval of the offense with respect for the person who committed the offense.

The goal is to bring about reconciliation among victims, offenders, and the community at large. With regard to offenders, the primary goal is to encourage feelings of shame regarding one's crimes, accepting responsibility, and sincerely apologizing for them; this is intended to restore the dignity of offenders. Essential to this process is the social connection that perpetrators feel to their family, friends, and community, so these parties are present at restorative justice hearings, along with the victim and his or her family and friends. All those present are involved in reconnecting the offender to a sense of moral responsibility in relation to the community as a whole. This reconnection, it is hoped, will enhance intrinsic motivation to engage in self-regulatory actions that will work against future transgressions of the law. The restorative

justice argument mirrors the concern in procedural justice research with developing both formal and informal legal procedures that strengthen the influence of social and moral values on people's law-related behavior (Tyler & Darley, 2000).

Research results support the facilitative role of restorative justice conferences (Roberts & Stalans, 2004; Sherman, 1999). Studies suggest that, at least with regard to some types of crime, participating in a restorative justice conference leads to greater cooperation with the law in the future (Nugent, Williams, & Umbreit, 2003; Poulson, 2003). Such conferences, it seems, do increase the motivation to accept the law and the decisions of legal authorities and to be a law-abiding citizen. Work on restorative justice and reintegrative shaming provides further evidence that internal moral values play a pivotal role in motivating compliance with the law (Sunshine & Tyler, 2003b; Tyler, 2006b; Tyler & Huo, 2002). This work also suggests future directions for the rehabilitation of criminals that is based on sound social psychological principles.

Corporate Wrongdoing and Organizational Legitimacy

As with regard to criminal conduct in general, the dominant approach to the problem of corporate crime has been to focus on increasing penalties aimed at deterrence (Tyler, 2005). These efforts have encountered many of the same problems that deterrence efforts have encountered in other domains, including major difficulties with regard to surveillance and the tendency for corporations to accept possible risks associated with being caught for accounting fraud, environmental abuses, and other crimes as costs of doing business (e.g., Moore & Loewenstein, 2004). Another approach to ethical scandals at Enron, Arthur Andersen, WorldCom, Adelphi, Tyco, and many other corporations has been to blame wrongdoing on a "few bad apples," when the reality is that much of the fault lies with the barrel itself (e.g., Bazerman & Banaji, 2004; Tenbrunsel & Messick, 2004).

In this section, we focus on two types of values that have ramifications for the problem of business corruption. The first is the conviction by employees that their organization's rules and authorities are *legitimate*. Legitimacy in this context refers to the view by employees that they are responsible for obeying organizational rules (e.g., that the organization is entitled to have its rules and policies obeyed). Thus, issues of legitimacy apply not only to perceptions of government and law (Tyler, 1999), but also to perceptions of work organizations (Elsbach, 2001; Selznick, 1969; Suchman, 1995; Tyler & Blader, 2005). In work settings, legitimacy refers to the judgment that "the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions" (Suchman, 1995, p. 574). If people feel that their organization has legitimacy, they will be motivated to defer to its rules and policies.

The second value is the conviction by employees that corporate policies are congruent with their own personal *moral values*. Because of relatively widespread corporate wrongdoing, this congruence is often missing. If an employee believes that personal and corporate value congruence exists, then their own moral values motivate them to follow corporate rules because they see those rules as being consistent with—and developed from—a set of moral values that they have internalized. In legal settings, research shows that an important motivation that encourages people to bring their behavior into line with the law is their belief that many behaviors that are illegal are also immoral (Carlsmith et al., 2002; Robinson & Darley, 1995, 1997; Tyler, 2006b). Similar moral values shape cooperation within experimental games (Kerr, 1995; Kerr, Garst, Kiehle, & Harris, 1997; Kerr & Kaufman-Gilliland, 1994). When people feel that their organization acts in ways that are consistent with their own social and moral values, they will be more strongly motivated to support their organization. Serious problems arise when business leaders fail to define illegal activity as morally unacceptable and to foster organizational norms that are consistent with that definition.

Recent calls for greater attention to ethics in business school curricula and for more attention to ethical issues in work cultures follow from the notion that employees' values can be developed and activated within organizational settings (Bowie, 1999; Schminke, 1998; Trevino & Weaver, 2003; Tyler, 2005; Tyler & Blader, 2005). Because internalized moral values strongly affect employee rule following, the challenge is to create organizational climates that foster genuinely ethical goals and then harness the motivational power of employees' values to attain those goals. Fair market ideology and other system-justifying beliefs pose a threat to these efforts to the extent that they minimize ethical scandals, lead to system undercorrection, and rationalize rule-breaking behavior in the name of profit-related goals (Jost, Blount, et al., 2003).

The efficacy of drawing on genuine organizational legitimacy and internalized moral values is suggested by the results of research. While the use of sanctions remains a traditional (but not very effective) management strategy to secure employee compliance with organizational rules and policies, recent studies directly examine whether activating employees' ethical values can be an effective management strategy for securing compliance. The use of a self-regulatory model such as the one we are proposing was advocated long ago in the context of discussing the legal regulation of business (Selznick, 1969), and it has been advanced with increasing frequency in recent years (Darley et al., 2003; Gunningham & Rees, 1997; King & Lenox, 2000; Rechtschaffen, 1998; Suchman, 1995; Tyler, 2001b; Tyler & Darley, 2000).

Research shows that ethical concerns do indeed motivate self-regulatory behavior in organizational settings (Aalders & Wilthagen, 1997). This includes work focused on legitimacy (Human & Provan, 2000; Suchman, 1995; Tyler, 2006b; Tyler & Blader, 2000, 2005; Zimmerman & Zeitz, 2002), morality (Paternoster & Simpson, 1996; Tyler, 2006b; Tyler & Blader, 2000, 2005), and the general role of fairness in shaping social behavior (Lerner, 2003; Rabin, 1993; Tyler & Blader, 2000; Vandenbergh, 2003).

Ethical values shape behavior when people believe that the rules of their organization are legitimate and hence ought to be obeyed, and/or that the values defining the organization are congruent with their own moral values, leading people to feel that they are morally obligated to support the organization.

At the organizational level, studies show that companies are reluctant to use their market power by lowering employee wages during recessions because they believe such an action will be viewed by employees as unfair (Bewley, 1999). In addition, companies often relinquish opportunities to press their market advantages when dealing with their customers due to concerns that they will be perceived as behaving unfairly (Kahneman, Knetsch, & Thaler, 1986). Ethical concerns also shape wage determination (Rees, 1993) and other parameters of employment relationships (Jolls, 2002). These studies suggest that companies are motivated to respond to serious ethical concerns because they believe that judgments of fairness shape people's reactions and behavior (Estreicher, 2002). Indeed, there is evidence suggesting that companies that are regarded as more ethical by employees, customers, and other constituencies also tend to be more profitable (Huselid, 1995; Margolis & Walsh, 2001) and, conversely, that more profitable companies tend to be perceived as more ethical (Jost, Blount, et al., 2003). Because, as we have suggested with regard to system justification and false consciousness, subjective appraisals of legitimacy can be wrong, the key to developing long-term ethical commitments is to maintain genuine rather than apparent organizational legitimacy that is based on normatively defensible standards of fairness.

CONCLUDING REMARKS

Our goal has been to demonstrate several ways in which psychological jurisprudence can inform a set of legal questions, especially those pertaining to the legitimate exercise of legal authority and the motivation of individuals and groups to adhere to the tenets of the legal system. It would be impossible for any review to consider all the relevant issues in the heterogeneous field of law and psychology, and we have not even tried to do so. Nor would it be possible to handle all of these issues within a single conceptual framework, especially given that the concerns of the legal system are what determine the agenda for the application of psychology to the law (e.g., Carson, 2003). We have opted for a more modest approach. We have considered a specific set of theoretical and empirical developments in social psychology that we think are broadly relevant to several areas of the law. These developments pertain to human motivation, social influence, procedural fairness, and system legitimacy and the relations among these constructs.

More specifically, we have suggested that the deterrence model makes a number of assumptions concerning human motivation to comply with the law and that at least some of these assumptions are untenable. A more realistic model takes as its starting point the notion that people are intrinsically motivated to follow internalized moral values that are supportive of a system that is perceived as legitimate. Because perceptions of legitimacy can be inaccurate due to a number of system-justifying biases, a legal system that delivers substantive justice (and not just the appearance of justice) must be based on something other than consensus. We have suggested that one of the most solid foundations for such a system is a genuine commitment to normative standards of procedural fairness. This formulation suggests a number of consequences for reactions to rule breakers, restorative sentencing, and corporate malfeasance. Our hope is that by taking seriously both subjective (i.e., social and psychological) and objective (i.e., normative) factors with regard to justice and the law—as well as the relationship between subjective and objective factors—it is possible to develop a more sophisticated understanding of how citizens in a society like ours can be motivated to follow norms for appropriate conduct and, indeed, why they should.

REFERENCES

Aalders, M., & Wilthagen, T. (1997). Moving beyond command and control: Reflexivity in the regulation of occupational safety and health and the environment. *Law and Policy*, *19*, 415–443.

Ayres, I. (2001). *Pervasive prejudice?: Unconventional evidence of race and gender discrimination*. Chicago: University of Chicago Press.

Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. Oxford, UK: Oxford University Press.

Bazerman, M. (1990). *Judgment in managerial decision making* (2nd ed.). New York: Wiley.

Bazerman, M., & Banaji, M. R. (2004). The social psychology of ordinary ethical failures. *Social Justice Research*, *17*, 111–115.

Bewley, T. (1999). *Why wages don't fall during a recession*. Cambridge, MA: Harvard University Press.

Blasi, G., & Jost, J. T. (2006). System justification theory and research: Implications for law, legal advocacy, and social justice. *California Law Review*, *94*, 1119–1168.

Blumstein, A., Cohen, J., & Nagin, D. (1978). *Deterrence and incapacitation*. Washington, DC: National Academy of Sciences.

Boll, T., Ferring, D., & Filipp, S. H. (2005). Effects of parental differential treatment on relationship quality with siblings and parents: Justice evaluations as mediators. *Social Justice Research*, *18*, 155–182.

Bond, D., & Park, J.-C. (1991). An empirical test of Rawls's theory of justice: A second approach, in Korea and the United States. *Simulation and Gaming*, *22*, 443–462.

Bowie, N. E. (1999). *Business ethics: A Kantian perspective*. Malden, MA: Blackwell.

Braithwaite, J. (1989). *Crime, shame, and reintegration*. Cambridge, UK: Cambridge University Press.

Braithwaite, J. (2002). *Restorative justice and responsive regulation*. Oxford, UK: Oxford University Press.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? *Journal of Personality and Social Psychology*, *83*, 284–299.

Carson, D. (2003). Psychology and law: A subdiscipline, an interdisciplinary collaboration or a project? In D. Carson & R. Bull (Eds.), *Handbook of psychology in legal contexts* (pp. 1–27). Chichester, UK: Wiley.

Citrin, J., & Muste, C. (1999). Trust in government. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of political attitudes* (pp. 465–532). San Diego, CA: Academic Press.

Cohn, E. S., & White, S. O. (1990). *Legal socialization: A study of norms and rules*. New York: Springer-Verlag.

Crosby, F. J. (1982). *Relative deprivation and working women.* New York: Oxford University Press.

Dahl, R. (1956). *A preface to democratic theory.* Chicago: University of Chicago Press.

Darley, J. M. (2004). Social production for the organization of evil. In J. T. Jost & J. Sidanius (Eds.), *Political psychology: Key readings* (pp. 383–410). New York: Psychology Press/Taylor & Francis. (Reprinted from *Psychological Inquiry*, *3*, 199–218)

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, *24*, 659–683.

Darley, J. M., Fulero, S., Haney, C., & Tyler, T. R. (2002). *Psychological jurisprudence*. New York: Plenum Press.

Darley, J. M., Tyler, T. R., & Bilz, K. (2003). Enacting justice: The interplay of individual and institutional perspectives. In M. A. Hogg & J. Cooper (Eds.), *The Sage handbook of social psychology* (pp. 458–476). London: Sage.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.

Dennett, D., & Haugeland, J. C. (1987). Intentionality. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 383–386). New York: Oxford University Press.

Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. New York: Cambridge University Press.

Dunning, D. (1999). A newer look: Motivated social cognition and the schematic representation of social concepts. *Psychological Inquiry*, *10*, 1–11.

Easton, D. (1965). *A systems analysis of political life*. Chicago: University of Chicago Press.

Easton, D. (1975). A reassessment of the concept of political support. *British Journal of Political Science*, *5*, 435–457.

Ellsworth, P. C., & Gross, S. R. (1994). Hardening of the attitudes: Americans' views on the death penalty. *Journal of Social Issues*, *50*, 19–52.

Ellsworth, P. C., & Mauro, R. (1998). Psychology and law. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 684–732). New York: McGraw-Hill.

Elsbach, K. S. (2001). The architecture of legitimacy. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 391–415). Cambridge, UK: Cambridge University Press.

Estreicher, S. (2002). Human behavior and the economic paradigm at work. *New York University Law Review*, *77*, 1–5.

Ewick, P., & Silbey, S. S. (1988). *The common place of law*. Chicago: University of Chicago Press.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Finkel, N. J. (1995). *Commonsense justice: Jurors' notions of the law*. Cambridge, MA: Harvard University Press.

Flanagan, O. (1991). *Varieties of moral personality: Ethics and psychological realism*. Cambridge, MA: Harvard University Press.

Flanagan, T. J., & Longmire, D. R. (1996). *Americans' view of crime and justice.* Thousand Oaks, CA: Sage.

Fox, D. R. (1999). Psycholegal scholarship's contribution to false consciousness about injustice. *Law and Human Behavior*, *23*, 9–30.

French, J. R. P., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Ann Arbor: University of Michigan Press.

Frey, B. S. (1994). How intrinsic motivation is crowded in and out. *Rationality and Society*, *6*, 334–352.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives. *American Economic Review*, *87*, 746–755.

Frohlich, N., & Oppenheimer, J. A. (1992). *Choosing justice: An experimental approach to ethical theory.* Berkeley: University of California Press.

Fuller, L. (1969). Human interaction and the law. *American Jurisprudence*, *14*, 1–36.

Garland, D. (2001). *The culture of control*. Chicago: University of Chicago Press.

Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, *29*, 1–17.

Greenstein, F. (1965). *Children and politics.* New Haven, CT: Yale University Press.

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, *94*, 945–967.

Greenwood, J. (1991). Introduction: Folk psychology and scientific psychology. In J. Greenwood (Ed.), *The future of folk psychology* (pp. 1–21). Cambridge, UK: Cambridge University Press.

Gunningham, N., & Rees, J. (1997). Industry self-regulation. *Law and Policy*, *19*, 363–414.

Gurr, T. R. (1970). *Why men rebel*. Princeton, NJ: Princeton University Press.

Hamilton, V. L., & Sanders, J. (1992). *Everyday justice: Responsibility and the individual in Japan and the United States*. New Haven, CT: Yale University Press.

Haney, C. (1982). Criminal justice and the nineteenth-century paradigm. *Law and Human Behavior*, *6*, 191–235.

Haney, C. (1991). The fourteenth amendment and symbolic legality: Let them eat due process. *Law and Human Behavior*, *15*, 183–204.

Haney, C. (1993). Psychology and legal change: The impact of a decade. *Law and Human Behavior*, *17*, 371–398.

Haney, C., & Zimbardo, P. (1998). The past and future of U.S. prison policy: Twenty-five years after the Stanford prison experiment. *American Psychologist*, *53*, 709–727.

Hanson, J., & Yosifon, D. (2004). The situational character: A critical realist perspective on the human animal. *Georgetown Law Journal*, *93*, 1–179.

Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, *99*, 315–331.

Hastie, R., Penrod, S. D., & Pennington, N. (1983). *Inside the jury*. Cambridge, MA: Harvard University Press.

Heuer, L. & Penrod, S. D. (1989). Instructing jurors: A field experiment with written and preliminary instructions. *Law and Human Behavior*, *13*, 409–430.

Hess, R. D., & Torney, J. (1967). *The development of political attitudes in children*. Chicago: Aldine Press.

Hibbing, J. R., & Theiss-Morse, E. (1995). *Congress as public enemy: Public attitudes toward American political institutions*. Cambridge, UK: Cambridge University Press.

Hibbing, J. R., & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Cambridge, UK: Cambridge University Press.

Hochschild, J. L. (1981). *What's fair*. Cambridge, MA: Harvard University Press.

Hoffman, M. (1977). Moral internalization: Current theory and research. *Advances in Experimental Social Psychology*, *10*, 85–133.

Human, S. E., & Provan, K. G. (2000). Legitimacy building in the evolution of small firm multilateral networks: A comparative study of success and demise. *Administrative Science Quarterly*, *45*, 327–365.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, *38*, 635–672.

Hyman, H. (1959). *Political socialization*. New York: Free Press.

Jolls, C. (2002). Fairness, minimum wage law, and employee benefits. *New York University Law Review*, *77*, 47–70.

Jolls, C., & Sunstein, C. R. (2006). The law of implicit bias. *California Law Review*, *94*, 969–996.

Jost, J. T. (1995). Negative illusions: Conceptual clarification and psychological evidence concerning false consciousness. *Political Psychology*, *16*, 397–424.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system justification and the production of false consciousness. *British Journal of Social Psychology*, *33*, 1–27.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*, 881–919.

Jost, J. T., Blount, S., Pfeffer, J., & Hunyady, G. (2003). Fair market ideology: Its cognitive-motivational underpinnings. *Research in Organizational Behavior*, *25*, 53–91.

Jost, J. T., & Hunyady, O. (2002). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, *13*, 111–153.

Jost, J. T., & Major, B. (2001). Emerging perspectives on the psychology of legitimacy. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 3–32). Cambridge, UK: Cambridge University Press.

Jost, J. T., Pelham, B. W., Sheldon, O., & Sullivan, B. N. (2003). Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged. *European Journal of Social Psychology*, *33*, 13–36.

Jost, J. T., & Ross, L. (1999). Fairness norms and the potential for mutual agreements involving majority and minority groups. In E. Mannix, M. Neale, & R. Wageman (Eds.), *Research on managing groups and teams: Vol. 2. Context* (pp. 93–114). Greenwich, CT: JAI Press.

Kahan, D. (1999). The secret ambition of deterrence. *Harvard Law Review*, *113*, 413–500.

Kahneman, D., Knetsch, J., & Thaler, R. (1986). Fairness and the assumptions of economics. *Journal of Business*, *59*, 5285–5300.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kang, J. (2005). Trojan horses of race. *Harvard Law Review*, *118*, 1489–1537.

Kang, J., & Banaji, M. R. (2006). Fair measures: A behavioral realist revision of "affirmative action." *California Law Review*, *94*, 1063–1118.

Kelman, H. C. (1969). Patterns of personal involvement in the national system: A social-psychological analysis of political legitimacy. In J. Rosenau (Ed.), *International politics and foreign policy* (pp. 276–288). New York: Free Press.

Kelman, H. C., & Hamilton, V. L. (1989). *Crimes of obedience*. New Haven, CT: Yale University Press.

Kerr, N. L. (1995). Norms in social dilemmas. In D. Schroeder (Ed.), *Social dilemmas: Perspectives on individuals and groups* (pp. 31–47). Westport, CT: Praeger.

Kerr, N. L., Garst, J., Kiehle, D., & Harris, S. (1997). That still, small voice: Commitment to cooperate as an internalized vs. a social norm. *Personality and Social Psychology Bulletin*, *23*, 1300–1311.

Kerr, N. L., & Kaufman-Gilliland, C. M. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology*, *66*, 513–529.

Kim, W. C., & Mauborgne, R. A. (1993). Procedural justice, attitudes, and subsidiary top management compliance with multinationals' corporate strategic decisions. *Academy of Management Journal*, *36*, 502–526.

King, A., & Lenox, M. (2000). Industry self-regulation without sanctions. *Academy of Management Journal*, *43*, 698–716.

Kitzmann, K. M., & Emery, R. E. (1993). Procedural justice and parents' satisfaction in a field study of child custody dispute resolution. *Law and Human Behavior*, *17*, 553–567.

Kluegel, J. R., & Smith, E. R. (1986). *Beliefs about inequality*. New York: Aldine de Gruyter.

Köhnken, G. (1996). Social psychology and the law. In G. R. Semin & K. Fiedler (Eds.), *Applied social psychology* (pp. 257–281). London: Sage.

Kramer, R. M., & Jost, J. T. (2002). Close encounters of the suspicious kind: Outgroup paranoia in hierarchical trust dilemmas. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 173–189). New York: Psychology Press/Taylor & Francis.

Kramer, R. M., & Tyler, T. R. (Eds.). (1996). *Trust in organizations*. Thousand Oaks: CA: Sage.

Krieger, L. H. (1995). The content of our categories: A cognitive bias approach to discrimination and equal employment opportunity. *Stanford Law Review*, *47*, 1161–1248.

Krieger, L. H., & Fiske, S. T. (2006). Behavioral realism in employment discrimination law: Implicit bias and disparate treatment. *California Law Review*, *94*, 997–1062.

Krislov, S., Boyum, K. O., Clark, J. N., Shaefer, R. C., & White, S. O. (1966). *Compliance and the law: A multi-disciplinary approach*. Beverly Hills, CA: Sage.

Lerner, M. J. (1980). *The belief in a just world*. New York: Plenum Press.

Lerner, M. J. (2003). The justice motive: Where social psychologists found it, how they lost it, and why they may not find it again. *Personality and Social Psychology Review*, *7*, 388–399.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Lind, E. A., Kulik, C. T., Ambrose, M., & de Vera Park, M. (1993). Individual and corporate dispute resolution. *Administrative Science Quarterly*, *38*, 224–251.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum Press.

Lipset, S. M. (1959). Some social requisites of democracy. *American Political Science Review*, *53*, 69–105.

Loewenstein, G. F., Prelec, D., & Shatto, C. (1998). *Hot/cold empathy intrapersonal gaps and the prediction of curiosity*. Unpublished manuscript, Carnegie Mellon University.

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*, 426–441.

Loftus, E. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

MacCoun, R. J. (1993). Drugs and the law: A psychological analysis of drug prohibition. *Psychological Bulletin*, *113*, 497–512.

Major, B. (1994). From inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group memberships. *Advances in Experimental Social Psychology*, *26*, 293–355.

Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.

Margolis, J. D., & Walsh, J. P. (2001). *People and profits? The search for a link between a company's social and financial performance*. Mahwah, NJ: Erlbaum.

Mastrofski, S. D., Snipes, J. B., & Supina, A. E. (1996). Compliance on demand: The public's responses to specific police requests. *Journal of Crime and Delinquency*, *33*, 269–305.

McCluskey, J. D., Mastrofski, S. D., & Parks, R. B. (1998). *To acquiesce and rebel: Predicting citizen compliance with police requests*. Unpublished manuscript.

Meares, T. L. (2000). Norms, legitimacy, and law enforcement. *Oregon Law Review*, *79*, 391–415.

Melton, G. B. (Ed.). (1985). *Nebraska Symposium on Motivation* (Vol. 33). Lincoln: University of Nebraska Press.

Merelman, R. J. (1966). Learning and legitimacy. *American Political Science Review*, *60*, 548–561.

Merry, S. (1990). *Getting justice and getting even: Legal consciousness among working-class Americans*. Chicago: University of Chicago Press.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*, 340–363.

Milgram, S. (1974). *Obedience to authority*. New York: Harper & Row.

Miller, D. T. (2001). Disrespect and the experience of injustice. *Annual Review of Psychology*, *52*, 527–553.

Mitchell, G., Tetlock, P. E., Mellers, B. A., & Ordonez, L. D. (1993). Judgments of social justice: Compromises between equality and efficiency. *Journal of Personality and Social Psychology*, *65*, 629–639.

Monahan, J., & Loftus, E. F. (1982). The psychology of law. *Annual Review of Psychology*, *33*, 441–475.

Moore, B., Jr. (1978). *Injustice: The social bases of obedience and revolt*. White Plains, NY: Sharpe.

Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research*, *17*, 189–202.

Nagin, D. S. (1998). Criminal deterrence research at the outset of the twenty-first century. In M. Tonry (Ed.), *Crime and justice* (Vol. 23, pp. 1–42). Chicago: University of Chicago Press.

Nagin, D. S., & Paternoster, R. (1991). The preventive effects of the perceived risk of arrest. *Criminology*, *29*, 561–585.

Niemi, R. G. (1973). Political socialization. In J. N. Knutson (Ed.), *Political psychology* (pp. 117–138). San Francisco: Jossey-Bass.

Nugent, W. R., Williams, M., & Umbreit, M. S. (2003). Participation in victim-offender mediation and the prevalence and severity of subsequent delinquent behavior. *Utah Law Review*, pp. 137–166.

Parkin, F. (1971). *Class inequality and political order*. New York: Praeger.

Parsons, T. (1967). Some reflections on the place of force in social process. In T. Parsons (Ed.), *Sociological theory and modern society* (pp. 264–296). New York: Free Press.

Paternoster, R. (1987). The deterrent effect of the perceived certainty and severity of punishment. *Justice Quarterly*, *4*, 173–217.

Paternoster, R. (1989). Decisions to participate in and desist from four types of common delinquency. *Law and Society Review*, *23*, 7–40.

Paternoster, R., Brame, R., Bachman, R., & Sherman, L. W. (1997). Do fair procedures matter? The effect of procedural justice on spouse assault. *Law and Society Review*, *31*, 163–204.

Paternoster, R., & Iovanni, L. (1986). The deterrent effect of perceived severity. *Social Forces*, *64*, 751–777.

Paternoster, R., Saltzman, L. E., Waldo, G. P., & Chiricos, T. G. (1983). Perceived risk and social control: Do sanctions really deter? *Law and Society Review*, *17*, 457–479.

Paternoster, R., & Simpson, S. (1996). Sanction threat and appeals to morality. *Law and Society Review*, *30*, 549–583.

Poulson, B. (2003). A third voice: A review of empirical research on the psychological outcomes of restorative justice. *Utah Law Review*, pp. 167–204.

Pruitt, D. G., Peirce, R. S., McGillicuddy, N. B., Welton, G. L., & Castrianno, L. M. (1993). Long-term success in mediation. *Law and Human Behavior*, *17*, 313–330.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*, 1281–1302.

Rasinski, K., Tyler, T. R., & Fridkin, K. (1985). Exploring the function of legitimacy: Mediating effects of personal and institutional legitimacy on leadership endorsement and system support. *Journal of Personality and Social Psychology*, *49*, 386–394.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rechtschaffen, C. (1998). Deterrence vs. cooperation and the evolving theory of environmental enforcement. *Southern California Law Review*, *71*, 1181–1272.

Rees, A. (1993). The role of fairness in wage determination. *Journal of Labor Economics*, *11*, 243–253.

Reich, W. (1989). What is class consciousness? In R. S. Gottlieb (Ed.), *An anthology of Western Marxism* (pp. 146–167). New York: Oxford University Press. (Original work published 1934)

Reiss, A. J. (1971). *The police and the public*. New Haven, CT: Yale University Press.

Roberts, J. V., & Stalans, L. J. (2004). Restorative sentencing: Exploring the views of the public. *Social Justice Research*, *17*, 315–334.

Robinson, P. H., & Darley, J. (1995). *Justice, liability, and blame*. Boulder, CO: Westview.

Robinson, P. H., & Darley, J. (1997). The utility of desert. *Northwestern University Law Review*, *91*, 453–499.

Ross, H. L. (1982). *Deterring the drinking driver*. Lexington, MA: Lexington Press.

Scheingold, S. A. (1974). *The politics of rights*. New Haven, CT: Yale University Press.

Schminke, M. (1998). *Managerial ethics: Moral management of people and processes*. Mahwah, NJ: Erlbaum.

Sears, D. O. (2003). The psychology of legitimacy. *Political Psychology*, *25*, 318–323.

Selznick, P. (1969). *Law, society, and industrial justice*. New York: Russell Sage.

Sherman, L. (1993). Defiance, deterrence, irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency*, *30*, 445–473.

Sherman, L. (1998, November). *Alternative prevention strategies and the role of policing*. Paper presented at the symposium "Beyond Incarceration: The Economics of Crime." Cambridge, MA: Harvard University.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Sparks, R., Bottoms, A., & Hay, W. (1996). *Prisons and the problem or order*. Oxford, UK: Clarendon.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, *20*, 571–610.

Sunshine, J., & Tyler, T. R. (2003a). Moral solidarity, identification with the community, and the importance of procedural justice. *Social Psychology Quarterly*, *66*, 153–165.

Sunshine, J., & Tyler, T. R. (2003b). The role of procedural justice and legitimacy in shaping public support for policing. *Law and Society Review*, *37*, 513–548.

Sunstein, C. R. (Ed.). (2000). *Behavioral law and economics*. Cambridge, UK: Cambridge University Press.

Tapp, J., & Levine, F. (1977). *Law, justice and the individual in society: Psychological and legal issues*. New York: Holt, Rinehart and Winston.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Teevn, J. J. (1975). Subjective perception of deterrence. *Journal of Research in Crime and Delinquency*, *13*, 155–164.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, *44*, 684–707.

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research*, *17*, 223–236.

Thaler, R. H. (1994). *Quasi-rational economics*. New York: Russell Sage.

Thibaut, J., & Walker, L. (1975). *Procedural justice*. Hillsdale, NJ: Erlbaum.

Travis, M. A. (2005). Recapturing the transformative potential of employment discrimination law. *Washington and Lee Law Review*, *62*, 3–92.

Trevino, L. K., & Weaver, G. R. (2003). *Managing ethics in business organizations*. Palo Alto, CA: Stanford University Press.

Tyler, T. R. (1994). Psychological models of the justice motive. *Journal of Personality and Social Psychology*, *67*, 850–863.

Tyler, T. R. (1997a). Citizen discontent with legal procedures. *American Journal of Comparative Law*, *45*, 869–902.

Tyler, T. R. (1997b). Compliance with intellectual property laws: A psychological perspective. *Journal of International Law and Politics*, *28*, 101–115.

Tyler, T. R. (1997c). Procedural fairness and compliance with the law. *Swiss Journal of Economics and Statistics*, *133*, 219–240.

Tyler, T. R. (1998). Public mistrust of the law: A political perspective. *University of Cincinnati Law Review*, *66*, 847–876.

Tyler, T. R. (1999). Why do people help organizations?: Social identity and pro-organizational behavior. In B. Staw & R. Sutton (Eds.), *Research on organizational behavior* (Vol. 21, pp. 201–246) Greenwich, CT: JAI Press.

Tyler, T. R. (2001a). A psychological perspective on the legitimacy of institutions and authorities. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 416–436). Cambridge, UK: Cambridge University Press.

Tyler, T. R. (2001b). Trust and law-abidingness: A proactive model of social regulation. *Boston University Law Review*, *81*, 361–406.

Tyler, T. R. (2004). Affirmative action in an institutional context: The antecedents of policy preferences and political support. *Social Justice Research*, *17*, 5–24.

Tyler, T. R. (2005). Promoting employee policy adherence and rule following in work settings: The value of self-regulatory approaches. *Brooklyn Law Review, 70,* 1287–1312.

Tyler, T. R. (2006a). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology, 57,* 375–400.

Tyler, T. R. (2006b). *Why people obey the law*. Princeton, NJ: Princeton University Press.

Tyler, T. R., & Blader, S. L. (2000). *Cooperation in groups: Procedural justice, social identity, and behavioral engagement*. Philadelphia: Psychology Press.

Tyler, T. R., & Blader, S. L. (2005). Can businesses effectively regulate employee conduct? The antecedents of rule following in work settings. *Academy of Management Journal, 48,* 1143–1158.

Tyler, T. R., Boeckmann, R. J., Smith, H. J., & Huo, Y. J. (1997). *Social justice in a diverse society*. Boulder, CO: Westview.

Tyler, T. R., Casper, J., & Fisher, B. (1989). Maintaining allegiance toward political authorities. *American Journal of Political Science, 33,* 629–652.

Tyler, T. R., & Darley, J. M. (2000). Building a law-abiding society: Taking public views about morality and the legitimacy of legal authorities into account when formulating substantive law. *Hofstra Law Review, 28,* 707–739.

Tyler, T. R., & Huo, Y. J. (2002). *Trust in the law*. New York: Russell Sage Foundation.

Tyler, T. R., & McGraw, K. M. (1986). Ideology and the interpretation of personal experience: Procedural justice and political quiescence. *Journal of Social Issues, 42,* 115–128.

Tyler, T. R., & Mitchell, G. (1994). Legitimacy and the empowerment of discretionary legal authority: The United States Supreme Court and abortion rights. *Duke Law Journal, 43,* 703–814.

Tyler, T. R., Rasinski, K., & McGraw, K. (1985). The influence of perceived injustice on support for political authorities. *Journal of Applied Social Psychology, 15,* 700–725.

Tyler, T. R., & Smith, H. J. (1998). Social justice and social movements. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 595–629). New York: McGraw-Hill.

U.S. Department of Justice. (2001). *Prisoners in 2000* (Bureau of Justice Statistics Bulletin NCJ 188207). Washington, DC: Department of Justice. Office of Justice Programs.

Vandenbergh, M. P. (2003). Beyond elegance: A test of social norms in corporate environmental compliance. *Stanford Environmental Law Journal, 22,* 55–143.

Wakslak, C., Jost, J. T., Tyler, T. R., & Chen, E. (in press). Moral outrage mediates the dampening effect of system justification on support for redistributive social policies. *Psychological Science*.

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Boston: Allyn & Bacon.

Weber, M. (1968). *Economy and society*. (G. Roth & C. Wittich, Eds.). Berkeley: University of California Press.

Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wells, G. L., Leippe, M. R., & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a line up. *Law and Human Behavior, 3,* 285–293.

Wissler, R. L. (1995). Mediation and adjudication in the small claims court. *Law and Society Review, 29,* 323–358.

Zelditch, M., Jr. (2001). Theories of legitimacy. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 33–53). Cambridge, UK: Cambridge University Press.

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. *Academy of Management Review, 27,* 414–43.

Zucker, L. G. (1977). The role of institutionalization in cultural persistence. *American Sociological Review, 42,* 726–743.

The Reciprocal Relation
between Principles and Practice
Social Psychology and Health Behavior

ALEXANDER J. ROTHMAN
PETER SALOVEY

From 1990 to 2003, the percentage of obese individuals in the United States has increased dramatically. In 1991, four states reported that between 15% and 19% of their population was obese and not a single state reported rates of obesity equal to or greater than 20%. By 2001, these numbers had increased to 20 and 29 states, respectively (Mokdad et al., 2001, 2003). These data offer a stark illustration of a daunting practical problem. Although the observed increase in rates of obesity reflect the interplay of myriad factors, fundamentally they are rooted in the consequences of people's behavior—the amount and type of food people eat and the amount and type of physical activity in which people are engaged. Efforts to reduce rates of obesity thus depend on strategies that can effectively elicit changes in diet and physical activity. These initiatives can take the form of structural initiatives that modify the physical environment to promote physical activity, improve access to healthier food, or reduce the cost of healthy foods in schools or other settings (Brownell & Horgen, 2004; French et al., 2001; Stokols, Grzywacz, McMahan, & Phillips, 2003) as well as behavioral initiatives that target individuals directly and provide them with a diet or exercise plan (Wadden, Vogt, Andersen, & Bartlett, 1997; Wing, 2002). Regardless of the focus of the initiative, its design is predicated—either implicitly or explicitly—on assumptions regarding the processes that regulate people's behavioral decisions. Thus, the design and implementation of initiatives to address obesity critically depend on the state of scientific knowledge regarding the basic principles that shape how people process and respond to information about their behavior.

The premise that theoretical principles afford invaluable guidance in the design and implementation of interventions is not unique to efforts to slow rising rates of obesity. The opportunities afforded by theoretical innovations underlie Miller's (1970) call to "give psychology away" and Kurt Lewin's (1951) often quoted assertion "that there is nothing so practical as a good theory" (p. 169). Although we share Lewin and others' confidence in the value of theory, we must not forget that such value is predicated on the existence of *good* theories (Rothman, 2004).

What makes for a good theory? Theory development is a dynamic process that involves the specification of relations between constructs, the testing of those relations, and then, in light of the obtained evidence, the respecification, if needed, of the initial proposition (Noar & Zimmerman, 2005; Weinstein & Rothman, 2005). The efforts of basic behavioral scientists typically focus on elucidating whether a predicted relation between two constructs *can* be observed and thus theory testing is typically conducted in a controlled setting that affords the social and behavioral version of a Petrie dish (Mook, 1983). However, from the perspective of an applied behavioral scientist, good theories must articulate not only what *can* happen but also what *does* and *does not* happen when basic principles operate in more complex social environments. Theory development depends on answers to questions regarding both what can and what does happen, and initiatives that rigorously test the ability of basic principles to specify the determinants of practical problems provide invaluable opportunities to pursue answers to these questions. The process of applying

basic principles sheds light on both what is known and well specified and what is not known or poorly specified. Yet, too often theories are treated—both by theorists and by practitioners—as fixed entities that need only to be disseminated. What is needed instead is a system comprised of an ongoing exchange between basic and applied behavioral scientists in which theory is treated as a dynamic entity whose value depends on it being not only applied and tested rigorously but also refined based on the findings afforded by those tests.

In this chapter, we explore how the interplay between basic and applied behavioral scientists has advanced not only our understanding of health behavior but also our ability to develop initiatives that can promote healthy behavioral practices effectively. Specifically, we examine how social psychological principles have informed the understanding of how people think and reason about their health and how efforts to analyze people's health practices have, in turn, enriched our understanding of basic social psychological principles. It is our hope that this perspective serves to illustrate what can be gained when we challenge our theoretical principles and seek to discern their scope, power, and applicability and to shed light on those areas in which investigators have successfully pursued the interplay between basic and applied behavioral science and those areas where work of this type remains sorely needed.

Our organization of the material in this chapter is grounded in the observation that the process by which people manage and change their behavior can be differentiated into a series of different phases, each of which involves a critical set of psychological and behavioral processes (Prochaska, DiClemente, & Norcross, 1992; Rothman, Baldwin, & Hertel, 2004; Weinstein, 1988). Specifically, we have chosen to focus on three phases. First, we examine how people *gather and interpret* health information and determine whether there is a health problem. Second, we examine how people *decide* whether to take action and the processes that regulate the implementation of their decision. Third, we examine the processes that underlie people's efforts to *initiate and maintain* their behavioral decisions. We believe that differentiating between these phases illustrates a fundamental social psychological perspective in that it reveals how the decision-making processes people rely on are sensitive to the situations in which they find themselves. Finally, in light of the present state of the science, we consider what investigators can do to capitalize on the opportunities emanating from collaborations between basic and applied behavioral scientists.

HEALTH INFORMATION PROCESSING: RECOGNIZING A HEALTH PROBLEM

Given the significant costs associated with unhealthy behaviors, one might expect that people would readily take advantage of opportunities to modify them. Yet, it has become increasingly clear that most people are neither ready nor willing to change their health behavior (Prochaska et al., 1992) and that a critical aspect of en-

couraging people to modify their behavior is to first draw their attention to a given health issue and then, when appropriate, help them to recognize its personal relevance. It is this phase of the behavior change process that we examine in this section.

Weinstein (1988; see also Prochaska et al., 1992) has proposed that there is value in recognizing the informational and psychological needs of people who have yet to consider whether to take action. Specifically, some individuals may not be aware of the health issue (e.g., "I've never heard of the avian flu"), whereas others may be aware of the issue but not see it as personally relevant ("I've heard of the avian flu, but it is not a problem where I live"). Weinstein has proposed that this distinction is best represented by two discrete stages and thus suggests that there are a discrete set of factors that facilitate the transition out of each stage (see Weinstein, Rothman, & Sutton, 1998, for a discussion of what defines a stage). Although there is not yet sufficient evidence to accept this premise, we feel it is worth considering the distinction between efforts to help people become aware of a health issue and efforts to help people recognize the personal relevance of a health issue.

Are You Aware of the Health Problem?

What is known about getting people to recognize the existence of a health problem? Although research has shown consistently that information alone is not enough to promote changes in behavior, awareness of a health issue is a critical first step in the behavior change process. Basic health information provides people with a cognitive framework or schema that helps them to monitor for and integrate new information about the issue (Cameron & Leventhal, 2003). In thinking about educational efforts, it may be important to distinguish between providing information about a new, emerging health issue (e.g., avian flu) and a health issue that has been around for a while but for which what is known about it has changed. People have been shown to be attentive to novel, particularly negative, information (Taylor, 1991), suggesting that information about a new threat will stand out. However, in the absence of any continued pressure to think through the issue, people may have little motivation to elaborate on and gather more information about it and thus will take advantage of available opportunities to minimize any potential danger afforded by this issue. One consequence of this may be that people find that they have heard of a health issue, but when pressed find they know little or nothing about it.

Yet, it may be even more challenging to get people to attend to information about a health problem of which they are already aware. A general feeling of familiarity with the issue may be taken as a sign that they have already considered the issue and thus have little need to look for more, new information. For example, in a recent survey 94% of smokers answered "yes" to the question, "Do you consider yourself adequately informed about the health risks of smoking?" (Cumming et al., 2001). It would be interesting to examine how beliefs such as these affect decisions about information seeking.

People may also be less responsive to information when they perceive a continual shift in what is known about an issue. The perception that there is uncertainty regarding what is understood about a health issue may undermine the value people attribute to new information and, in some cases, may induce a sense of overload. For example, the repeated onslaught of information regarding the impact of diet and exercise on cancer may induce the belief that everything causes cancer and, in turn, may lead people to pay less attention to cancer information. Under these circumstances, people may actively resist attending to new information. Although researchers in the area of health communication continue to examine where and how people gather health information (e.g., Dutta-Bergman, 2004), there is a need for further work that specifies the processes that regulate people's willingness to seek out and evaluate new health information.

Do I Need to Be Concerned about This?

Perceptions of personal risk are thought to play a critical role in shaping the choices people make about their health (Salovey, Rothman, & Rodin, 1998; Weinstein, 1993). Although people are exposed to health risk information repeatedly—for example, that a specific pattern of behavior (e.g., smoking) will lead to a specific health problem (e.g., lung cancer)—this information does not always have a systematic effect on perceptions of personal risk. People appear to be consistently more willing to recognize the existence of a health problem than to recognize their own personal vulnerability (Weinstein, 1998) and to report that their own risk is below that of the average person (Klein & Weinstein, 1997; see Chambers & Windschitl, 2004, for a discussion of the processes that may underlie this effect).

The observation that people may not always recognize the risk posed by their behavior is thought to reflect the fact that people are not passive, unbiased processors of information about their health status (Kunda, 1990; Taylor & Brown, 1988). They respond to health information strategically and are generally receptive to favorable information about their health but often engage in strategies that minimize or discount unfavorable health information (e.g., Ditto & Lopez, 1992; Liberman & Chaiken, 1992; see Rothman & Kiviniemi, 1999, for a review).

Investigators have utilized a range of paradigms to illustrate the ways in which people can resist the unfavorable implications of health information. Laboratory studies have shown that when people review information that may indicate they are at risk for a health problem (e.g., when a heavy coffee drinker reads an article about the potential health risks associated with caffeine intake), they actively counterargue the information provided and are quick to recognize any potential problems with it (Kunda, 1987; Liberman & Chaiken, 1992). The same information is given limited scrutiny when it poses little risk (e.g., when read by someone who does not consume much caffeine). In a similar manner, people emphasize differences between themselves and others who have developed a health problem (e.g., HIV; Gump & Kulik,

1995). The more readily people are able to differentiate themselves from someone with a health problem, the less at risk they consider themselves to be. In a similar manner, the belief that a health problem is consistently associated with a particular social group to which one does not belong can facilitate the inference that it "cannot happen to me." The strategic reliance on social comparison information, in particular the reliance on comparisons to people worse off than themselves, may similarly offer people a way to manage the potential threat posed by a health problem (Taylor & Lobel, 1989).

Potentially threatening health risk information can also lead people to seek out additional information that might undermine its implications (Ditto & Lopez, 1992). For example, in an innovative paradigm that has provided investigators the opportunity to study people's reactions to a diagnostic test, people are asked to screen themselves for an enzyme deficiency that could lead to a mild but irritating pancreatic disorder (Croyle & Ditto, 1990). When people discover that they have tested negative for the enzyme deficiency, they are quick to accept the results. However, when people learn they have tested positive for the deficiency, they spend significantly more time scrutinizing the test result, question the accuracy of the test, and seek out opportunities to retest themselves (Ditto & Lopez, 1992; Ditto, Munro, Apanovitch, Scepansky, & Lockhart, 2003). People have also been shown to alter their beliefs about the health issue selectively in order to minimize its negative implications for the self. For example, when faced with information indicating that they may have a health problem, people reduce their perceptions of the severity of the problem and increase their perceptions of the prevalence of the problem (e.g., Croyle, Sun, & Louie, 1993; McCaul, Thiesse-Duffy, & Wilson, 1992). People have also been shown to shift the perceived importance of health-threatening information. For example, as adolescents engage in risk behaviors more frequently (e.g., smoking and drinking) they recognize that they have put themselves at greater risk for adverse outcome, but at the same time they lower the perceived importance of this information for their behavioral decisions (Gerrard, Gibbons, Benthin, & Hessling, 1996).

People may also have selective memory for health information. When people are faced with evidence of a potential health problem, they selectively recall prior experiences that might render the diagnostic test unreliable (e.g., I haven't slept much the last few days), which could represent a memory bias or a reporting bias (Ditto & Lopez, 1992). Do people have better memory for information that places their behavior in a favorable light? For example, if smokers who eat a healthy diet heard two news reports on the radio, one on the dangers posed by smoking and a second on the dangers posed by eating an unhealthy diet, might they have better memory for one report over the other? In a recent study, Kiviniemi and Rothman (2006) provided people with a series of health recommendations (e.g., eat at least five servings of fruit and vegetables a day) and then conducted a surprise memory test for the recommendations. Participants demonstrated a pronounced bias in their memory for

the recommendations. They were more likely to recall recommendations that indicated that their current behavior was healthy than those that indicated that their current behavior was unhealthy. Returning to the smokers in the prior example, one would expect them to have better memory for the report on diet than the one on smoking. Taken together with the findings from a second study that demonstrated that people's attitudes toward regular alcohol consumption affected their memory for arguments either in favor of or against consuming alcohol (Kiviniemi & Rothman, 2006), there would appear to be evidence of a congeniality bias in people's memory for health information (see Eagly, Chen, Chaiken, & Shaw-Barnes, 1999, for a meta-analysis regarding the effect of attitudes on memory).

There are two dominant explanations for the observation that people are strategic in their response to favorable and unfavorable health information. One perspective emphasizes the motivational concerns that may underlie the observed effects and asserts that because people are motivated to maintain a favorable sense of self, they actively strive to avoid or discount threatening health information. The second perspective emphasizes the degree to which people process information inconsistent with their prior expectations (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998; Ditto et al., 2003; Renner, 2004). When people receive information that is inconsistent or unexpected, it is processed systematically, whereas information that is consistent with one's prior expectations is processed heuristically. Because unfavorable health risk information is generally unexpected, it typically will receive extensive scrutiny. However, according to this framework, unexpected favorable health information would also be scrutinized and, in fact, there is some evidence to support this premise (Renner, 2004). Although investigators typically have approached people's responses to health information from either a motivational or a cognitive perspective, in the end these views are more likely to represent integrated parts of an explanation than two alternative accounts (see Ditto et al., 1998 and 2003, for a thoughtful discussion of these issues). The cognitive activity identified by investigators such as Ditto and his colleagues may provide a description of the mechanisms by which people act on the goals or motives that regulate their preferences.

Although it is clear that motivational concerns have a significant effect on how people respond to potentially unfavorable health information, the failure to appropriately or systematically respond to health information may also reflect the fact that people have difficulty understanding the information, especially statistical information about health risks. There is considerable evidence that people fail to appreciate the fundamental assumptions that impart meaning to numerical estimates of risk (e.g., Black, Nease, & Tosteson, 1995; Shaklee & Fischhoff, 1990) and are sensitive to the specific number used to describe a risk (e.g., 1:10 vs. 10:100; Denes-Raj, Epstein, & Cole, 1995; Slovic, Fischhoff, & Lichtenstein, 1982). In fact, not only are people's understanding of basic probability and numerical concepts quite poor, but their ability to interpret numerical information regard-

ing breast cancer and mammography accurately is systematically dependent on their general ability to deal with numerical information (i.e., numeracy; Schwartz, Woloshin, Black, & Welch, 1997).

One challenge posed by people's innumeracy is that they find it difficult to assign meaning to a given level of risk. Although informing a 40-year-old pack-a-day smoker that quitting will reduce his risk of dying from lung cancer from 131 of 10,000 to 56 of 10,000 is quite precise (see http://cancercontrol.cancer.gov/tcrb/smokersrisk), how people react to that information may not be equally precise. There is some evidence that people spontaneously transform numerical information into discrete categories (e.g., high or low risk; Bottorff, Ratner, Johnson, Lovato, & Joab, 1998) and that the labels assigned to the risk are the best predictors of people's responses to numerical information (Windschitl & Wells, 1996). The difficulty people have assigning meaning to numerical information about their health may help to explain their sensitivity to comparison information (Festinger, 1954). A range of studies has demonstrated that available comparison information regulates how people react to information about their health risk or their health behavior (e.g., Klein, 1997; Rothman, Haddock, & Schwarz, 2001; Sandman, Weinstein, & Miller, 1994; Windschitl & Weber, 1999; see Klein & Weinstein, 1997, for a review). Moreover, when possible, people may selectively recruit comparison standards that place them in a favorable light (Wood, Taylor, & Lichtman, 1985).

Finally, how people respond to health information may be affected by the features they believe characterize a specific health problem. One particularly interesting feature of people's mental model of a health problem (Leventhal, Nerenz, & Steele, 1984) is the social group they associate with the problem. To the extent that people associate an illness with a specific group, they may be less likely to recognize signs of that illness when presented by someone not in that group and may be too quick to recognize signs of that illness when presented by someone who does fit the stereotype. Across a series of studies, Martin and her colleagues have provided a compelling case that the belief that coronary heart disease is more strongly associated with men than women affects the likelihood that lay people and health professionals recognize signs of a heart attack when presented by a woman (Martin, Gordon, & Lounsbury, 1998; Martin & Lemos, 2002). Based on interviews conducted with men and women who had recently been hospitalized due to myocardial infarction (MI), they observed that women were more surprised than men to find that they had an MI and that prior to seeing a doctor, women were less likely to have attributed their initial symptoms to an MI (Martin et al., 2004). In a parallel line of work, Hunt and her colleagues have observed that illness stereotypes regarding who develops eating disorders affect how both lay people and doctors respond to diagnostic information (Hunt, McGorty, Iyer, & Rothman, 2005; Hunt & Rothman, 2004, respectively). Despite the presentation of the exact same symptom information, people are more likely to infer the presence of an eating problem when it is reported by a patient who is Caucasian and/or

a woman. Moreover, a separate set of studies has shown that the presentation of words that cue an illness automatically activates thoughts about members of social groups stereotypically associated with the illness and inhibit thoughts about members of social groups unassociated with the illness (Hunt, Rothmann, Rothman, Iyer, & McGorty, in press). Findings such as these not only illustrate how information about social category membership may regulate responses to health information but also point toward a potential set of mechanisms that might contribute to the disparities in clinical care that have been consistently observed (Smedley, Stith, & Nelson, 2003; van Ryn & Fu, 2003).

Do People Always Resist Health Risk Information?

Given the observation that people may resist unfavorable health information, it is critical to consider what is known about the factors that may facilitate or undermine resistance. Perhaps the first thing that should be noted is that people's hesitancy to accept or recognize unfavorable health information is bounded by rationality (Kunda, 1990). People may capitalize on uncertainties or weakness in evidence that indicate the presence of a health problem but generally do not make up weaknesses that are not present. Similarly, people may take longer to accept an unfavorable than a favorable test result, but when faced with consistent evidence of a health problem, they will generally accept it.

As noted earlier, one factor that may facilitate people's resistance to health risk information is the difficulty they have processing and understanding the information. To the extent that people need to work actively to infer the meaning of new health information, there is greater opportunity for them to engage in biased processing. This would suggest that there is particular value in ensuring that initial messages about a health risk are clear and easy to understand.

Beliefs about and interactions with similarly minded peers may also serve to reinforce the tendency to either discount the health risk information or augment the perceived value of a particular unhealthy behavior (Crandall, 1988; Prentice & Miller, 1993; Suls & Green, 2003; Suls, Wan, & Sanders, 1988). In fact, people may choose to rely on the real or perceived attitudes of their peers rather than thinking through the implications of an issue for themselves. This might be particularly true for people who are dispositionally inclined not to elaborate on information (Cacioppo, Petty, Feinstein, & Jarvis, 1996).

Another factor that may facilitate resistance to health risk information is that concerns about health and illness may not always be the primary motive that guides people's health decisions (McGuire, 1991). As Leary, Tchividjian, and Kraxberger (1994) have noted, people's health behaviors are often regulated by interpersonal, self-presentational concerns. For example, young adults think about tanning more in terms of the social benefits it affords (e.g., enhancing physical appearance) than its associated health risks. Although these concerns may render people less receptive to health risk information,

they may afford other opportunities to convince people that unprotected sun exposure is a problem. For instance, Jones and Leary (1994) found that a pamphlet that describes how the sun can harm one's appearance elicited lower interest in tanning than did a pamphlet describing the health risks posed by sun exposure. In a different study, an ultraviolet photo that revealed evidence of sun damage was able to increase intentions to use sunscreen but elicited behavior change (i.e., a reduction in the number hours spent sunbathing over the next month) only when combined with a brochure on photoaging (Mahler, Kulik, Gibbons, Gerrard, & Harrell, 2003).

Are there factors that render people more receptive to threatening health information? Prior experience with a health problem has a pronounced effect on perceptions of future risk (Weinstein, 1989). Although for health problems that arise as independent acute events (e.g., cavities and warts), people can capitalize on the lessons afforded by their earlier behavior or the lack thereof, in most cases this is not a satisfying solution. There is some evidence that people who are optimistic may be more willing to attend to potentially threatening health information (Aspinwall & Brunhart, 1996). Optimism may help minimize the feelings of distress elicited by consideration of the information or may help people maintain a sense of confidence or control that should they conclude that there is a health problem, they will be able to do something about it. More generally, positive affect may offer a similar set of psychological resources that enable them to confront potential challenges more effectively (see Aspinwall, 1998; Fredrickson, 2001; Trope & Neter, 1994), although the manner with which it influences the processing of health risk information has not been clearly elucidated.

Several studies have also shown that when people are provided an opportunity to affirm their self-worth, they are more open to health-threatening information (Reed & Aspinwall, 1998; Sherman, Nelson, & Steele, 2000). For example, Reed and Aspinwall (1998) adapted a paradigm that had previously been used to demonstrate that people will resist information that might indicate that their behavior is risky (Kunda, 1987; Liberman & Chaiken, 1992) and observed that when people had initially affirmed their self-worth, they were quicker to attend to information that confirmed that their behavior was risky and perceived this information to be more convincing than information that indicated their behavior was *not* risky. As with optimism, self-affirmation may provide people with the psychological resources to cope with and confront distressing information and may afford them the self-confidence that they can deal with these challenges.

Interpreting How People Respond to Health Information: The Importance of Assessing Multiple Indicators

Across a broad array of studies, investigators have reported evidence regarding people's resistance to accepting health information. Although these findings offer a

compelling picture of an individual who, when possible, strives to minimize or undermine the implications of threatening health information, there has been little discussion of what measures are needed to provide a comprehensive assessment of how people respond to health information (Rothman & Kiviniemi, 1999). Should investigators focus on people's knowledge of a risk factor, their perception of personal risk, feelings of concern or worry, or perceptions of the prevalence and the severity of a health problem? Should they assess people's interest in additional information about the issue? How important is the time course of people's reactions? If people's initial reaction is to derogate the information, does this negative view persist over time? The answers to these questions are critically important as the outcome criteria an investigator chooses to emphasize will affect what conclusions are drawn.

For example, think back to an earlier finding we discussed regarding the cognitive consequences elicited by changes in adolescent risk behavior (Gerrard, Gibbons, Benthin, & Hessling, 1996). In this study, Gerrard and her colleagues assessed several distinct beliefs (i.e., perceptions of personal risk, perceived prevalence of the behavior, and impact of health and safety concerns). However, imagine that the investigators had assessed only perceptions of personal risk. Because they found that participants' risk perceptions shifted in line with changes in their behavior, the authors might have concluded that adolescents do not respond defensively to health risk information. However, because they also assessed the perceived prevalence of the risk behaviors and the impact that health and safety beliefs have on decision making, it became clear that how people respond to changes in their risk behavior is quite complex. Although the adolescents acknowledged that their actions placed them at greater risk, they altered other beliefs that served to undermine the implications of the risk information for future behavioral decisions.

To develop a comprehensive understanding of how people respond to threatening health information, it may be critical for investigators to assess a broad range of outcome measures. As discussed earlier, there is considerable evidence that people actively work to minimize the implications of threatening health information (e.g., Ditto & Lopez, 1992; Jemmott, Ditto, & Croyle, 1986). However, a few of these studies included indicators of not only whether people minimized the threat posed by the problem (e.g., measures of severity and prevalence) but also whether they were interested in future opportunities to address the problem (e.g., how the problem could be treated). These studies revealed that even though people minimize the threat posed by the problem, this does not preclude them from being interested in how to deal with the problem (e.g., Croyle & Hunt, 1991; Croyle et al., 1993; McCaul, Glasgow, & O'Neill, 1992). For example, people who learned that they had borderline high cholesterol perceived high cholesterol to be a *less* serious health problem than did those told they had low cholesterol, but they still expressed greater interest in taking steps to lower their cholesterol (Croyle et al., 1993). There is also evidence that potentially threatening

health information can affect people's behavioral intentions, even in the absence of changes in perceived personal risk (e.g., Evers, Bishop, Gerhan, & Weisse, 1997; Gump & Kulik, 1995). For example, heterosexual students increased their perceptions of personal risk for HIV infection in response to a presentation by an HIV+ speaker, but only when the speaker was said to be heterosexual. Yet, they increased their intentions to be HIV tested regardless of the speaker's sexual orientation (Evers et al., 1997).

The observation that people can both minimize and be responsive to threatening health information is consistent with the parallel response model (Leventhal, 1970; Leventhal et al., 1984), which suggests that people independently regulate their affective and behavioral response to a health threat. Specifically, when faced with threatening health information, people may respond in ways that alleviate the distress elicited by the information—thereby attenuating any change in perceived risk and/or perceived severity—and in ways that serve to reduce the likelihood that they might develop the health problem—thereby eliciting changes in behavior. By reducing feelings of distress, people may find themselves in a better position to deal with the challenges posed by the health issue and, in fact, possess the same psychological resources that allow people who are optimistic or have a strong sense of self-worth to be more receptive to health-threatening information (Aspinwall & Brunhart, 1996; Reed & Aspinwall, 1998; Sherman et al., 2000).

Yet, the broader implications of these findings remain uncertain. If people are able to alleviate the feelings of distress elicited by a health problem, will this serve to undermine their motivation to take corrective actions? In situations in which the steps people need to take are relatively simple and require only a single course of action (e.g., being vaccinated for a health threat), there may be little downside to reducing felt distress. However, in situations in which people need to engage in a sustained series of behaviors (e.g., changes in diet or physical activity), the absence of distress may serve to undermine the perceived value of engaging in these new behaviors.

Recent research by Klein and his colleagues also indicates the importance of assessing the interrelationships between people's health beliefs (Klein, 2005; Klein, Blier, & Janz, 2001). Klein has suggested that self-affirmation may allow people to more readily examine threatening health information but that the nature of this examination may not be as thoughtful as one might have hoped. Specifically, Klein and colleagues found that the risk perceptions reported by people who had self-affirmed prior to receiving threatening health information were associated with their self-esteem but not with their standing on relevant risk factors (Klein et al., 2001). People who had not affirmed prior to receiving the threatening information reported risk perceptions that were correlated with their risk factor scores. Thus, self-affirmation may help people pay attention to threatening health information but may lead them to form poorly structured beliefs about the issue. To date, the implications of these findings are uncertain. A critical issue for future research will

be to examine the immediate and long-term consequences of self-affirmation, in particular, the manner in which the beliefs that self-affirmed people form predict subsequent behavioral decisions.

Generally, research in this area would benefit from a closer examination of the structure of the relationships among the thoughts and feelings people hold about a health issue. Research on mental representations of illnesses (Lau & Hartman, 1983; Leventhal et al., 1984) may provide a useful starting point for these investigations as researchers have identified a set of themes that consistently characterize these representations (e.g., identity, cause, consequence, time line, and cure). Although considerable work has gone into describing these themes and developing measures that can assess them reliably (Moss-Morris et al., 2002), tests of the predictive value of these constructs are much more limited and when conducted have primarily examined how beliefs within a particular category are related to outcomes (see Hagger & Orbell, 2003, for a recent review). Greater consideration needs to be paid to the dynamic nature of these beliefs—how are these beliefs formed? How do they affect people's receptivity to new information about an illness? How does change along one dimension (e.g., time line) affect beliefs about another dimension (e.g., consequences)? And, how do beliefs affect and how are they affected by changes in behavior? For example, people's beliefs about the time line of an illness (i.e., acute, cyclic, or chronic) have been shown to be associated with how people treat their condition (Meyer, Leventhal, & Guttman, 1985) and how they feel about their condition (Rabin, Leventhal, & Goodin, 2004). To date, this work has focused exclusively on people with the relevant health problem. Yet, there may be a reason to predict that people's construal of the time line for a health problem, for example, might affect how healthy people process and respond to information about a potential health problem. For example, people who construe an illness as an acute condition may be quite receptive to information about screening for the condition but feel less of a need to learn about how it can be prevented, whereas those who construe an illness as a chronic condition may be quite interested in information about prevention but see little value in information about screening.

Taken together, it is clear that health beliefs are multifaceted and moderated by characteristics of the individual as well as his or her social environment. Should it surprise us, then, that the relation between health beliefs and health behavior is more complex than might be anticipated? It is to these issues that we turn next.

MOVING FROM THOUGHT TO ACTION

People may find that a health issue elicits a rich array of thoughts and feelings. They might know that regular exercise is associated with a range of health benefits and affords an important way to manage one's weight but find that thinking about exercise elicits feelings of boredom and discomfort. At a certain point, though, these thoughts and feelings need to be integrated, and people

need to decide whether or how to act. In this section, we first examine the processes that shape how people make behavioral decisions (e.g., should I test my home for radon?) and then turn to those that influence the transition from intention to action. Because a fundamental practical challenge is designing initiatives that promote healthy behavioral decisions, we conclude this section with an overview of several communication strategies (e.g., message tailoring and message framing) that appear to be effective.

There are presently a substantial number of models that purport to elucidate how people determine whether to adopt a given behavior (health belief model [Rosenstock, Strecher, & Becker, 1988]; precaution adoption process model [Weinstein, 1988]; protection motivation theory [Maddux & Rogers, 1983]; social cognitive theory [Bandura, 1986]; theory of planned behavior [Ajzen, 1991]; theory of reasoned action [Ajzen & Fishbein, 1980]; transtheoretical model of behavior change [Prochaska et al., 1992]). Although these models differ in assumptions regarding how health decisions are structured (e.g., some rely on a stage-based framework, whereas others rely on a continuum-based framework; Weinstein et al., 1998) and in the specific set of beliefs thought to be most closely associated with a decision to take action (see Sutton, 2002b, and Weinstein, 1993, for insightful comparisons of theories), they all rely on some version of a utility-based framework such that the decision to adopt a given behavior is predicated on an analysis of the relative costs and benefits associated with different courses of action (Salovey et al., 1998; Weinstein, 1993). In this chapter, we do not provide a systematic overview of each of the dominant models but instead refer the reader to several recent narrative and meta-analytic reviews (e.g., Albaracin, Johnson, Fishbein, & Mullerleile, 2001; Armitage & Conner, 2001; Conner & Norman, 1996, 2005; Milne, Sheeran, & Orbell, 2000; Sutton, 1998).

Deciding Whether to Take Action

What determines whether a person decides to engage in a new behavior? A critical assumption of nearly all models is that people conclude that the outcomes afforded by engaging in the new behavior (e.g., exercising at the gym 3 days a week) are more desirable than those afforded by one's current behavior (e.g., no regular exercise) (Bandura, 1997; Rothman, 2000; Salovey et al., 1998). Although there is considerable variability in the potential outcomes people consider, one consistent theme is the risk of unwanted health problems. The opportunity to reduce one's risk of a health problem is believed to be a critical motivation to adopt a health behavior. Although risk perceptions are explicitly identified in some models (e.g., the protection motivation theory), they can be readily integrated into other models that do not explicitly reference them (e.g., theory of reasoned action) (Weinstein, 1993).

The premise that people decide to take precautions to reduce the risk of adverse outcomes is widely shared, but it has received surprisingly inconclusive empirical sup-

port. This state of affairs primarily reflects the difficulty of capturing the causal effect of perceived risk on behavioral decisions (Gerrard, Gibbons, & Bushman, 1996; Weinstein, Rothman, & Nicolich, 1998). Because people's behavioral decisions affect their level of risk (e.g., having quit smoking for a year, I conclude that my risk of developing lung cancer has dropped), investigators need to be careful to differentiate between the effect that perceptions of risk have on behavior and the effect that behavior has on perceptions of risk (see Brewer, Weinstein, Cuite, Herrington, & Hayes, 2004, for an excellent empirical example of both of these effects).[1]

It has been suggested that people assess their risk for an unwanted event based on their ability to imagine it happening to them (Armor & Taylor, 1998; Kahneman & Tversky, 1982; Raghubir & Menon, 1998; Rothman & Schwarz, 1998; Sherman, Cialdini, Schwartzman, & Reynolds, 1985). To this end, people may be particularly responsive to information that facilitates this process, such as how the actions they do (or do not) take can affect the onset of a health problem (i.e., antecedents) and what it would be like to experience the health problem (i.e., consequences; for a more comprehensive discussion of this issue, see Rothman & Kiviniemi, 1999). Fear appeals have been one approach to providing people with vivid information about both how a health problem may arise and its severe, unpleasant consequences (Leventhal, 1970). Appeals that elevate feelings of concern about a health issue are effective as long as they provide people with an effective means to reduce their risk (Witte & Allen, 2000).[2] Although investigators have at times predicted that at a certain point elevating feelings of risk or concern about an issue can be counterproductive (e.g., Janis, 1967), there has been little evidence of a curvilinear relation between perceptions of risk and people's behavioral decisions (McCaul, Branstetter, Schroeder, & Glasgow, 1996). In fact, fear appeals lead people to be favorably biased in their evaluation of potential solutions to their health problem (Das, de Wit, & Stroebe, 2003; de Hoog, Stroebe, & de Wit, 2005). Once people recognize that they must deal with a personally relevant health problem, they may be motivated to believe that there is something that can be done about it.

Research on fear appeals would appear to suggest that action is predicated on people recognizing the existence of a significant personal health risk and believing that there is something they can do to effectively reduce their risk. However, the predictive power of these two constructs is almost always assessed simultaneously, and there has been little reason to consider the timing of these effects as well as whether these two classes of beliefs make unique contributions to the behavior change process. The precaution process adoption model (PPAM) (Weinstein, 1988) challenges this view and has argued that investigators need to differentiate between deciding whether to take action and implementing the decision to act (or not act as the case may be), a distinction that is consistent with the perspective offered by several other investigators (Gollwitzer, 1999; Sheeran, 2002). According to the PPAM, information about personal risk is a critical determinant of whether people decide to take ac-

tion, but once someone has decided to do so, additional risk information has limited impact. The availability of information on how, when, and where to take action—and the resulting inference that one is able to perform the desired behavior (i.e., self-effficacy) (Bandura, 1997)—is the critical determinant of whether people move from decision to action. Evidence for the premise that different types of information (i.e., about personal risk or about one's ability to take action) have more impact at specific points in the behavior change process was obtained in an informational intervention designed to promote home radon testing (Weinstein, Lyon, Sandman, & Cuite, 1998).

The observation that as people make their way through the behavior change process they shift from a consideration of the potential outcomes (e.g., the likelihood that one might develop cancer) to a consideration of the potential solution (e.g., the confidence in one's ability to perform the behavior) is strikingly similar to several predictions generated by construal state theory (Trope & Liberman, 2003). According to this theory, when people are considering an event in the future—which may be analogous to deciding whether or not to take action—they focus on its desirability, but when the are considering an event that may soon occur—which may be analogous to formally implementing one's decision—they focus on its feasibility (Liberman & Trope, 1998). This shift in focus from desirability to feasibility is consistent with the premise that beliefs about one's ability to take action become critical as the time for action gets closer. Furthermore, the notion that people initially focus on the desirability of the potential outcome serves as an important reminder that the risk of an unwanted outcome may not always be the critical determinant of the decision to take action.

Because Weinstein chose to focus on decisions regarding precautionary behaviors, he emphasized personal risk and the opportunity to reduce one's risk as the primary way to assess a behavior's value. The perceived value of other classes of behavior may rest not on their ability to avoid or reduce unwanted outcomes (i.e., reduce one's risk of cancer) but, rather, on their ability to afford or approach desirable outcomes (i.e., improve one's appearance). For example, weight loss efforts appear to be motivated more by a desire to improve one's appearance and physical fitness than to avoid health problems (Foster, Wadden, Vogt, & Brewer, 1997). To the extent that we can specify the factors that regulate the perceived value of a behavior, or ideally classes of behaviors, this information could be used to generate a more refined set of predictions regarding the determinants of the behavioral decision-making stage identified in the PPAM.

Generally, what is striking about the conceptual framework generated by Weinstein is that rather than providing a list of constructs that predict a behavioral outcome, it offers a series of predictions regarding the specific conditions under which a construct will or will not facilitate behavior change. By specifying when a construct has an effect, investigators should find it easier to describe the processes that underlie its impact, in turn stimulating the

development of a more comprehensive theoretical framework. This type of information should also enhance the ability of practitioners to develop theoretically grounded intervention strategies (Jeffery, 2004; Rothman, 2004).

Moving from Intention to Action

According to the dominant models of health behavior, once people decide to take action, their decision, perhaps represented by a stated intention to act, produces a pattern of action. Yet, there is considerable empirical evidence that intentions do not always translate into behavior, and a common complaint is that the process by which intentions lead to actions is not well articulated (Bagozzi, 1992; Sheeran, 2002).

What factors might regulate the observed relation between intentions and behavior? From a methodological perspective, it is critical that there is a clear correspondence between the stated intention and the observed behavioral outcome (Ajzen & Fishbein, 1980). Investigators need to be thoughtful from the outset regarding the behavioral outcome they are primarily interested in—is it a specific behavior (e.g., eating fruits and vegetables for snack when at work), a broad class of behaviors (e.g., eating a healthy diet), or a goal or consequence of behavior (e.g., losing weight) (Sutton, 2004)? Because investigators are frequently interested in the consequences of the behavior, they test whether a set of beliefs predict a desired outcome (e.g., change in weight). Yet, in doing so, they frequently fail to consider the behavioral practices that underlie the desired outcome. For example, Linde, Rothman, Baldwin, and Jeffery (2006) have illustrated how the impact of self-efficacy on weight loss outcomes critically depends on the ability of self-efficacy to predict effective weight loss practices (i.e., changes in diet and physical activity).

By adding perceived behavioral control to the theory of reasoned action, Ajzen (1991) strove to better specify the relation between intentions and behavior. Perceptions of behavioral control are thought to represent the degree to which people believe they have control over their actions and thus their confidence in their ability to turn their intentions into actions. Although perceived behavioral control has been shown to enhance investigators' ability to predict behavior, the specific manner by which it is expected to affect behavior is uncertain (Armitage & Conner, 2001). It could have a direct effect on behavior, an indirect effect on behavior through its effect on intentions, a moderating effect on the relation between intention and behavior, or perhaps some combination of effects (see Sutton, 2002a and 2004, for an excellent discussion of these issues)? The premise that control would regulate people's ability to act on their intentions would seem to be consistent with viewing its status as a moderating factor, but when effects have been observed, they almost always have taken the form of a direct effect of control on behavior. Interestingly, a study of the determinants of behavior change in patients with coronary heart disease revealed that patients' perceptions of behavioral control regarding exercise and smok-

ing cessation predicted objective changes in behavior (i.e., increases in exercise and decreases in smoking), whereas their stated intention to change did not predict behavior (Johnston, Johnston, Pollard, Kinmonth, & Mant, 2004). Although one must be cautious in drawing strong conclusions from a single study, this pattern of results might suggest that once people have decided to make a change (e.g., quit smoking), the degree to which they want to change may be less important than their confidence that they can act on their desire.

A particularly active line of inquiry regarding the relation between intentions and behavior has been to examine what characterizes intentions that people are more likely to carry out as compared to those they are less likely to act on. In particular, investigators have been examining the stability or strength of people's intentions (e.g., Conner, Norman, & Bell, 2002; Conner, Sheeran, Norman, & Armitage, 2000; Sheeran, Orbell, & Trafimow, 1999; see Cooke & Sheeran, 2004, for a review). Consistent with research on attitude strength (Petty & Krosnick, 1995), strong intentions are thought to be better predictors of behavior. Intention strength has been operationalized by assessing its stability over time (Sheeran et al., 1999), and empirical evidence has consistently found strong intentions to be a better predictor of behavior. In fact, Conner and colleagues (2002) demonstrated that, controlling for initial behavior, strong intentions regarding food choices (which were based on assessments of stability over a 6-month time window) were associated with healthy eating habits 6 years later.[3]

What is particularly striking about research on intention strength is that it may help to explain why a diverse set of factors moderate the relation between intentions and behaviors. Specifically, Sheeran and Abraham (2003) have shown that the moderating effect of factors such as intention certainty, self-schemas, and anticipated regret on the intention–behavior relation is mediated by intention strength. Thus, people who are certain about their intentions, who see a behavior as important to their sense of self, or who anticipate feelings of regret if they do not act, all form intentions that are strong and stable. The observation that the same mechanism underlies a disparate array of moderated effects provides a reassuring sense that it may be possible to develop parsimonious explanations for the relation between intentions and behavior. Moreover, it suggests there may be value in intervention strategies that can promote stable intentions. The challenge to future endeavors will be to delineate the processes that promote the formation of strong intentions.

One method that augments the translation of intention into actions is the use of implementation plans (Gollwitzer, 1999; Gollwitzer, Bayer, & McCulloch, 2005). Several studies have shown that asking people to specify how, when, and where they will implement their intentions increases the likelihood that they will perform the behaviors such as making healthy food choices, taking a vitamin supplement, or screening for cervical cancer (e.g., Armitage, 2004; Sheeran & Orbell, 1999, 2000). Implementation plans are thought to affect people's

sense of control over their behavior, but not their motivation to act. Consistent with this view, Sheeran, Webb, and Gollwitzer (2005) demonstrated that the beneficial effect of implementation plans is contingent upon people having initially endorsed the relevant behavioral goal.

One thing that an implementation plan does is render salient the structural and behavioral factors that may underlie people's ability to take action. For example, if people formulate a plan to eat fruits and vegetables as an afternoon snack at work, they are more likely to recognize what they need to do in order to act (e.g., buy fruits and vegetables and prepare and pack them each morning). The formulation of an implementation plan may also help people better specify their initial intentions. For example, a smoker might shift from the general intention of cutting down on the number of cigarettes she smokes to the more specific intention of not smoking any cigarettes after dinner. The principle of correspondence that helps investigators detect the effect of beliefs on behavior at the sample level should similarly help an individual maximize the chance he or she will act on his or her intention.

In a recent application of the theory of planned behavior, Bryan, Fisher, and Fisher (2002) reported that the model fit the data best when intentions to use condoms were allowed to predict a set of preparatory behaviors (i.e., buying condoms, carrying condoms, and discussing condoms with partner), which in turn predict condom use. What is particularly noteworthy about this study is that it illustrates an attempt to unpack the intervening steps that underlie behavioral decisions. To the extent that a behavioral outcome depends on a series of behavioral decisions, investigators need to consider how the experiences afforded by the initial set of behaviors may affect the beliefs that, in turn, predict subsequent behavioral decisions. For example, after forming an intention to use condoms, someone might attempt to purchase condoms in his or her neighborhood grocery store. If he or she finds purchasing condoms to be difficult or embarrassing, this is likely to affect the status of the initially stated intention.

Are All Actions Intentional?

Up to this point, we have implicitly assumed that all behavioral decisions reflect intentional, planned choices. Although this assumption is likely appropriate for decisions about most health behaviors, Gibbons and Gerrard have proposed that it may not provide an accurate account of the decisions adolescents and young adults make to engage in risky behaviors (Gibbons, Gerrard, & Lane, 2003). According to the prototype/willingness model, there are two paths that guide people's behavioral decisions (Gibbons, Gerard, Blanton, & Russell, 1998). People's behavior may reflect their intentions as specified by models such as the theory of planned behavior, but they may also reflect their willingness to engage in a particular behavior under particular circumstances. Measures of behavioral willingness tap the fact that people may not intend to engage in a given behavior (e.g., people don't leave the house intending to drive home drunk

at the end of the evening) but may respond favorably to the presence of situational factors that afford that behavior (e.g., at the end of a night out, faced with limited options, people may be willing to drive drunk). According to the model, behavioral willingness is predicated on the perceived acceptability of the relevant behavior, which is typically assessed as people's perception of the prototypical person who engages in the risky behavior (Gibbons & Gerrard, 1995).

Across several studies, Gibbons, Gerrard, and their colleagues have demonstrated that measures of behavioral willingness are significant predictors of the onset and increases in risky health behaviors such as drinking, substance use, smoking, and unsafe sexual behaviors, even when controlling for behavioral intentions (see Gibbons et al., 2003, for a review). Consistent with the broader model, the impact of behavioral willingness is particularly pronounced in samples of adolescents (e.g., Gibbons et al., 2004). As people age, patterns of risky behavior, such as those noted above, appear to be more strongly a function of their behavioral intentions (Gibbons, Gerrard, Pomery, & Reimer, in press). Yet, this does not necessarily mean that the behaviors of older adults are always the result of intentional, planned choices. Behavioral willingness may prove to be an important predictor of a new set of adult-oriented risk behaviors such as adultery, tax evasion, or insider trading.

The impact of behavioral willingness has also been shown to be more sensitive to context than the impact of behavioral intentions. For example, in a sample of young adolescents, behavioral willingness predicted changes in substance use, but this effect was particularly strong when those kids lived in high-risk neighborhoods that provided the situational affordances for the kids to respond to (Gibbons et al., 2004). The effect of behavioral intentions on substance use was the same across high- and low-risk neighborhoods.

COMMUNICATION STRATEGIES

A primary method for motivating people to change their behavior has been to provide them with information that will persuade them to alter it (Eagly & Chaiken, 1993). To be effective, communication strategies are needed that get people not only to attend to the message but also to process it in a manner that maximizes its impact on how they think and feel about the issue (Petty & Wegener, 1998). In this section, we consider two communication strategies, message tailoring and message framing, that have been shown to motivate behavior change effectively.

Message Tailoring

Message tailoring is based on the premise that people will pay more attention to and be more persuaded by information that speaks directly to their own personal concerns.[4] For example, a smoker who is worried about the social stigma of smoking would receive information on that topic, whereas a smoker who is worried about how secondhand smoke may affect her family's health would

receive information about that topic. Tailored health messages have been shown to be more effective than generic messages that provide all individuals with the same information (Brug, Glang, van Assema, Kok, & van Breukelen, 1998; Dijkstra, De Vries, & Roijackers, 1998a, 1998b; Kreuter, Oswald, Bull, & Clark, 2000; see Skinner, Campbell, Remer, Curry, & Prochaska, 1999, for review).

Why might tailored messages be more effective? Contemporary models of attitude change and persuasion offer a framework for understanding how and when tailored messages are maximally effective (Petty, Barden, & Wheeler, 2002). To the extent that tailored messages are perceived to be personally relevant, the message will be processed more extensively, which, in turn, should increase its influence on people's thoughts and feelings about the health issue. All else being equaled, greater elaboration of a strong, health message is desirable, as well-reasoned attitudes are more stable over time and better predictors of behavior (Petty & Wegener, 1998). Consistent with this perspective, studies have demonstrated that tailored messages are more likely to be read and remembered than are nontailored messages (e.g., Skinner, Strecher, & Hospers, 1994), are more likely to be discussed with others (e.g., Brug, Steenhuis, van Assema, & de Vries, 1996), and are perceived as more interesting and engaging (e.g., Brug et al., 1996; Kreuter, Bull, Clark, & Oswald, 1999).

The principle of tailoring messages is consistent with several lines of research in social psychology that have demonstrated the benefits of matching a message to an aspect of a person's personality or motivational concerns (e.g., Williams-Piehota, Pizzaro, Navarro, Mowad, & Salovey, 2006; Williams-Piehota, Pizzaro, Schneider, Mowad, & Salovey, 2005; Williams-Piehota, Schneider, Pizzaro, Mowad, & Salovey, 2004). Higgins (2000) has proposed that messages that correspond to a person's dominant self-regulatory concern—the tendency to be preoccupied moreso with accomplishments than safety or vice versa—benefit from the subjective experience of fit. Moreover, people appear to enjoy working on tasks that are structured to match their self-regulatory concerns (e.g., Freitas & Higgins, 2002) and may even perform better (Spiegel, Grant-Pillow, & Higgins, 2004). It may be that when people receive a tailored message they are more likely to construe their motivation to modify their behavior as reflecting intrinsic concerns, which have been shown to be predictive of behavior change (e.g., Williams, Ryan, Rodin, Grolnick, & Deci, 1998).

A critical challenge to further advances in tailoring is the development of a conceptual framework that can help determine or prioritize the dimensions along which information should be tailored. Theories of health behavior could potentially guide decisions regarding the type of information that should be tailored. For example, as discussed earlier, there may be benefits to tailoring message content (e.g., information about personal risk vs. information about self-efficacy) based on where people are in the decision-making process (Dijkstra, De Vries, Roijackers, & van Breukelen, 1998; Weinstein, Rothman, & Sutton, 1998). To the extent there is an ad-

vantage to matching information to where people are in the decision-making process, consideration could be given to whether there is any added advantage to tailoring the presentation of that material to an individual's unique concerns.

To date, investigators of tailored health messages have assumed that people's motivations or goals for changing their behavior are all equally effective. However, there is a substantial body of research in academic settings that has shown that approach goals (i.e., goals that are characterized by a desire to reach a favorable goal state) are more strongly associated with desirable outcomes than are avoidance goals (i.e., goals that are characterized by a desire to stay away from an unfavorable goal state; Elliot & Church, 1997; Elliot & McGregor, 1999). This would suggest that tailoring messages to an approach goal (e.g., dieting to improve one's appearance) might prove to be more effective than tailoring a message to an avoidance goal (e.g., dieting to avoid future health problems).

Only one study has tested the differential impact of approach and avoidance goals on health behavior change (Worth, Sullivan, Hertel, Jeffery, & Rothman, 2005). Although the type of goal people generated at the outset of a smoking cessation program predicted their ability to quit smoking, smokers who had generated more *avoidance* goals were more likely to initiate a quit attempt. The observation that avoidance goals were beneficial is quite striking and may be due to two critical factors. First, the avoidance goals that predicted behavior change were those that focused on curing a current problem (e.g., getting rid of the smell in one's car), a subset of goals that have not been considered in prior research. Consistent with prior findings, avoidance goals that focused on preventing a future unwanted outcome (e.g., not developing cancer) did not predict success. Second, the behavioral task at hand may affect the relative impact of different goals. For example, smoking cessation is typically construed in terms of trying not to be a smoker rather than to be a nonsmoker (Gibbons & Eggleston, 1996). This construal may lead avoidance goals to better fit the task at hand, a finding that would be consistent with other demonstrations of correspondence. This is clearly an area that needs further investigation.

Finally, tailoring initiatives may benefit from consideration of how the broader social context affects the perceived value of information about a given health practice. According to deviance regulation theory (DRT; Blanton & Christie, 2003), people will be more responsive to information about deviant than normative behavioral practices. Because health messages can emphasize either engaging in a particular behavior (e.g., smoking) or not engaging in the behavior (e.g., not smoking), DRT predicts that people will be more responsive to messages regarding the behavior that is not considered normative. For example, when students believe that getting a flu shot is normative, they are more responsive to information about people who do not get vaccinated, but when they believe that getting a flu shot is not normative, they are more responsive to information about people who do get vaccinated (Blanton, Stuart, & VandenEijden, 2001). This framework would suggest that efforts to tai-

lor messages to people's dominant concerns (e.g., their thoughts and feelings about dieting) would benefit from considering people's normative beliefs about the behavior.

By definition, tailored messages provide people with information that addresses their personal needs and concerns. The value of such an approach would appear to be self-evident. Yet, any evaluation of tailoring must weigh the implementation costs associated with assessing people's personal profiles as well as the reliability and validity of the instruments used to identify a person's needs. Interestingly, Dijkstra (2005) has observed that the benefits of tailoring may be obtained with only minimal information. He found that a personalized smoking cessation message (which included the participant's name and some basic smoking history information) was as effective as one that provided tailored feedback, and both messages were more effective than a generic message. It may be that the relative utility of different levels of tailoring may depend on the goal of the researcher. If the researcher needs only to heighten how engaged someone is when reading a brochure, personalization may be sufficient. However, if the researcher needs to provide people with strategies that can help them overcome barriers to action, more extensive tailored feedback may be necessary. Delineating the dose–response relationship between tailored information and behavioral outcomes may be an important next step in this area.

Message Framing

Messages designed to promote a health behavior can be constructed to focus on the benefits of performing the behavior (a gain-framed appeal) or the costs of failing to perform the behavior (a loss-framed appeal). For example, a gain-framed brochure designed to promote mammography would emphasize the health benefits afforded by screening, whereas a loss-framed brochure would emphasize the health costs of failing to be screened. According to prospect theory (Tversky & Kahneman, 1981), people act to avoid risks when considering the potential gains afforded by a decision (they are risk averse in their preferences) but are more willing to take risks when considering the potential losses afforded by a decision (they are risk seeking in their preferences). This premise has led investigators to assess the relative impact of gain- and loss-framed health messages have on people's behavioral practices.

The relative impact of gain- and loss-framed messages is thought to be based on how people construe the risk posed by performing the targeted behavior (Rothman, Kelly, Hertel, & Salovey, 2003; Rothman & Salovey, 1997). If performing a behavior is perceived to involve some risk, loss-framed appeals should be more persuasive, but if a behavior is perceived to afford a relatively certain outcome, gain-framed appeals should be more persuasive. Rothman and Salovey (1997) observed that the function served by a health behavior can operate as a reliable heuristic to predict whether people perceive engaging in a behavior to be risky. Specifically, detection behaviors serve to detect the presence of a health prob-

lem, and because they can inform people that they may be sick, initiating the behavior may be considered a risky decision. In contrast, prevention behaviors such as the regular use of sun screen or condoms forestall the onset of an illness and maintain a person's current health status. In fact, these behaviors are risky only to the extent that one chooses *not* to take action. Taken together, this distinction suggests that loss-framed appeals would be more effective in promoting the use of detection behaviors but gain-framed appeals would be more effective in promoting the use of prevention behaviors.

Empirical evidence obtained across both laboratory and field studies have provided strong support for this framework. Loss-framed messages have been shown to elicit greater interest in or performance of detection behaviors (e.g., Banks et al., 1995; Meyerowitz & Chaiken, 1987; Schneider et al., 2001; but see Lauver & Rubin, 1990; Lerman et al., 1992), whereas gain-framed messages have been shown to elicit greater interest in or performance of prevention behaviors (e.g., Detweiler, Bedell, Salovey, Pronin, & Rothman, 1999; Rothman, Salovey, Antone, Keough, & Martin, 1993; see Rothman, Stark, & Salovey, 2006, for a more comprehensive review of this literature). Moreover, systematically manipulating whether people believed a behavior prevented or detected a health problem has been shown to moderate the relative influence of gain- and loss-framed appeals (Rothman, Martino, Bedell, Detweiler, & Salovey, 1999).

Despite the fact that the observed distinction between prevention and detection behaviors rests on how people construe the behavior, investigators have paid limited attention to variability along this dimension. A few studies have examined whether people's perceptions of a screening behavior moderate the influence of framed appeals and have obtained results consistent with the guiding framework (Apanovitch, McCarthy, & Salovey, 2003; Meyerowitz, Wilson, & Chaiken, 1991). Some initial experimental work designed to manipulate how people construe a behavior—for example, is the test designed to detect a favorable or an unfavorable outcome?—has also provided supporting evidence (for a more complete discussion of these findings, see Rothman, Bartels, Wlaschin, & Salovey, 2006). Future work in this area would benefit from a better understanding of the factors that shape how a behavior is construed. For example, Orbell, Perugini, and Rakow (2004) have observed that people who are inclined to focus on the future consequences of a behavior will tend to think about the long-term benefits of a screening test (i.e., that it affords one the opportunity to treat a health problem), whereas those who are inclined to focus on more immediate consequences will tend to think about the immediate costs of a screening test (i.e., the possibility of learning one has a health problem).

Although research on message framing has primarily focused on the correspondence between the frame of the message and features of the targeted behavior, differences in individual's sensitivity to favorable and unfavorable outcomes has been shown to moderate people's response to gain- and loss-framed messages (Cesario, Grant, & Higgins, 2004; Lee & Aaker, 2004; Mann,

Sherman, & Updegraff, 2004). The primary assumption that underlies this work is that a message will prove to be more persuasive to the extent it fits or is compatible with how the recipient thinks and reasons about his or her environment. For example, gain-framed appeals will be more effective for people whose primary self-regulatory strategy is to focus on hopes and aspirations (i.e., a promotion focus), whereas loss-framed appeals will be more effective for people whose primary self-regulatory strategy is to focus on duties and obligations (i.e., a prevention focus) (Cesario et al., 2004; Lee & Aaker, 2004).

An important challenge for future research will be to integrate these recent findings with prior evidence that the function of the behavior moderates the influence of framed appeals. We have recently suggested that these findings may be based on a similar set of processes (Rothman, Bartels, et al., 2006). Specifically, performing a screening behavior may be considered a duty or obligation and may serve to induce people into a prevention-focus mindset (facilitating the impact of loss-framed messages), whereas performing a prevention behavior may be considered a choice rather than an obligation and may serve to induce people into a promotion-focus mindset (facilitating the impact of gain-framed messages). Although the premise that health behaviors can systematically invoke a promotion or a prevention mindset is compelling, further research is clearly needed to specify the predicted associations.

Evaluating Effectiveness

The impact of the communication strategies we have examined is typically determined by assessing changes in people's beliefs about the health issue (people's attitude toward or intention to get a mammogram become more favorable) or changes in people's behavior (rates of mammography utilization increase). Although these measures have clear and important value, it may be wise for investigators to begin to consider using a broader set of indicators. In particular, there is growing evidence that structural features of people's beliefs have important behavioral implications. As was reviewed earlier, the strength of a person's behavioral intention was a critical determinant of whether it led to a change in behavior (e.g., Sheeran & Abraham, 2003). This would suggest that messages need to be able to elicit both favorable and strong beliefs. Research on the consequences of resistance to persuasion also suggests that investigators need to attend to more than just the overall evaluation of a person's attitude. When people feel they have resisted the influence of a strong persuasive message, they become more certain about their initial attitude, which leads them to be more resistant to subsequent persuasive efforts (Tormala & Petty, 2002, 2004). This finding would suggest that health messages that appear to have no meaningful impact on respondents' beliefs might in fact have had more negative effect than had previously been thought.

Investigators may also want to consider how responses to messages unfold over time. The persistence of an individual's response to a health message has not received much systematic attention as investigators have primarily relied on between-subjects comparison strategies (see Sutton, 2004, for a broader discussion of the how an overreliance on between-subjects comparisons has constrained progress in health behavior theory). If people's responses to a health message are tracked over time, investigators will have to consider how their subsequent experience with the health issue, and perhaps the targeted health behavior, affects their response to the initial message. This is likely to be a particularly critical issue when a message is developed to promote a behavior that requires an ongoing series of actions (Rothman et al., 2004; Rothman, Hertel, Baldwin, & Bartels, in press). Recent work by Albaracin, Cohen, and Tumkale (2003) illustrates how the actions people take after having been exposed to a message can have dramatic effects on their response to the initial message. Specifically, people who engaged in a behavior (e.g., trying a new alcoholic beverage) after having been exposed to an abstinence-based message subsequently held more favorable attitudes toward the behavior than did those who had initially received a message promoting moderation or those who had not engaged in the behavior.

SUSTAINING PATTERNS OF BEHAVIOR

If changes in people's behavior are to afford improvements in health outcomes, those changes need to be sustained over time. People not only must lose weight or start a program of regular exercise but also must maintain their lower weight or continue to exercise regularly. Similarly, screening behaviors typically involve standardized regimens. Yet, behavioral maintenance remains discouragingly elusive. Even people who have successfully initiated a new pattern of behavior more often than not fail to sustain it over time (e.g., diet and exercise to produce weight loss [Jeffery et al., 2000], smoking cessation [Ockene et al., 2000]; substance abuse [Hunt, Barnett, & Branch, 1971; Marlatt & Gordon, 1985]).

To date, researchers have primarily focused on predicting a single behavioral outcome such as whether a woman has obtained a mammogram (e.g., Aiken, West, Woodward, & Reno, 1994), a smoker has quit (e.g., Norman, Conner, & Bell, 1999), or a man has modified his diet (e.g., Conner et al., 2002). Although these outcomes may include assessments of a series of behaviors (e.g., food choices over a 4-week period), very limited attention has been paid to the psychological experiences and decision-making processes that underlie behavioral patterns. To a certain extent, this reflects the fact that the dominant theoretical models assume either implicitly or explicitly that the factors that underlie a decision to initiate a behavior are the same as those that underlie the decision to maintain it (Rothman, 2000; Rothman et al., 2004).

Even models that were developed to describe how the behavior change process unfolds over time (e.g., Prochaska et al., 1992; Weinstein, 1988) offer little insight into the processes that might guide the ongoing performance of behavior. Although these models in-

clude a distinction between initial and maintained behavior, it rests on the length of time a behavior has been performed rather than the possibility that the decision to maintain a new behavior is distinct from the decision to implement the change in the first place (Prochaska & Velicer, 1997). In a similar manner, investigators have provided thoughtful descriptions of habits as a sequence of successfully enacted behaviors (Ajzen, 2002; Ouellette & Wood, 1998) but have not delineated the processes that determine whether or how an initially enacted behavior becomes a habit.

Several studies have revealed that it is difficult to account for certain patterns of behavior with current models of behavioral decision making. For example, Norman and colleagues (1999) observed that predictors from the theory of planned behavior were able to predict smokers' intentions to quit and whether they subsequently quit but were unable to explain how long smokers were able to sustain their abstinence. Similarly, in a study of predictors of attendance at a regular yearly physical over a 2-year period, constructs from the theory of planned behavior were able to predict whether or not people ever sought a physical but were unable to differentiate between different patterns of attendance (i.e., people who went the first year but not the second, people who went the second year but not the first, and people who went both years; Sheeran, Conner, & Norman, 2001). Intervention studies have also observed that methods that are effective at helping people initiate changes in their behavior do not have a similar effect on efforts to maintain those changes (e.g., McCaul, Glasgow, & O'Neill, 1992; Perri, Nezu, Patti, & McCann, 1989).

Rothman and colleagues (Rothman, 2000; Rothman et al., 2004, in press) have proposed that there is theoretical and practical value to more precisely delineating the processes that guide the transition from behavioral initiation to behavioral maintenance and that, in doing so, investigators will be able to discern factors that facilitate (and inhibit) initial behavioral efforts and those that facilitate (and inhibit) sustained behavioral efforts. To this end, they have proposed distinguishing between the following four phases of the behavior change process: *initial response*, *continued response*, *maintenance*, and *habit*. These phases capture the behavioral processes that begin once someone implements a decision to take action (e.g., enrolling in a smoking cessation program), transitioning out of what Prochaska and colleagues (1992) have characterized as the preparation stage.

The initial response phase captures the translation of an intention into an initial successful action (e.g., a smoker successfully remains smoke-free for a week) and thus the dominant issues regarding this phase have been discussed in a prior section of this chapter. The continued response phase represents the period of time in which people must sustain their initial beliefs about the benefits of the new behavior and their ability to perform the behavior in the face of the experiences—both pleasant and unpleasant—afforded by their actions. The length of time people remain in this phase is likely to differ across both behavioral domain and person. Some people may find it easy to master the new pattern of behavior, whereas others may find it a continual struggle. Similarly, some behavioral domains such as dietary change involve a complex series of behavioral modifications—which should lengthen this phase, whereas other domains such as taking a medication daily involve a more limited set of challenges—which should shorten this phase.

The maintenance and habit phases capture the processes that determine whether a successfully enacted change in behavior is sustained. During the maintenance phase, the decision to maintain a behavior is based on a regular assessment of whether the behavior is worth the effort, whereas during the habit phase the behavior is maintained without any consideration of a behavioral alternative (for a more comprehensive discussion of these phases, see Rothman et al., 2004, in press). We have chosen to use these phases to organize the issues reviewed in this section of the chapter.

Continued Response

Given that successfully initiating a change in behavior is predicated on people's confidence in their ability to execute the behavior and their belief that engaging in the new pattern of behavior will meaningfully improve their lives (Bandura, 1997; Salovey et al., 1998), the primary challenge to continuing these initial efforts is sustaining these beliefs in the face of their experiences with the new behavior. To the extent that people find the new behavior to be unpleasant or feel that it demands substantial mental and/or physical energy, their commitment to and confidence in the behavior may weaken, thereby undermining the likelihood that the behavior will be sustained (Baumeister, Heatherton, & Tice, 1994; Schmeichel & Baumeister, 2004). Thus, how the costs and rewards afforded by a given behavior unfold over time affects people's ability to negotiate this phase of the behavior change process. For example, because the costs associated with a behavior are often tied to performing the behavior (e.g., having to continually monitor what one eats or having to get up early to exercise), they tend not only to appear with the onset of the behavior but also to continue as long as the behavior is performed (Jeffery, Kelly, Rothman, Sherwood, & Boutelle, 2004). These salient, persistent costs may challenge the optimism that motivated people to initiate a change in behavior.

What factors might help people sustain their motivation and confidence in their ability to continue a new pattern of behavior? There is reason to believe that people need to possess sufficient psychological resources to maintain their beliefs. Baumeister and his colleagues (Baumeister et al., 1994; Schmeichel & Baumeister, 2004) have consistently shown that deficits in self-regulatory strength contribute to failures to regulate one's behavior. Thus, individual differences in the ease with which people can generate and sustain favorable expectations may be important. People who score high on self-esteem and those who score high on dispositional optimism have been shown to be more likely to generate positive expectations (e.g., Affleck et al., 2001; Aspinwall & Taylor, 1993; Gibson & Sanbonmatsu, 2004; Helgeson,

2003) and, moreover, are better able to sustain those beliefs in the face of difficult experiences (e.g., Campbell & Fairey, 1985; Gibson & Sanbonmatsu, 2004). At the same time, people who have difficulty not attending to thoughts or images associated with their prior (unwanted) behavior may have considerable difficulty negotiating their way through this phase of the process. For example, Waters and colleagues (2003) observed that the extent to which smokers, after having quit for a day, exhibited a Stroop-based bias for smoking-related stimuli predicted whether they subsequently lapsed and the time to the lapse.

Models of relapse prevention have proposed that people's attributions for any difficulties they have changing their behavior will regulate the transition from a lapse to a relapse (Marlatt & Gordon, 1985; Witkiewitz & Marlatt, 2004). Although the premise that people might cope better with a lapse that can be attributed to a controllable, external event is compelling, there is little evidence that people's attributions predict their response to a lapse in their behavior (Shiffman et al., 1996). However, Shiffman and his colleagues (2000) have shown that after a lapse, temporal shifts in smokers' confidence in their ability to quit do predict whether a lapse progressed into a relapse. Although not directly linked to people's ability to cope with difficulties, there is considerable evidence that people are more likely to change their behavior if their efforts are intrinsically or autonomously motivated (e.g., Williams, Gagne, Ryan, & Deci, 2002; Williams, McGregor, Zeldman, Freedman, & Deci, 2004). Moreover, Williams and colleagues (2004) have shown that autonomous motivation promotes behavior change due to its ability to enhance people's confidence in their ability to perform the desired behavior.

Taken together, there are a number of factors that likely regulate people's ability to sustain the beliefs that underlie the change in their behavior. Yet, greater consideration needs to be given to the patterns of change over time, in both behavior and beliefs, which are associated with these factors. Further research should also examine the interrelations between these factors, with a particular emphasis on whether the effects attributed to different constructs such as dispositional optimism and self-esteem are, in fact, distinct. Finally, consideration should be given to how features of the behavior in question affect the ease or difficulty with which people are able to successfully continue a new pattern of behavior. A rich description of how the rewards and costs afforded by a behavior typically unfold over time would be particularly valuable.

Maintenance and Habit

What determines whether a successfully initiated pattern of behavior is maintained? Rothman and colleagues have proposed that the decision criteria that underlie behavioral maintenance are distinct from those that underlie behavioral initiation (Rothman, 2000; Rothman et al., 2004, in press). Specifically, whereas decisions regarding behavioral initiation are based on expectations about future outcomes, decisions regarding behavioral mainte-

nance involve a consideration of the experiences afforded by the new pattern of behavior and a determination of whether those experiences are sufficiently satisfying to warrant continued action. The feeling of satisfaction indicates that the initial decision to change the behavior was correct, and it serves to justify people's efforts to sustain the behavior. Given that people who have reached this phase in the behavior change process have shown they can successfully perform the behavior, they may pay less attention to questions regarding their ability to engage in the behavior (i.e., "Can I do it?") and more attention to questions regarding their evaluation of the behavior (i.e., "Do I want to do it?").

Although it seems intuitive that the more satisfied a person is with a behavior, the more likely he or she will continue to perform it; the prospective effect of satisfaction on behavior has received only limited empirical attention. Dijkstra and Borland (2003) reported that the more ex-smokers missed experiences from when they smoked, the more likely they were to relapse (especially when they were low in self-efficacy). Evidence that perceived satisfaction with behavior change predicts sustained behavior change has also been obtained in two intervention studies designed to promote weight loss (Finch et al., 2005) and smoking cessation (Hertel et al., 2005). What is particularly noteworthy about these studies is that they have disentangled predictors of behavioral maintenance from predictors of behavioral initiation. Typically, claims regarding the determinants of behavioral maintenance are based on testing whether a psychological state (e.g., self-efficacy at baseline) predicts a distal behavioral outcome (e.g., smoking status 12 months later). However, this analytic model cannot determine whether, in the current example, people's initial feelings of self-efficacy contribute to their willingness to maintain their behavior over and above its effect on their initial behavioral efforts. To test the premise that the predictors of behavioral initiation and behavioral maintenance are distinct, investigators must attend to where people are in the behavior change process. For example, Baldwin and colleagues (2006) examined predictors of behavior in a sample of smokers who had just participated in a smoking cessation program. Among those smokers who had failed to quit during the program, perceived self-efficacy predicted whether they were able to initiate another quit attempt, whereas among those who had successfully quit during the program, perceived satisfaction with being tobacco free (and not perceived self-efficacy) predicted whether someone maintained smoking cessation.

What determines whether people are satisfied with their behavioral experience? Although some experiences may prove easy to evaluate (e.g., severe side effects from a treatment), in many cases people may find it difficult to determine how much of a change is sufficient to feel satisfied? For instance, if a man enrolls in a weight loss program in order to be able to wear clothes in a smaller size, what size does he need to reach to be satisfied with his efforts? There is considerable laboratory evidence that people often examine their experiences in light of their expectations (Gollwitzer, 1996; Schwarz & Strack, 1991). Thus, how one feels about being able to wear pants with a

30-inch waist depends on whether you initially expected to reach a 28-inch or a 32-inch waist. One implication of this framework is that while raising people's expectations regarding the benefits of behavior change will help motivate them to initiate a change in their behavior, it will make it much more difficult for them to meet those expectations, which serves to undermine their motivation to maintain the behavior (Rothman, 2000). Although there is evidence in the consumer behavior literature that people's satisfaction with a product is predicated on their experiences meeting their expectations (Szymanski & Henard, 2001), to date, two intervention studies designed to promote healthy behavior have failed to support the predicted, ironic effect of expectations on feelings of satisfaction (Finch et al., 2005; Hertel et al., 2005).

There are a number of other factors that may shape people's feelings of satisfaction. For instance, given that many health behaviors afford myriad outcomes, some people—perhaps those high in self-esteem or in dispositional optimism—may be better able to selectively or strategically monitor for desirable outcomes. For example, when some people find that weight loss did not improve their social life in the way that they had hoped it would, they are able to alter their priorities and derive satisfaction from the improvements in physical health they experienced. People may also find a behavior more satisfying to the extent that the process of engaging in the behavior fits their dominant self-regulatory orientations (Higgins, 1999). To the extent that satisfaction with behavior change is a critical determinant of behavioral maintenance, a priority for future research will be to delineate the processes that shape people's experience of satisfaction (see Rothman et al., in press, for a more complete discussion of this issue).

During the behavioral maintenance phase, people will engage in a sustained pattern of behavior. How is this distinct from a habit, which is similarly represented by a sustained pattern of behavior? According to the present framework, an important distinction is that people in the maintenance phase continue to monitor the value of the behavior, whereas those in the habit phase do not. This would suggest that people in the habit phase will be less sensitive to new information or new experiences regarding the behavior (e.g., Ferguson & Bibby, 2002). It is important to note that this does not mean that people in the habit phase do not value the behavior. They just no longer need to verify or test its value (Wood, Quinn, & Kashy, 2002). Habits are likely to persist as long as there is not a significant shift in the environment in which the habit is operating. Consistent with this perspective, Wood, Tam, and Guerrero-Witt (2005) found that students transferring to a new school will sustain their exercise habits if they are able to preserve continuity in the environment (e.g., they can exercise at the same time of day), but if there is a change in the environment (e.g., they have to exercise at a different time of day), the likelihood that they continue to exercise is strongly dependent on their intentions to exercise. One might speculate that the need to restructure when to exercise led these students to reassess their satisfaction with regular exercise, thereby modulating their intention to continue to exercise. Efforts such as this that elucidate the factors that regulate habitual behaviors not only have important theoretical implications but also can inform the design of interventions to help people sustain healthy behavioral practices (as well as undermine stable, unhealthy behaviors).

WITH AN EYE TO THE FUTURE: OPPORTUNITIES FOR INNOVATION AND INTEGRATION

The scope of research that examines issues at the intersection of social psychology and health extends considerably beyond the domains examined in the present chapter. For example, we have not attempted to cover the vast literatures on stress and coping (e.g., Martin & Brantley, 2004), psychoneuroimunology (e.g., Marsland, Bachen, Cohen, Rabin, & Manuck, 2002), social relationships and health (e.g., Cohen, 2004), or socioeconomic status and health (e.g., Adler & Snibbe, 2003). In fact, there are important issues in these areas that have implications for understanding the determinants of people's health practices, especially actions that are taken in order to care for a health problem (an area of study we have not emphasized). For example, people's physiological response to their psychological states and to their behavior likely provides critical information that shapes how people think about a health issue and their willingness to change or continue with a behavior.

Despite the volume of research activity concerning people's health practices, a number of investigators have expressed concern that these efforts have not been as productive or innovative as hoped (Noar & Zimmerman, 2005; Rothman, 2004; Sutton, 2004; Weinstein, 2004; Weinstein & Rothman, 2005). Too many findings are constrained by the limitations posed by cross-sectional as well as prospective correlational data and investigators have not adequately pursued opportunities to test, refine, and if need be refute theories of health behavior. In this final section, we consider these challenges with an eye toward several issues that we believe should be kept in mind as investigators pursue the next generation of research activity on the social psychological determinants of health and health behavior.

Expanding the Use of Experimental Methods

Although experimental methods have been used extensively to examine how people process health information, research on the determinants of health behavior has relied primarily on nonexperimental data to inform our understanding of the processes that regulate people's behavioral decisions. In several recent papers, both Sutton (2002, 2004) and Weinstein (2004) have thoughtfully reviewed how an overreliance on nonexperimental data constrains the conclusions that investigators are able to draw. Although nonexperimental data afford opportunities to make inferences regarding factors that predict behavior, conclusions regarding causal mechanisms rest on investigators being able to satisfy a strong

set of assumptions (e.g., that a regression coefficient for a predictor is an unbiased estimate of its causal effect on an outcome; see Sutton, 2004, for a thorough discussion of these issues). A greater reliance on experimental methods is needed to complement the nonexperimental research activities currently under way (Salovey & Williams-Piehota, 2004; Sutton, 2004; Weinstein, 2004). The observed impact of an experimentally manipulated variable can provide critical, converging evidence for the impact of predictors observed in correlational data sets. Thus, theory-based interventions provide an invaluable opportunity to clarify the strengths and weakness of a proposed theoretical model (Rothman, 2004).

Although we strongly endorse a greater reliance on experimental manipulation, in doing so investigators need to remain mindful of several issues. First, an experimentally induced belief (e.g., a smoker is led to hold an optimistic view of smoking cessation) may not prove as influential as a naturally occurring belief (e.g., a smoker who enrolls in a program with an optimistic view of smoking cessation). In our own work, we have observed that naturally occurring variation in people's expectations about the benefits of weight loss had a more pronounced effect on people's weight loss practices than did experimentally induced variation in expectations (Finch et al., 2005). Williams and colleagues have similarly observed stronger effects for measured as opposed to manipulated levels of autonomous motivation (Williams et al., 2002, 2004). One explanation for these findings is that it is difficult to induce a set of beliefs that are sufficiently strong and stable to motivate behavior (we elaborate on this point later). Alternatively, these findings may indicate that investigators need to better specify the defining features of the construct hypothesized to motivate behavior change. For example, the observed variation in people's expectations about weight loss may merely be a marker for the critical underlying factor or it may be that a facet of a belief's underlying structure—such as its degree of accessibility—could determine whether it affects behavior.

Second, investigators may find that experimentally manipulating how people think and reason about their health or their health behavior can prove to be a challenge when operating outside the controlled environment of the laboratory. Shifting the relative salience of beliefs about an attitude object (e.g., heightening the salience of the benefits of cessation) may prove to have a greater impact on participants' beliefs when implemented in a laboratory setting where participants' have limited experience or involvement with the issue at hand than when implemented as part of a smoking cessation program where participants have extensive experience and are quite involved with the issue at hand. Knowing that an intervention strategy can work may not necessarily mean that it does work in more complex settings (cf. Mook, 1983). This is not to suggest that such methods cannot be of value but, rather, that their effectiveness may prove to be contingent upon factors such as people's prior experience with the issue. For example, we observed that an intervention designed to focus people's attention solely on the benefits of cessation made smokers more optimistic about cessation, but only if they had previously had some prior success quitting. For those smokers who had never been able to quit previously, the optimistic intervention actually made them more pessimistic in their outlook. In fact, these latter smokers were more optimistic when they were instructed to focus their attention on both the benefits and the costs of cessation (Hertel et al., 2005).

Finally, as presently formulated, nearly all models of behavioral decision making provide limited guidance regarding the processes that shape the factors that are thought to predict behavior change (Rothman, 2004; Sutton, 2004). For example, we proposed that satisfaction with the outcomes afforded by a pattern of behavior is a critical determinant of behavioral maintenance, but we have offered minimal guidance as to what can be done to heighten the satisfaction people derive from changes in their behavior (Rothman et al., in press). The failure to address this issue is a frequent complaint that interventionists have regarding theories of health behavior (Jeffery, 2004). In fact, given these practical needs, it is not surprising that interventionists are more likely to rely on theories that provide guidance as to how to construct an intervention protocol (e.g., social cognitive theory; Bandura, 1997).

Theory Testing in Ecologically Complex Contexts

Most theories of health behavior specify a list of factors that inform people's behavioral decision but offer limited guidance as to the conditions that might moderate the applicability of these factors (cf. McGuire, 2004). For example, from the outset the theory of reasoned action (Ajzen & Fishbein, 1980) explicitly stated that the degree to which people's intentions rest on their attitudes and/or their perceived social norms will vary across setting and from domain to domain. It is striking that more than 25 years later, researchers have yet to refine further our understanding of the parameters that regulate the impact of attitudes and subjective norms on intentions. More generally, there has yet to be a systematic effort to clarify how features of health behaviors and the setting in which they are performed moderate the applicability of principles that are derived from our theories. Can we delineate how features of a behavior, or the setting in which it is most likely elicited, inhibit or facilitate the impact of particular antecedents or decision rules? For example, do people have a particularly difficult time maintaining behaviors that at the outset are normatively characterized with a high degree of optimism (e.g., weight loss behaviors) (King, Rothman, & Jeffery, 2002)? The proposition by Gibbons and Gerrard (Gibbons et al., 2003) that intentions may not guide the performance of adolescent risk behaviors offers an excellent example of how specifying an aspect of the behavioral domain can improve investigators ability to identify the determinants of behavior.

To specify the contexts in which a theoretical principle may operate, investigators need a much more comprehensive understanding of people's experiences engaging in different health practices. Such work will not only pro-

vide information about how people's experiences of the rewards and costs of the behavior shift over time (e.g., Jeffery et al., 2004) but also advance our understanding of the time lag between when people come to a decision about an issue (e.g., a person decides they are at high risk of getting the flu) and when they take action (e.g., get a flu shot) (Sutton, 2004). Although our proposition is consistent with calls for more descriptive research (e.g., Rozin, 2002), we believe it is critical that investigators attend to how gathering this descriptive information can advance our ability to apply, test, and refine our theoretical models. The challenge will be to identify a core set of parameters that can be used to organize these descriptions.

Because people's health is regulated by ongoing behavioral decision making, time is a critical issue in the study of health behavior. Although investigators may be interested in whether a particular communication strategy can promote the use of a screening behavior such as mammography, it is important that investigators not lose sight of the fact that the final outcome assessed in the study (e.g., rates of mammography utilization) is but a step in a sequence of behavioral decisions. For example, several years ago we demonstrated that providing women with a message that affirmed it was their responsibility to take care of their health increased the likelihood that they would get a mammogram (Rothman, Salovey, Turvey, & Fishkin, 1993). Although encouraging women to take personal responsibility for their health increased screening, the downstream consequences of this message was never considered. In particular, could heightening a woman's sense of personal responsibility for having a mammogram affect her reaction to learning that the mammogram has found an abnormality that might be cancer? Would it cause her to feel more responsible for her situation and more distressed about the possibility of cancer? The overall value of this communication strategy clearly depends on the answer to this question. Thus, in conducting our work we need to be mindful of the broader context in which our study operates. In particular, we need to develop a better understanding of people's psychological reactions to the consequences of their health behaviors (e.g., Andrykowski, Boerner, Salsman, & Pavlik, 2004; Gibbons, Eggleston, & Benthin, 1997) and the longer-term psychological and behavioral consequences of these reactions.

Stimulating Innovation

We believe that the opportunities afforded by using social psychological theory and the method to explore important health decisions can stimulate significant advances in theory construction and methods development, while stimulating innovations in methods to promote health and prevent illness. In many ways, the programs of research described in this chapter represent wonderful examples of the Lewinian ideal of moving between theory and practice in social psychology. However, as research and practice in this area continue, it is critical that investigators continue to refine, revise, and, when needed, reject the theoretical principles derived from our models of health behavior. We need to determine whether new perspectives are needed as investigators grapple with issues such as the nonlinear combination of variables, the role of "irrational" processes in decision making, and the scope of conditions that limit the applicability of general principles. However, the traditional tools of social psychology—and the new emphasis on translational research within the various funding agencies supporting health research—should stimulate just this kind of effort. There has never been a better time to reassert the relevance of theory, experimental research, and the addressing of important social problems, the three-legged stool on which our field was built.

NOTES

1. How investigators operationalize risk perceptions is another critical issue. Estimates of the likelihood of an event may not be sufficient indicators of whether an event is something they want to avoid (Slovic, 1987). Reports of how worried or concerned people are about a potential outcome may reveal the meaning people ascribe to a given level of risk and thus might serve as a better predictor of behavior (e.g., McCaul, Schroeder, & Reid, 1996).
2. It would appear that the label "fear appeals" is a bit misleading as there has been minimal evidence that the beneficial effects of these messages are due to increase in fear per se.
3. Several studies have observed that when intentions are not stable, prior behavior proves to be a better predictor of future behavior (e.g., Conner et al., 2002). Questions concerning the processes by which prior behavior influences future behavior are considered later in this chapter.
4. Messages can also be matched to information about group-level characteristics (e.g., why women resist exercising), an approach that has been labeled "message targeting."

REFERENCES

Adler, N. E., & Snibbe, A. C. (2003). The role of psychosocial processes in explaining the gradient between socioeconomic status and health. *Current Directions in Psychological Science*, *12*, 119–123.

Affleck, G., Tennen, H., Zautra, A., Urrows, S., Abeles, M., & Karoly, P. (2001). Women's pursuit of personal goals in daily life with fibromyalgia: A value-expectancy analysis. *Journal of Counseling and Clinical Psychology*, *69*, 587–596.

Aiken, L. S., West, S. G., Woodward, C. K., & Reno, R. R. (1994). Health beliefs and compliance with mammography screening recommendations in asymptomatic women. *Health Psychology*, *13*, 122–129.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Ajzen, I. (2002). Residual effects of past on later behavior: Habitua-

tion and reasoned action perspectives. *Personality and Social Psychology Review*, *6*, 107–122.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Albaracin, D., Cohen, J. B., & Tumkale, G. T. (2003). When communications collide with recipients' actions: Effects of post message behavior on intentions to follow the message. *Personality and Social Psychology Bulletin*, *29*, 834–845.

Albaracin, D., Johnson, B. T., Fishbein, M., & Muellerleile, P. A. (2001). Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychological Bulletin*, *127*, 142–161.

Andrykowski, M. A., Boerner, L. M., Salsman, J. M., & Pavlik, E. (2004). Psychological response to test results in an ovarian cancer screening program: A prospective longitudinal study. *Health Psychology*, *23*, 622–630.

Apanovitch, A. M., McCarthy, D., & Salovey, P. (2003). Using message framing to motivate HIV testing among low-income, ethnic minority women. *Health Psychology*, *22*, 60–67.

Armitage, C. J. (2004). Evidence that implementation intentions reduce dietary fat intake: A randomized trial. *Health Psychology*, *23*, 319–323.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*, 471–499.

Armor, D. A., & Taylor, S. E. (1998). Situated optimism: Specific outcome expectancies and self-regulation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 309–379). New York: Academic Press.

Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, *22*, 1–32.

Aspinwall, L. G., & Brunhart, S. M. (1996). Distinguishing optimism from denial: Optimistic beliefs predict attention to health threats. *Personality and Social Psychology Bulletin*, *22*, 993–1003.

Aspinwall, L. G., & Taylor, S. E. (1993). Effects of social comparison direction, threat, and self-esteem on affect, self-evaluation, and expected success. *Journal of Personality and Social Psychology Bulletin*, *64*, 708–722.

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions, and behaviour. *Social Psychology Quarterly*, *55*, 178–204.

Baldwin, A. S., Rothman, A. J., Hertel, A. W., Linde, J. A., Jeffery, R. W., Finch, E. A., et al. (2006). Specifying the determinants of behavior change initiation and maintenance: An examination of self-efficacy, satisfaction, and smoking cessation. *Health Psychology*, *25*, 626–634.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Banks, S., Salovey, P., Greener, S., Rothman, A., Moyer, A., Beauvais, J., et al. (1995). The effects of message framing on mammography utilization. *Health Psychology*, *14*, 178–184.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego, CA: Academic Press.

Black, W. C., Nease, R. F., & Tosteson, A. N. A. (1995). Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. *Journal of the National Cancer Institute*, *87*, 720–731.

Blanton, H., & Christie, C. (2003). Deviance regulation: A theory of action and identity. *Review of General Psychology*, *7*, 115–149.

Blanton, H., Stuart, A. E., & VandenEijnden, R. J. J. M. (2001). An introduction to deviance regulation theory: The effect of behavioral norms on message framing. *Personality and Social Psychology Bulletin*, *27*, 848–858.

Bottorff, J. L., Ratner, P. A., Johnson, J. L., Lovato, C. Y., & Joab, S. A. (1998). Communicating cancer risk information: The challenges of uncertainty. *Patient Education and Counseling*, *33*, 67–81.

Brewer, N., Weinstein, N. D., Cuite, C. L., Herrington, J., & Hayes, N. (2004). Measuring risk perception and its relation to risk behavior. *Annals of Behavioral Medicine*, *27*, 125–130.

Brownell, K. D., & Horgen, K. B. (2004). *Food fight: The inside story of the food industry, America's obesity crisis, and what we can do about it*. New York: McGraw-Hill.

Brug, J., Glanz, K., van Assema, P., Kok, G., & van Breukelen, G. J. P. (1998). The impact of computer-tailored feedback and iterative feedback on fat, fruit, and vegetable intake. *Health Education and Behavior*, *25*, 517–531.

Brug, J., Steenhuis, I., van Assema, P., & de Vries, H. (1996). The impact of a computer-tailored nutrition intervention. *Preventative Medicine*, *25*, 236–242.

Bryan, A., Fisher, J. D., & Fisher, W. A. (2002). Tests of the mediational role of preparatory safer sexual behavior in the context of the theory of planned behavior. *Health Psychology*, *21*, 71–80.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Cameron, L. D., & Leventhal, H. (2003). *The self-regulation of health and illness behavior*. London: Routledge.

Campbell, J. D., & Fairey, P. J. (1985). Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. *Journal of Personality and Social Psychology*, *48*, 1097–1111.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right." *Journal of Personality and Social Psychology*, *86*, 388–404.

Chambers, J. R., & Windschitl, P. (2004). Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, *130*, 813–838.

Cohen, S. (2004). Social relationships and health. *American Psychologist*, *59*, 676–684.

Conner, M., & Norman, P. (Eds.). (1996). *Predicting health behaviour: Research and practice with social cognition models*. Buckingham, UK: Open University Press.

Conner, M., & Norman, P. (Eds.). (2005). *Predicting health behaviour: Research and practice with social cognition models* (2nd ed.). Buckingham, UK: Open University Press.

Conner, M., Norman, P., & Bell, R. (2002). The theory of planned behavior and healthy eating. *Health Psychology*, *21*, 194–201.

Conner, M., Sheeran, P., Norman, P., & Armitage, C. J. (2000). Temporal stability as a moderator of relationship in the theory of planned behavior. *British Journal of Social Psychology*, *39*, 469–493.

Cooke, R., & Sheeran, P. (2004). Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology*, *43*, 159–186.

Crandall, C. (1988). Social contagion of binge eating. *Journal of Personality and Social Psychology*, *55*, 588–598.

Croyle, R. T., & Ditto, P. H. (1990). Illness cognition and behavior: An experimental approach. *Journal of Behavioral Medicine*, *13*, 31–52.

Croyle, R. T., & Hunt, J. R. (1991). Coping with health threat: Social influence processes in reactions to medical test results. *Journal of Personality and Social Psychology*, *60*, 382–389.

Croyle, R. T., Sun, Y. C., & Louie, D. H. (1993). Psychological minimization of cholesterol test results: Moderators of appraisal in college students and community residents. *Health Psychology*, *12*, 503–507.

Cumming, K. M., Giovino, G., Bansal, M. A., Hyland, A., Hastrip, J. & Yost, B. (2001). *Are smokers misinformed about the relative risks and benefits of different nicotine delivery products*. Paper presented at the annual grantee meeting of the Robert Wood Johnson Foundation Substance Abuse Policy Research Program, St. Augustine, FL.

Das, E. H., de Wit, J. F., & Stroebe, W. (2003). Fear appeals moti-

vate acceptance of action recommendations: Evidence for a positive bias in the processing of persuasive messages. *Personality and Social Psychology Bulletin, 29*, 650–664.

de Hoog, N., Stroebe, W., & de Wit, J. F. (2005). The impact of fear appeals on processing and acceptance of health messages. *Personality and Social Psychology Bulletin, 31*, 24–33.

Denes-Raj, V., Epstein, S., & Cole, J. (1995). The generality of the ratio-bias phenomenon. *Personality and Social Psychology Bulletin, 21*, 1083–1092.

Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sun screen use: Gain-framed messages motivate beach-goers. *Health Psychology, 18*, 189–196.

Dijkstra, A. (2005). Working mechanisms of computer-tailored health education: Evidence from smoking cessation. *Health Education Research, 20*, 527–539.

Dijkstra, A., & Borland, R. (2003). Residual outcome expectations and relapse in ex-smokers. *Health Psychology, 22*, 340–346.

Dijkstra, A., De Vries, H., & Roijackers, J. (1998a). Computerized tailored feedback to change cognitive determinants of smoking: A Dutch field experiment. *Health Education Research, 13*, 197–206.

Dijkstra, A., De Vries, H., & Roijackers, J. (1998b). Long-term effectiveness of computer generated tailored feedback in smoking cessation. *Health Education Research, 13*, 207–214.

Dijkstra, A., De Vries, H., Roijackers, J., & van Breukelen, G. (1998). Tailored interventions to communicate stage-matched information to smokers in different motivational stages. *Journal of Consulting and Clinical Psychology, 66*, 549–557.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63*, 568–584.

Ditto, P. H., Munro, G. D., Apanovitch, A. M., Scepansky, J. A., & Lockhart, L. K. (2003). Spontaneous skepticism: The interplay of motivation and expectation in responses to favorable and unfavorable medical diagnoses. *Personality and Social Psychology Bulletin, 29*, 1120–1132.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference inconsistent information. *Journal of Personality and Social Psychology, 75*, 53–69.

Dutta-Bergman, M. J. (2004). Primary sources of health information: Comparisons in the domains of health attitudes, health cognitions, and health behaviors. *Health Communication, 16*, 273–288.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin, 125*, 64–89.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232.

Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 76*, 628–644.

Evers, K. E., Bishop, C. H., Gerhan, L. S., & Weisse, C. S. (1997). AIDS educator effectiveness as a function of sexual orientation and HIV status. *Journal of Applied Social Psychology, 27*, 902–914.

Ferguson, E., & Bibby, P. A. (2002). Predicting future blood donor returns: Past behavior, intentions, and observer effects. *Health Psychology, 21*, 513–518.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.

Finch, E. A., Linde, J. A., Jeffery, R. W., Rothman, A. J., King, C. M., & Levy, R. L. (2005). The effects of outcome expectations and satisfaction on weight loss and maintenance: Correlational and experimental analyses. *Health Psychology, 24*, 608–616.

Foster, G. D., Wadden, T. A., Vogt, R. A., & Brewer, G. (1997). What is a reasonable weight loss?: Patients' expectations and evaluations of obesity treatment outcomes. *Journal of Consulting and Clinical Psychology, 65*, 79–85.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218–226.

Freitas, A. L., & Higgins, E. T. (2002). Enjoying goal-directed action: The role of regulatory fit. *Psychological Science, 13*, 1–6.

French, S. A., Jeffery, R. W., Story, M., Breitlow, K. K., Baxter, J. S., Hannan, P., et al. (2001). Pricing and promotion effects on low-fat vending snack purchases: The CHIPS study. *American Journal of Public Health, 91*, 112–117.

Gerrard, M., Gibbons, F. X., Benthin, A. C., & Hessling, R. M. (1996). A longitudinal study of the reciprocal nature of risk behaviors and cognitions in adolescents: What you do shapes what you think, and vice versa. *Health Psychology, 16*, 344–354.

Gerrard, M., Gibbons, F. X., & Bushman, B. J. (1996). Relation between perceived vulnerability to HIV and precautionary sexual behavior. *Psychological Bulletin, 119*, 390–409.

Gibbons, F. X., & Eggleston, T. J. (1996). Smoker networks and the "typical smoker": A prospective analysis of smoking cessation. *Health Psychology, 15*, 469–477.

Gibbons, F. X., Eggleston, T. J., & Benthin, A. C. (1997). Cognitive reactions to smoking relapse: The reciprocal relation between dissonance and self-esteem. *Journal of Personality and Social Psychology, 72*, 184–195.

Gibbons, F. X., & Gerrard, M. (1995). Predicting young adults' health risk behavior. *Journal of Personality and Social Psychology, 69*, 505–517.

Gibbons, F. X., Gerrard, M., Blanton, H., & Russell, D. W. (1998). Reasoned action and social reaction: Willingness and intention as independent predictors of health risk. *Journal of Personality and Social Psychology, 74*, 1164–1180.

Gibbons, F. X., Gerrard, M., & Lane, D. J. (2003). A social reaction model of adolescent health risk. In J. M. Suls & K. Wallston (Eds.), *Social psychological foundations of health and illness* (pp. 107–136). Oxford, UK: Blackwell.

Gibbons, F. X., Gerrard, M., Pomery, E. A., & Reimer, R. A. (in press). Health decision making: Reasoned vs. reactive processes. In D. de Ridder & J. de Wit (Eds.), *Self-regulation of health behavior*. Hoboken, NJ: Wiley.

Gibbons, F. X., Gerrard, M., Vande Lune, L. S., Wills, T. A., Brody, G., & Conger, R. D. (2004). Context and cognitions: Environmental risk, social influence, an adolescent substance use. *Personality and Social Psychology Bulletin, 30*, 1048–1061.

Gibson, B., & Sanbonmatsu, D. M. (2004). Optimism, pessimism, and gambling: The downside of optimism. *Personality and Social Psychology Bulletin, 30*, 149–160.

Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognitions and motivation to behavior* (pp.287–312). New York: Guilford Press.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*, 493–503.

Gollwitzer, P. M., Bayer, U. C., & McCulloch, K. C. (2005). The control of the unwanted. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 485–515). New York: Oxford University Press.

Gump, B. B., & Kulik, J. A. (1995). The effect of a model's HIV status of self-perceptions: A self-protective similarity bias. *Personality and Social Psychology Bulletin, 21*, 827–833.

Hagger, M. S., & Orbell, S. (2003). A meta-analytic review of the common-sense model of illness representations. *Psychology and Health, 18*, 141–184.

Helgeson, V. S. (2003). Cognitive adaptation, psychological adjustment, and disease progression among angioplasty patients: 4 years later. *Health Psychology, 22*, 30–38.

Hertel, A. W., Finch, E., Kelly, K., King, C., Lando, H., Linde, J., et al. (2005). *The impact of outcome expectations and satisfaction on the initiation and maintenance of smoking cessation: An experimental test*. Unpublished manuscript, University of Minnesota, Minneapolis.

Higgins, E. T. (1999). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 503–526). New York: Guilford Press.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*, 1217–1230.

Hunt, J. S., McGorty, E. K., Iyer, S. N., & Rothman, A. J. (2005). *The influence of patient sex, race, and socioeconomic status on judgments about symptoms related to anorexia nervosa*. Unpublished manuscript, University of Nebraska, Lincoln.

Hunt, J. S., & Rothman, A. J. (2004). *What does eating light mean? How illness schemas influence health judgments*. Unpublished manuscript, University of Nebraska, Lincoln.

Hunt, J. S., Rothmann, T. L., Rothman, A. J., Iyer, S. N., & McGorty, E. K. (in press). Implicit and explicit associations between social groups and health problems. *Psychology and Health*.

Hunt, W. A., Barnett, L. W., & Branch, L. G. (1971). Relapse rates in addiction programs. *Journal of Clinical Psychology*, *27*, 455–456.

Janis, I. L. (1967). Effects of fear arousal on attitude change: Recent developments in theory and experimental research. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 3, pp. 166–224). New York: Academic Press.

Jeffery, R. W. (2004). How can health behavior theory be made more useful for intervention research? *International Journal of Behavioral Nutrition and Physical Activity*, *1*, 10.

Jeffery, R. W., Drewnowski, A., Epstein, L. H., Stunkard, A. J., Wilson, G. T., Wing, R. R., et al. (2000). Long-term maintenance of weight loss: Current status. *Health Psychology*, *19*, 5–16.

Jeffery, R. W., Kelly, K. M., Rothman, A. J., Sherwood, N. E., & Boutelle, K. N. (2004). The weight loss experience: A descriptive analysis. *Annals of Behavioral Medicine*, *27*, 100–106.

Jemmott, J. B., III, Ditto, P. H., & Croyle, R. T. (1986). Judging health status: Effects of perceived prevalence and personal relevance. *Journal of Personality and Social Psychology*, *50*, 899–905.

Johnston, D. W., Johnston, M., Pollard, B., Kinmonth, A. L., & Mant, D. (2004). Motivation is not enough: Prediction of risk behavior following diagnosis of coronary heart disease from the theory of planned behavior. *Health Psychology*, *23*, 533–538.

Jones, J. L., & Leary, M. R. (1994). Effects of appearance-based admonitions again sun exposure on tanning intentions in young adults. *Health Psychology*, *13*, 86–90.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge, UK: Cambridge University Press.

King, C. M., Rothman, A. J., & Jeffery, R. W. (2002). The challenge study: Theory based interventions for smoking and weight loss. *Health Education Research*, *17*, 522–530.

Kiviniemi, M. T., & Rothman, A. J. (2006). What am I supposed to do about my health? Selective memory biases in individuals' memory for health-related information and behavior recommendations. *Psychology and Health*, *21*, 247–272.

Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, *72*, 763–774.

Klein, W. M. (2005, January). *Self-affirmation and the processing of health information*. Paper presented at the annual convention of the Society for Personality and Social Psychology, New Orleans, LA.

Klein, W. M., Blier, X., & Janz, X. (2001). Maintaining positive self-evaluations: Reducing attention to diagnostic but unfavorable social comparison information when general self-regard is salient. *Motivation and Emotion*, *25*, 23–40.

Klein, W. M., & Weinstein, N. D. (1997). Social comparison and unrealistic optimism about personal risk. In B. P. Buunk & F. X. Gibbons (Eds.), *Health, coping, and well-being: Perspectives from social comparison theory* (pp. 25–61). Mahwah, NJ: Erlbaum.

Kreuter, M. K., Bull, F. C., Clark, E. M., & Oswald, D. L. (1999). Understanding how people process health information: A comparison of tailored and non-tailored weight-loss materials. *Health Psychology*, *18*, 487–494.

Kreuter, M. K., Oswald, D. L., Bull, F. C., & Clark, E. M. (2000). Are tailored health education materials always more effective than non-tailored materials? *Health Education Research*, *15*, 305–315.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*, 636–647.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.

Lau, R. R., & Hartman, K. A. (1983). Common sense representations of common illnesses. *Health Psychology*, *2*, 167–185.

Lauver, D., & Rubin, M. (1990). Message framing, dispositional optimism, and follow-up for abnormal papanicolau tests. *Research in Nursing and Health*, *13*, 199–207.

Leary, M. R., Tchividjian, L. R., & Kraxberger, B. E. (1994). Self-presentation can be hazardous to your health: Impression management and health risk. *Health Psychology*, *13*, 461–470.

Lee, A. Y., & Aaker, J. L. (2004). Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion. *Journal of Personality and Social Psychology*, *86*, 205–218.

Lerman, C., Ross, E., Boyce, A., Gorchov, P., McLaughlin, R., Rimer, B., et al. (1992). The impact of mailing psychoeducational materials to women with abnormal mammograms. *American Journal of Public Health*, *82*, 729–730.

Leventhal, H. (1970). Findings and theory in the study of fear communications. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 5, pp. 119–186). New York: Academic Press.

Leventhal, H., Nerenz, D. R., & Steele, D. J. (1984). Illness representations and coping with health threats. In A. Baum, S. E. Taylor, & J. Singer (Eds.), *Handbook of psychology and health* (Vol. 4, pp. 219–252). Hillsdale, NJ: Erlbaum.

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers*. New York: Harper & Row.

Liberman, A., & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin*, *18*, 669–679.

Liberman, N., & Trope, Y. (1998). The role feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, *75*, 5–18.

Linde, J. A., Rothman, A. J., Baldwin, A. S., & Jeffery, R. W. (2006). The impact of self-efficacy on behavior change and weight change among overweight participants in a weight loss trial. *Health Psychology*, *25*, 282–291.

Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology*, *19*, 469–479.

Mahler, H. I. M., Kulik, J. A., Gibbons, F. X., Gerrard, M., & Harrell, J. (2003). Effects of appearance based interventions on sun protection intentions and self-reported behaviors. *Health Psychology*, *22*, 199–209.

Mann, T., Sherman, D., & Updegraff, J. (2004). Dispositional motivations and message framing: A test of the congruency hypothesis in college students. *Health Psychology*, *23*, 330–334.

Marlatt, G. A., & Gordon, J. R. (1985). *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. New York: Guilford Press.

Marsland, A. L., Bachen, E. A., Cohen, S., Rabin, B., & Manuck, S. B. (2002). Stress, immune reactivity and susceptibility to infectious disease. *Physiology and Behavior*, *7*, 711–716.

Martin, P. D., & Brantley, P. J. (2004). Stress, coping, and social support in health and behavior. In J. M. Raczynski & L. C. Leviton (Eds.), *Handbook of clinical health psychology: Vol. 2. Disorders of behavior and health* (pp. 233–267). Washington, DC: American Psychological Association.

Martin, R., Gordon, E. E. I., & Lounsbury, P. (1998). Gender disparities in attribution of cardiac-related symptoms: Contribution of common sense models of illness. *Health Psychology*, *17*, 346–357.

Martin, R., & Lemos, K. (2002). From heart attacks to melanoma:

Do common sense models of somatization influence symptom interpretation for female victims? *Health Psychology*, *21*, 25–32.

Martin, R., Lemos, K., Rothrock, N., Bellman, S. B., Russell, D., Tripp-Reimer, T., et al. (2004). Gender disparities in common sense models of illness among myocardial infarction victims. *Health Psychology*, *23*, 345–353.

McCaul, K. D., Branstetter, A. D., Schroeder, D. M., & Glasgow, R. E. (1996). What is the relationship between breast cancer risk and mammography screening? A meta-analytic review. *Health Psychology*, *15*, 423–429.

McCaul, K. D., Glasgow, R. E., & O'Neill, H. K. (1992). The problem of creating habits: Establishing health-protective dental behaviors. *Health Psychology*, *11*, 101–110.

McCaul, K. D., Schroeder, D. M., & Reid, P. A. (1996). Breast cancer worry and screening: Some prospective data. *Health Psychology*, *15*, 430–433.

McCaul, K. D., Thiese-Duffy, E., & Wilson, P. (1992). Coping with medical diagnosis: The effects of at-risk versus disease labels over time. *Journal of Applied Social Psychology*, *22*, 1340–1355.

McGuire, W. J. (1991). Using guiding-idea theories of the person to develop educational campaigns against drug abuse and other health-threatening behavior. *Health Education Research*, *6*, 173–184.

McGuire, W. J. (2004). A perspectivist approach to theory construction. *Personality and Social Psychology Review*, *8*, 173–182.

Meyer, D., Leventhal, H., & Guttman, M. (1985). Common-sense models of illness: The example of hypertension. *Health Psychology*, *4*, 115–135.

Meyerowitz, B., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*, 500–510.

Meyerowitz, B. E., Wilson, D. K., & Chaiken, S. (1991, June). *Loss-framed messages increase breast self-examination for women who perceive risk*. Paper presented at the annual convention of the American Psychological Society, Washington, DC.

Miller, G. A. (1970). Psychology as a means of promoting human welfare. *American Psychologist*, *24*, 1063–1075.

Milne, S., Sheeran, P., & Orbell, S. (2000). Prediction and intervention in health-related behavior: A meta-analytic review of protection motivation theory. *Journal of Applied Social Psychology*, *30*, 106–143.

Mokdad, A. H., Bowman, B. A., Ford, E. S., Vinicor, F., Marks, J. S., & Koplan, J. P. (2001). The continuing epidemics of obesity and diabetes in the United States. *Journal of the American Medical Association*, *286*, 1195–1200.

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Marks, J. S., et al. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, *Journal of the American Medical Association*, *289*, 76–79.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, *38*, 379–387.

Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D., & Buick, D. (2002). The revised illness perception questionnaire (IPQ-R). *Psychology and Health*, *17*, 1–16.

Noar, S. M., & Zimmerman, R. S. (2005). Health behavior theory and cumulative knowledge regarding health behaviors: Are we moving in the right direction? *Health Education Research*, *20*, 275–290.

Norman, P., Conner, M., & Bell, R. (1999). The theory of planned behavior and smoking cessation. *Health Psychology*, *18*, 89–94.

Ockene, J. K., Emmons, K. M., Mermelstein, R. J., Perkins, K. A., Bonollo, D. S., Vorhees, C. C., et al. (2000). Relapse and maintenance issues for smoking cessation. *Health Psychology*, *19*, 17–31.

Orbell, S., Perugini, M., & Rakow, T. (2004). Individual differences in sensitivity to health communications: Consideration of future consequences. *Health Psychology*, *23*, 388–396.

Ouellete, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*, 54–74.

Perri, M. G., Nezu, A. M., Patti, E. T., & McCann, K. L. (1989). Effect of length of treatment on weight loss. *Journal of Consulting and Clinical Psychology*, *57*, 450–452.

Petty, R. E., Barden, J., & Wheeler, S. C. (2002). The elaboration likelihood model of persuasion: Health promotions that yield sustained behavioral change. In R. J. DiClemente, R. A. Crosby, & M. C. Kegler (Eds.), *Emerging theories in health promotion practice and research* (pp. 71–99). San Francisco: Jossey-Bass.

Petty, R. E., & Krosnick, J. A. (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 323–390). New York: McGraw-Hill.

Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*, 243–256.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, *47*, 1102–1114.

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, *12*, 38–48.

Rabin, C., Leventhal, H., & Goodin, S. (2004). Conceptualization of disease timeline predicts posttreatment distress in breast cancer patients. *Health Psychology*, *23*, 407–412.

Raghubir, P., & Menon, G. (1998). AIDS and me, never the twain shall meet: The effects of information accessibility on judgments of risk and advertising effectiveness. *Journal of Consumer Research*, *25*, 52–63.

Reed, M. B., & Aspinwall, L. (1998). Self-affirmation reduces biased processing of health-information. *Motivation and Emotion*, *22*, 99–132.

Renner, B. (2004). Biased reasoning: Adaptive responses to health risk feedback. *Personality and Social Psychology Bulletin*, *30*, 384–396.

Rosentstock, I. M., Strecher, V. J., & Becker, M. H. (1988). Social learning theory and the health belief model. *Health Education Quarterly*, *15*, 175–183.

Rothman, A. J. (2000). Toward a theory-based analysis of behavioral maintenance. *Health Psychology*, *19*, 1–6.

Rothman, A. J. (2004). Is there nothing more practical than a good theory?: Why Innovations and advances in health behavior change will arise if interventions are more theory-friendly. *International Journal of Behavioral Nutrition and Physical Activity*, *1*, 11.

Rothman, A. J., Baldwin, A., & Hertel, A. (2004). Self-regulation and behavior change: Disentangling behavioral imitation and behavioral maintenance. In R. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 130–148). New York: Guilford Press.

Rothman, A. J., Bartels, R., Wlaschin, J., & Salovey, P. (2006). The strategic use of framed messages to promote healthy behavior: How theory can inform practice. *Journal of Communication*, *56*, S202–S220.

Rothman, A. J., Haddock, G., & Schwarz, N. (2001). "How many partners is too many?": Shaping perceptions of vulnerability. *Journal of Applied Social Psychology*, *31*, 2195–2214.

Rothman, A. J., Hertel, A. W., Baldwin, A. S., & Bartels, R. (in press). Integrating theory and practice: Understanding the determinants of health behavior change. In J. Shah & W. Gardner (Eds.), *Handbook of motivation science*. New York: Guilford Press.

Rothman, A. J., Kelly, K. M., Hertel, A., & Salovey P. (2003). Message frames and illness representations: Implications for interventions to promote and sustain healthy behavior. In L. D. Cameron & H. Leventhal (Eds.), *The self-regulation of health and illness behavior* (pp. 278–296). London: Routledge.

Rothman, A. J., & Kiviniemi, M. (1999). "Treating people with health information": An analysis and review of approaches to communicating health risk information. *Journal of the National Cancer Institute Monographs*, *25*, 44–51.

Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., &

Salovey, P. (1999). The systematic influence of gain-and loss-framed messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin*, *25*, 1355–1369.

Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, *121*, 3–19.

Rothman, A. J., Salovey, P., Antone, C., Keough, K., & Martin, C. D. (1993). The influence of message framing on intentions to perform health behaviors. *Journal of Experimental Social Psychology*, *29*, 408–433.

Rothman, A. J., Salovey, P., Turvey, C., & Fishkin, S. A. (1993). Attributions of responsibility and persuasion: Increasing mammography utilization among women over forty with an internally-oriented message. *Health Psychology*, *12*, 37–45.

Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality and Social Psychology Bulletin*, *24*, 1053–1064.

Rothman, A. J., Stark, E., & Salovey, P. (2006). Using message framing to promote healthy behavior: A guide to best practices. In J. Trafton (Ed.), *Best practices in the behavioral management of chronic diseases, Volume 3* (pp. 31–48). Los Altos, CA: Institute for Disease Management.

Rozin, P. (2002). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, *5*, 2–14.

Salovey, P., Rothman, A. J., & Rodin, J. (1998). Health behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 633–683). Boston: McGraw-Hill.

Salovey, P., & Williams-Piehota, P. (2004). Field experiments in social psychology: Message framing and the promotion of health protective behaviors. *American Behavioral Scientist*, *47*, 488–505.

Sandman, P. M., Weinstein, N. D., & Miller, P. (1994). High risk or low: How location on a "risk ladder" affects perceived risk. *Risk Analysis*, *14*, 35–45.

Schmeichel, B. J., & Baumeister, R. (2004). Self-regulatory strength. In K. Vohs & R. Baumeister (Eds.), *The handbook of self-regulation* (pp. 84–98). New York: Guilford Press.

Schneider, T., Salovey, P., Apanovitch, A., Pizarro, J., McCarthy, D., Zullo, J., et al. (2001). The effects of message framing and ethnic targeting on mammography use among low-income women. *Health Psychology*, *20*, 256–266.

Schwarz, N., & Strack, F. (1991). Evaluating one's life: A judgmental model of subjective well-being. In F. Strack, M. Argyle, & N. Schwarz, (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 27–47). Oxford, UK: Pergamon Press.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*, 966–972.

Shaklee, H., & Fischhoff, B. (1990). The psychology of contraceptive surprises: Cumulative risk and contraceptive effectiveness. *Journal of Applied Social Psychology*, *20*, 385–403.

Sheeran, P. (2002). Intention–behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, *12*, 1–36.

Sheeran, P., & Abraham, C. (2003). Mediator of moderators: Temporal stability of intention and the intention–behavior relation. *Personality and Social Psychology Bulletin*, *29*, 205–215.

Sheeran, P., Conner, M., & Norman, P. (2001). Can the theory of planned behavior explain patterns of health behavior change? *Health Psychology*, *20*, 12–19.

Sheeran, P., & Orbell, S. (1999). Implementation intentions and repeated behavior: Augmenting the predictive validity of the theory of planned behavior. *European Journal of Social Psychology*, *29*, 349–369.

Sheeran, P., & Orbell, S. (2000). Using implementation intentions to increase attendance for cervical cancer screening. *Health Psychology*, *19*, 283–289.

Sheeran, P., Orbell, S., & Trafimow, D. (1999). Does the temporal stability of behavioral intentions moderate intention–behavior and past behavior–future behavior relations? *Personality and Social Psychology Bulletin*, *25*, 721–730.

Sheeran, P., Webb, T. L., & Gollwitzer, P. M. (2005). The interplay between goal intentions and implementation intentions. *Personality and Social Psychology Bulletin*, *31*, 87–98.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, *26*, 1046–1058.

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, *11*, 118–127.

Shiffman, S., Balabanis, M. H., Paty, J. A., Engberg, J., Gwaltney, C. J., Liu, K. S., et al. (2000). Dynamic effects of self-efficacy on smoking lapse and relapse. *Health Psychology*, *19*, 315–323.

Shiffman, S., Hickcox, M., Paty, J. A., Gnys, M., Kassel, J. D., & Richards, T. J. (1996). Progression from a smoking lapse to relapse: Prediction from abstinence violation effects, nicotine dependence, and lapse characteristics. *Journal of Consulting and Clinical Psychology*, *64*, 993–1002.

Skinner, C. S., Campbell, M. K., Rimer, B. K., Curry, S., & Prochaska, J. O. (1999). How effective is tailored print communication? *Annals of Behavioral Medicine*, *21*, 290–298.

Skinner, C. S., Strecher, V. J., & Hospers, H. (1994). Physicians' recommendations for mammography: Do tailored messages make a difference? *American Journal of Public Health*, *84*, 43–49.

Slovic, P. (1987). Perception of risk. *Science*, *236*, 280–285.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 463–489). Cambridge, UK: Cambridge University Press.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care*. Washington, DC: National Academies Press.

Spiegel, S., Grant-Pillow, H., & Higgins, E. T. (2004). How regulatory fit enhances motivational strength during goal pursuit. *European Journal of Social Psychology*, *34*, 39–54.

Stokols, D., Grzywacz, J. G., McMahan, S., & Phillips, K. (2003). Increasing the health promotive capacity of human environments. *American Journal of Health Promotion*, *18*, 4–13.

Suls, J., & Green, P. (2003). Pluralistic ignorance and college student perceptions of gender-specific alcohol norms. *Health Psychology*, *22*, 479–486.

Suls, J., Wan, C. K., & Sanders, G. (1988). False consensus and false uniqueness in estimating the prevalence of health-protective behaviors. *Journal of Applied Social Psychology*, *18*, 66–79.

Sutton, S. (1998). Predicting and explaining intentions and behaviour: How well are we doing? *Journal of Applied Social Psychology*, *28*, 1317–1338.

Sutton, S. (2002a). Testing attitude–behaviour theories using non-experimental data: An examinations of some hidden assumptions. *European Review of Social Psychology*, *13*, 293–323.

Sutton, S. (2002b). Using social cognition models to develop health behaviour interventions: Problems and assumptions. In D. Rutter & L. Quine (Eds.), *Changing health behaviour: Intervention and research with social cognition models* (pp. 193–208). Buckingham, UK: Open University Press.

Sutton, S. (2004). Determinants of health-related behaviours: Theoretical and methodological issues. In S. R. Sutton, A. Baum, & M. Johnston (Eds.), *The Sage handbook of health psychology* (pp. 94–126). London: Sage.

Szymanski, D. M., & Henard, D. H. (2001). Customer satisfaction: A meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science*, *29*, 16–35.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*, 67–85.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social

psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*, 569–575.

Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resistance to persuasion on attitude certainty. *Journal of Personality and Social Psychology*, *83*, 1298–1313.

Tormala, Z. L., & Petty, R. E. (2004). Resistance to persuasion and attitude certainty: The moderating role of elaboration. *Personality and Social Psychology Bulletin*, *30*, 1446–1457.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–421.

Trope, Y., & Neter, E. (1994). Reconciling competing motives in self-evaluation: The role of self-control in feedback seeking. *Journal of Personality and Social Psychology*, *66*, 646–657.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the rationality of choice. *Science*, *221*, 453–458.

van Ryn, M., & Fu, S. S. (2003). Paved with good intentions: Do public health and human service providers contribute to racial/ethnic disparities in health? *American Journal of Public Health*, *93*, 248–255.

Wadden, T. A., Vogt, R. A., Andersen, R. E., & Barlett, S. J. (1997). Exercise in the treatment of obesity: Effects of four interventions on body composition, resting energy expenditure, appetite, and mood. *Journal of Consulting and Clinical Psychology*, *65*, 269–277.

Waters, A. J., Shiffman, S., Sayette, M., Paty, J. A., Gwaltney, C. J., & Balabanis, M. H. (2003). Attentional bias predicts outcome in smoking cessation. *Health Psychology*, *22*, 378–387.

Weinstein, N. D. (1988). The precaution adoption process. *Health Psychology*, *7*, 355–386.

Weinstein, N. D. (1989). Effects of personal experience on self-protective behavior. *Psychological Bulletin*, *105*, 31–50.

Weinstein, N. D. (1993). Testing four competing theories of health-protective behavior. *Health Psychology*, *12*, 324–333.

Weinstein, N. D. (1998). Accuracy of smokers' risk perceptions. *Annals of Behavioral Medicine*, *20*, 1–8.

Weinstein, N. D. (2004). *Misleading tests of health behavior theories*. Unpublished manuscript, Rutgers University, New Brunswick, NJ.

Weinstein, N. D., Lyon, J. E., Sandman, P. M., & Cuite, C. L. (1998). Experimental evidence for stages of health behavior change: The precaution adoption process model applied to home radon testing. *Health Psychology*, *17*, 445–453.

Weinstein, N. D., & Rothman, A. J. (2005). Revitalizing research on health behavior theories. *Health Education Research*, *20*, 294–297.

Weinstein, N. D., Rothman, A. J., & Nicolich, M. (1998). Using correlational data to examine the effects of risk perceptions on precautionary behavior. *Psychology and Health*, *13*, 479–501.

Weinstein, N. D., Rothman, A. J., & Sutton, S. R. (1998). Stage theories of health behavior. *Health Psychology*, *17*, 290–299.

Williams, G. C., Gagne, M., Ryan, R. M., & Deci, E. L. (2002). Facilitating Autonomous motivation for smoking cessation. *Health Psychology*, *21*, 40–50.

Williams, G. C., McGregor, H. A., Zeldman, A., Freedman, Z. R., & Deci, E. L. (2004). Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. *Health Psychology*, *23*, 58–66.

Williams, G. C., Ryan, R. M., Rodin, G. C., Grolnick, W. S., & Deci, E. L. (1998). Autonomous regulation and long-term medication adherence in adult outpatients. *Health Psychology*, *17*, 269–276.

Williams-Piehota, P., Pizzaro, J., Navarro, S., Mowad, L., & Salovey, P. (2006). The impact of messages tailored to need for cognition on increasing fruit and vegetable intake among callers to the cancer information service. *Health Communication*, *19*, 75–84.

Williams-Piehota, P., Pizzaro, J., Schneider, T. R., Mowad, L., & Salovey, P. (2005). Matching health messages to monitor-blunter coping styles to motivate screening mammography. *Health Psychology*, *24*, 58–67.

Williams-Piehota, P., Schneider, T. R., Pizarro, J., Mowad, L., & Salovey, P. (2004). Matching health messages to health locus of control beliefs for promoting mammography utilization. *Psychology and Health*, *19*, 407–423.

Windschitl, P. D., & Weber, E. U. (1999). The interpretation of "likely" depends on the context, but "70%" is 70%—Right?: The influence of associative processes on perceived certainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1514–1533.

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied*, *2*, 343–364.

Wing, R. R. (2002). Behavioral weight control. In T. A. Wadden & A. J. Stunkard (Eds.), *Handbook of obesity treatment* (pp. 301–316). New York: Guilford Press.

Witkiewitz, K., & Marlatt, G. A. (2004). Relapse prevention for alcohol and drug problems: That was zen, this is tao. *American Psychologist*, *59*, 224–235.

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education and Behavior*, *27*, 591–615.

Wood, J. V., Taylor, S. E., & Lichtman, R. R. (1985). Social comparison in adjustment to breast cancer. *Journal of Personality and Social Psychology*, *49*, 1169–1183.

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*, 1281–1297.

Wood, W., Tam, L., & Guerrero-Witt, M. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, *88*, 918–933.

Worth, K., Sullivan, H., Hertel, A. W., Jeffery, R. W., & Rothman, A. J. (2005). Are there times when avoidance goals can be beneficial?: A look at smoking cessation. *Basic and Applied Social Psychology*, *27*, 107–116.

CHAPTER 37

Contributions of Social Psychology to Clinical Psychology

Three Views of a Research Frontier

Timothy J. Strauman
Philip R. Costanzo
Neil P. Jones
Amy Noll McLean
Kari A. Merrill

The vast expanse of conceptual territory adjoining the disciplines of clinical psychology and social psychology remains a largely uncharted world. Although previous edited volumes have provided important summaries of the social/clinical interface (e.g., Leary & Kowalski, 1999; Ruble, Costanzo, & Oliveri, 1992), most scholars in each field remain largely unaware of important developments in the corresponding field (Strauman & Merrill, 2004). The extent of unexplored territory is particularly ironic because the flagship journal for many social psychologists, the *Journal of Personality and Social Psychology*, was once entitled the *Journal of Abnormal and Social Psychology*. Nonetheless, what may be most intriguing about the surprising scarcity of explicitly acknowledged connections between these two fields is that both claim a kinship with personality psychology (Snyder, Tennen, Affleck, & Cheavens, 2000). The linkage of social and clinical psychology through their mutual emphasis on understanding individual variation is a theme that has yet to be fully explored, but psychologists have been at least implicitly aware of such a mutual kinship since the very beginnings of clinical psychology. This chapter considers recent contributions of social psychology to clinical psychology and in doing so draws heavily on this often-implicit pathway between the two disciplines as a basic translational principle, with an eye on the clinical implications of social psychology's emerging understanding about "what matters in the social world" (Higgins, 2000).

The translation of the situational idiom and focus of social psychology into the language of individual variability often obscures the intimate connection between so-cial and clinical research but also allows social psychology to be highly relevant for clinical psychologists. In turn, the dominant questions within clinical psychology (particularly those related to the study of psychopathology and psychotherapy broadly construed) become relevant to the interests of social psychologists when they are "backtranslated" from an individual differences framework into the language of situational reasoning. For example, it is not only the case that individuals who are prone to anxiety are reliably distinguishable from other individuals less prone to such disruptive phenomena, but it is also clear that such vulnerability can be understood in situational as well as characterological terms (Tennen & Affleck, 2002). Situations that generate anxiety, such as those in which the perceived controllability of behavior is limited, are precisely the kinds of situations that elicit the chronic lack of control experienced by those who struggle with anxiety disorders. In this instance, articulating characteristics of people vulnerable to anxiety draws heavily on social psychological research and, in turn, allows for effective theory-based intervention (Bitran & Barlow, 2004).

This chapter considers the advances that might be achieved by an appreciation of the conceptual, theoretical, and empirical kinship between clinical and social psychology. A better understanding of this kinship is particularly critical given the unfortunate tendency for contemporary research in psychopathology and treatment to take a reductionistic neurobiological perspective (Fuchs, 2004; Millon, 2004). Ironically, as theories of psychopathology have become more dimensional and less

strictly categorical, they increasingly draw situational factors into the diagnostic equation (Widiger, 2005). Psychopathology does not solely lie under the skin, in the brain, or in each individual's unique past; it also is manifest in reactivity to ordinary and extraordinary circumstances, in how people are differentially exposed to pathogenic situations, and in tendencies to engage in activities that do not fit well with an individual's underlying proclivities (Caspi, 2000).

Although clinicians still contend over which psychotherapeutic interventions are most efficacious, they are coming to realize that different structures of intervention might best fit different people (Beutler, Alomohamed, Moleiro, & Romanelli, 2002). To match people to therapies, it is most important to understand not only differences between people but the differential impact of the kinds of situations in which they live. Without an understanding of the ways in which a focus on individual variability links social and clinical psychology, public health problems such as matching psychological interventions to the needs of different individuals will not be solved.

Social psychology's exploration of clinically relevant issues demands a systematic examination beyond the scope of what a single chapter could provide. Although we are selective by necessity, we return repeatedly to a discipline-level analysis to illustrate how the two fields continue to be drawn together—the basic principle of translation that allows social psychology to have a continuous productive influence on clinical psychology. What is critical in this movement is the tendency for social psychologists to frame theories and hypothesis of situational impact in ways that are equally compatible with describing distinctions among people (Higgins, 1990). Hence the assertion with which we began: that the two disciplines are linked by their mutual compatibility with an individual differences perspective that, in turn, facilitates the application of social psychological principles to diagnosis, intervention, and prevention. We also wish to note that due to space limitations we are not able to focus on health psychology, a domain in which social psychology has been enormously influential (Strobe, 2000).

Social psychologists have been particularly adept at discerning maladaptive behavioral styles by noticing how they deviate from typical behavior in well-characterized situations. Thus, for instance, while individuals generally tend to inflate the magnitude of personal control they believe they exert over the good things that happen to them, the subset of people who do not evidence such a pattern tend to experience dysphoric affect and to be differentially vulnerable to episodes of depression (Roberts & Monroe, 1999). In essence, they are the exceptions to the "rule" of positive self-illusions that both highlight the importance of the basic social psychological insights and raise the question of whether such illusions are paradoxically adaptive. This reciprocal pattern illustrates the best of what has become known as *translational* research, and social psychologists have been leaders in the effort to bring basic science to bear on the diagnosis, treatment, and prevention of mental disorders (National Institute of Mental Health, 1999).

To present a representative survey of the frontier that lies between social psychology and clinical psychology, we use three distinct lenses. The first lens is one through which social psychologists might peer; the second, another through which clinical psychologists might peer; and the third is speculative. To be more explicit:

• From the perspective of social psychologists, what have been their field's most important contributions to clinical psychology? This perspective can be illuminated by examining the many clinically relevant hypotheses tested by research in social psychology. We already have alluded to several, but more are examined in subsequent sections. We note that this first lens has a particularly critical diffracting property—that is, viewing social psychology as a level of analysis neither reducible to nor dependent upon other equally valid levels.
• From the perspective of clinical psychologists, what have been the most important contributions of social psychological research to their own models, theories, and practices? To identify these contributions, we examine relevant clinical literatures that have directly, or indirectly, deployed social psychological concepts or models. We expect that social psychologists will be pleasantly surprised with regard to how these contributions have influenced clinical theorizing and practice.
• From our own perspective on the yet-uncharted lands between the two fields, what developments in theory and research are likely to be most influential in the near future? New ideas in the psychology of self-regulation and social neuroscience represent potentially important facets of this third lens.

In all three instances, we attempt to define principles underlying the conjunction of the two fields by examining the translational space between the situational "common person" models of social psychologists and the individual-differences-based process models of clinical psychology (Ruble, Costanzo, & Higgins, 1992). Although we acknowledge that a number of research literatures might be viewed equally productively from all three lenses, we limit our discussion of those literatures to a single perspective, recognizing that the same territory can be viewed from multiple vantage points.

CONTRIBUTIONS TO CLINICAL PSYCHOLOGY: SOCIAL PSYCHOLOGY'S VIEW

This chapter is being written at a time when social psychology's contributions to the public health missions of the National Institutes of Health (including, but not limited to, the diagnosis, treatment, and prevention of mental disorders) appear to be underappreciated (Holden, 2004). There are many reasons for this unfortunate state of affairs, including the current popularity of reductionistic approaches to brain science (Cacioppo, Berntson, Sheridan, & McClintock, 2000) and the reluctance of social psychologists to demand the spotlight on their contributions to public health (Muehrer et al., 2002). While the "decade of the brain" has been valuable

for advancing knowledge of the neurobiological aspects of psychiatric disorders (National Advisory Mental Health Council, 1988), social psychology provides complementary and equally critical levels of analysis in understanding and treating such disorders (Cacioppo & Gardner, 1999). It also is worth noting that numerous other critical problems in our society, such as violence, sexual assault/abuse, cults, and terrorism, likewise will require a social-psychological analysis as part of a comprehensive understanding (Beck, 2002).

In this section, we review some of the most important social-psychological contributions to clinical psychology from the perspective of social psychologists. These contributions span several diagnostic categories and involve both etiology and treatment but have a common element: that social psychologists are more widely aware of their bases in the study of social cognition and behavior than are clinicians (Widiger, 2001).

Depression and Self-Esteem

Social psychology has had its greatest impact on clinical psychology in the domain of depression (Dodge, 1993). Of all facets of social psychology with potential relevance to depression, *self-esteem* has been the most extensively studied. Clinical, epidemiological, social, and personality researchers have long theorized that inadequate or fragile self-esteem can lead to (Roberts & Monroe, 1994; Southall & Roberts, 2002), exacerbate (Kernis et al., 1998; Roberts & Gotlib, 1997), and result from (Hayes, Harris, & Carver, 2004; Roberts & Gamble, 2001) episodes of unipolar depression. Although the origins of self-esteem are complex and not completely understood (Hoyle, Kernis, Leary, & Baldwin, 1999), the implications of self-esteem for psychopathology and health are undeniable (Kemeny, Gruenewald, & Dickerson, 2004).

It should be acknowledged that over the past century, self-esteem has been claimed as a causal explanation for a wide variety of social ills, including violence, delinquency, teenage pregnancy, eating disorders, and marital discord (Dawes, 1994). In many cases, decreased self-esteem is more likely to be a consequence of negative social or health outcomes than a causal factor (Baumeister, 1998). In contrast, recent research has criticized the emphasis in earlier theorizing on trait self-esteem, suggesting instead that self-esteem is a dynamic phenomenon that cannot be understood simply as a trait (Dweck, Higgins, & Grant-Pillow, 2003; Gramzow, Sedikides, Panter, & Insko, 2000). For example, Crocker and Wolfe (2001) discussed the importance of both trait and state self-esteem in general and perceived contingencies of self-worth in particular. In this model, fluctuations in self-esteem occur depending on successes and failures within specific domains in which the individual's self-worth is contingent, such as appearance, academic achievement, or popularity.

There is substantial empirical evidence that low self-esteem constitutes a significant risk factor for depression, beginning with the groundbreaking work of Brown and Harris (1978). In a series of retrospective and prospective surveys, Brown and his colleagues demonstrated that chronic low self-esteem was a statistically and clinically significant factor in vulnerability to depression (e.g., Andrews & Brown, 1993, 1995). Brown proposed that self-esteem was a critical *proximal* (immediate) causal locus for the indirect effects of *distal* vulnerability factors on affective vulnerability, including childhood loss or abuse experiences, temperament, and socioeconomic factors (Brown, Andrews, Harris, Adler, & Bridge, 1986). Roberts and Monroe (1994) argued that overall level of self-esteem was not a robust predictor of the onset of depression. They proposed a multidimensional model of self-esteem in depression that included an emphasis on structural deficits within the self (such as few, rigid, or externally based sources of self-worth), low levels of self-esteem exacerbated by dysphoric mood, stressful events, or schema-congruent experiences, and temporal instability of self-worth. This model, in turn, was elaborated by Crocker and Wolfe (2001) and other investigators.

The assertion that self-esteem processes are proximal influences on mental health that can carry risk from distal factors is consistent with much current research at the social/clinical interface (Moretti, Higgins, & Feldman, 1990; Strauman & Segal, 2001). Subsequent investigations using larger community samples and more sophisticated statistical modeling lead to similar conclusions. Kendler, Gardner, and Prescott (2002) found that among women, low self-esteem was associated with risk for depression in the period following a significant life event (usually, but not necessarily, negative), though the same group also reported that self-esteem was a less robust predictor of depression than neuroticism (Roberts & Kendler, 1999). Continued study of how individuals are motivated to maintain their self-worth and how self-worth becomes contingent upon certain domains would greatly improve clinical assessment and inform treatment planning (Showers, Limke & Zeigler-Hill, 2004; Strauman & Kolden, 1997).

Depression and Attributional Style

Personality and clinical psychologists have long noticed that variability in explanatory style was associated with psychopathology. Social psychologists, in turn, have determined that people explain their outcomes through reference to a wide variety of causal factors. A vast body of evidence indicates that Heider's (1958) classic four factors (ability, effort, luck, and task difficulty) are among the most frequently offered explanations for positive or negative outcomes. Weiner's theory of attribution (Weiner, 1985) has been a dominant model of how individuals interpret their successes and failures for the past 30 years. Attribution theorists emphasize that individuals' interpretations of their outcomes determine subsequent strivings (or lack thereof) as well as the psychological consequences of perceived success and failure (Anderson, Krull, & Weiner, 1996). Weiner's model classified attributions into three causal dimensions: locus of control, stability, and controllability. In turn, each of these causal dimensions has identifiable influences on motivation, behavior, and affect. Some degree of self-serving attributional bias (a tendency to attribute posi-

tive, but not negative, outcomes to one's own efforts, ability, etc.) appears to be universal (Mezulis, Abramson, Hyde, & Hankin, 2004), and greater positive bias is associated with greater psychological and physical health (Peterson & Seligman, 1987). Social-psychological theories of explanatory style and attributional biases have been enormously influential in the development of cognitive models for depression, anxiety, and related psychological problems (Bell-Dolan & Anderson, 1999).

The hopelessness theory of depression (Abramson, Metalsky, & Alloy, 1989) takes an explanatory style approach to the psychological processes associated with vulnerability to depression. Hopelessness theory proposes that specific attributional styles predispose individuals to depression when they encounter negative events. The theory focuses on three distinct depressogenic explanatory styles: attributions about the self (the tendency to draw negative inferences about one's worth, desirability, abilities, etc.), about consequences of one's actions (the tendency to catastrophize the consequences of negative events and to view these consequences as important, unlikely to change, and as affecting many areas of one's life), and about causes (the tendency to attribute negative events to global and stable causes). According to Abramson and colleagues, individuals with depressogenic attributional styles are more likely than individuals without such tendencies to make depressogenic inferences about negative events they encounter, which increases the likelihood that hopelessness (and subsequent depression) will result. The hopelessness theory is a powerful conceptual framework for predicting vulnerability to depression with substantial empirical support (Gibb, Alloy, Abramson, Beevers, & Miller, 2004). Along with Beck's cognitive theory of depression (Clark, Beck, & Alford, 1999), itself borrowing heavily from social cognition research, hopelessness theory represents a major success story in the social/clinical interface.

Thought Suppression and Psychopathology

The paradoxical tendency for unwanted suppressed thoughts to return to consciousness with increased frequency is a phenomenon that has ignited a plethora of research since the first research demonstration of the postsuppression rebound effect (Wegner, Schneider, Carter, & White (1987). Since its identification, researchers have identified three classes of suppression-related effects: (1) the immediate occurrence of target thoughts after a period of suppression; (2) an immediate suppression-induced surge in target thoughts; and (3) an intensification of intrusions during suppression when under cognitive load (Wenzlaff & Wegner, 2000).

Theoretical accounts exist that shed light on aspects of the process of mental control; however, to date, no single theory explains all the empirically observed results. The theory of ironic processes of mental control (Wegner & Wenzlaff, 1996) proposes that two processes are used when exerting mental control: (1) an intentional operating process that searches for mental contents that will yield the desired mental state and (2) an ironic monitoring process that searches for mental contents that signal

failure to achieve the desired state. For example, an intentional operating process may search for thoughts unrelated to romantic relationships, while the ironic monitoring process searches for mental contents that are related to romantic relationships. The operating process is described as effortful and requiring conscious thought, whereas the monitoring process is unconscious and does not require significant mental effort. This monitoring process is considered ironic because it is actively searchers for cues related to the thoughts to be suppressed. The theory goes on to state that even when the operating process is voluntarily terminated or is altered by increasing cognitive load, the monitoring process continues its vigilance for unwanted thoughts. This account explains the tendency to experience immediate unwanted thought intrusions and the increase in these intrusions while under high cognitive load (the postsuppression rebound effect).

Thought suppression research has the potential to influence both understanding and treatment of various forms of psychopathology, including obsessive–compulsive disorder, posttraumatic stress disorder, mood disorders, eating disorders, and borderline personality disorder. For example, Hardy and Brewin (2005) reported that efforts at thought suppression lead to an increase in obsessional thinking and symptoms, and Purdon, Rowa, and Antony (2005) noted a reliable effect of thought suppression on the frequency of unwanted thoughts (as well as associated distress) in individuals meeting criteria for obsessive–compulsive disorder according to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). Thought suppression has a demonstrable negative impact on eating behavior among dieters, restrained eaters, and individuals with eating disorders (Boon, Stroebe, Schut, & Ijntema, 2002; Mann & Ward, 2001). Similarly, Cheavens and colleagues (2005) observed that thought suppression moderates the impact of perceived criticism and distress on maladaptive behavior in borderline personality disorders. Studies of risk for unipolar depression suggest that thought suppression can be beneficial under some circumstances but is associated with greater symptoms in the presence of high levels of life stress (Beevers & Meyer, 2004; Rude & McCarthy, 2003).

Many of the effective treatments for these disorders focus on identifying and challenging cognitions that are maladaptive, distressing, intrusive, and perceived as uncontrollable. Research taking into account ironic mental processes has suggested that interventions focus on assisting patients in identifying effective distractors, pursuing alternate goals, and practicing acceptance of unwanted intrusive thoughts. For example, Roemer and Salters (2004) explored the effects of teaching rape victims to actively suppress intrusive thoughts on the short-term incidence of those thoughts. Dialectical behavior therapy and the associated skills training modules is an example of a second-generation cognitive-behavioral treatment that focuses on practicing acceptance of unwanted thoughts and feelings as a way of teaching distress tolerance and emotion regulation skills (Linehan, 2000; Robins, 2002).

Social-Cognitive Aspects of Body Image Disturbance and Eating Disorders

The emergence of body image dissatisfaction, maladaptive eating behaviors, and frank eating disorders ranks as one of the most challenging public health crises of the past century (Currin, Schmidt, Treasure, & Jick, 2005). With the exception of models focusing on metabolic abnormalities and physiological sequelae of disordered eating (e.g., Siegfried, Berry, Hao, & Avraham, 2003), the theoretical models that dominate the body image, eating disorders, and obesity literatures are social-cognitive in nature. In particular, three themes originating with social psychology pervade these literatures: social-cognitive and interpersonal influences on body image, social/interpersonal influences on food intake, and disordered eating as a means for coping with self-esteem, identity, and control issues.

In an influential review article, Polivy and Herman (2002) noted that social psychological principles were essential to understanding the intrapersonal and sociocultural causes of eating disorders. We endorse their conclusion and proposed to expand on it as follows. Any of the social-cognitive processes that influence self-perception and self-evaluation have the potential to lead to a distorted internal representation of one's body and appearance and/or a chronic sense of dissatisfaction with one's appearance. Other social-cognitive processes influence food intake and thus may be implicated in the development of disordered eating. Although genetic and biological factors have a role to play in vulnerability to eating disorders, the critical causal factors are predominantly social.

Social-Cognitive and Interpersonal Influences on Body Image

The psychology of body image formation, maintenance, and disturbance incorporates many of the most popular theories in social psychology (Cash, 2004). Individual differences (and intraindividual variation) in body image, along with gender differences in the determinants of body satisfaction, have been the focus of literally thousands of studies in the laboratories of social and clinical psychologists (Brownell, 1991). Given the enormous influence of symbolic interactionism and social constructionism on contemporary social psychology, it is hardly surprising that in order to understand body image, theorists have considered the role of the society (and its internalized representations) in determining how individuals perceive and evaluate their bodies (Reischer & Koo, 2004). In turn, body image distortion and dissatisfaction represents a major risk and maintenance factor for eating disorders, including bulimia, anorexia, binge eating disorder, and obesity (Stice & Shaw, 2002).

Body image can be conceptualized as a multidimensional construct that represents how individuals think, feel, and behave with regard to their own physical attributes (Muth & Cash, 1997). Body image is undoubtedly a social and cultural construction, and both historical and contemporary data indicate that self-esteem is greatly influenced by the extent to which the individual perceives his or her body as congruent with societal and local ideals (Van Wolputte, 2004). Beginning in childhood and extending into adolescence, both boys and girls come to evaluate their bodies in reference to cultural standards; however, in general, the data suggest that body image tends to be more critical for girls' self-esteem than for boys (Frost & McKelvie, 2004). Interestingly, recent studies suggest that social comparison theory, a mainstay of social psychology, is useful in explaining the dynamics of body image evaluation by the time children reach adolescence (Morrison, Kalin, & Morrison, 2004). Those dynamics include a tendency for girls and women to adopt observers' perspectives on their bodies and appearance, resulting in a greater vulnerability to distress and body shape/size distortion (Frederickson, Roberts, Noll, Quinn, & Twenge, 1998).

In turn, although both boys and girls can experience body dissatisfaction, the observed gender differences in body image tendencies predicts that more adolescent girls and adult women will report significant dissatisfaction with their bodies and appearance (Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999). Body dissatisfaction appears to be pervasive in American culture and constitutes a risk factor for eating disorders, depression, and substance use (Polivy & Herman, 2002). For both genders, behaviorally and clinically significant body dissatisfaction typically begins after puberty; a combination of biological, interpersonal, and sociocultural factors predicts dissatisfaction with appearance but does so more strongly for girls (Stice & Whitenton, 2002).

Both cross-sectional and longitudinal studies indicated that attitudes toward one's appearance are both influenced by and influence the individual's interpersonal context. In addition to the continuous presence of socially sanctioned ideal body types as presented by mass media, children's and adolescents' attitudes toward their appearance are substantially influenced by their parents and peers (Stice & Shaw, 2002). The belief that one is unattractive (in general) or overweight (in particular) significantly impairs the development and maintenance of intimate relationships and leads to an attachment style that reflects both negative self-evaluation and excessive concern with evaluation by others (Cash, Theriault, & Annis, 2004). Interestingly, the act of comparing one's appearance or body characteristics with those of others may constitute a critical psychological mechanism by which the influence of social and cultural factors is translated into body dissatisfaction, excessive dieting, and maladaptive eating habits (van den Berg, Thompson, Obremski-Brandon, & Coovert, 2002). Further support for this assertion comes from studies in which exposure to more realistic body images protects individuals from the negative consequences of thinness and restricted-eating expectancies (Fister & Smith, 2004).

Social and Interpersonal Influences on Food Intake

Anecdotal evidence and common sense both suggest that the presence of others influences how much people eat, but it has been social psychologists who have done much

of the research on these influence processes. As summarized by Herman, Roth, and Polivy (2003), that research supports three conclusions: (1) when people eat in groups, they tend to eat more than they do when alone (a social facilitation effect); (2) when individuals eat in the presence of others who consistently eat a lot or a little, those individuals likewise tend to eat a lot or a little (a modeling effect); and (3) when people eat in the presence of others whom they believe are evaluating or observing their eating, they tend to eat less (an impression management effect). Herman and colleagues proposed an inhibitory norm model to account for the social circumstances under which the same individuals would eat less or more than what they otherwise would eat when alone. It is interesting to note that this model incorporates both situational factors and individual differences, providing a clear illustration of how general social psychological principles are applicable to fundamental human activities.

Indeed, a broad range of social psychological processes have been implicated in food intake, including self-efficacy (e.g., Steptoe et al., 2003), regulatory focus (e.g., Förster, 2003), self-enhancement (e.g., Mills, Polivy, Herman, & Tiggemann, 2002), and conformity (e.g., Herman, Fitzgerald, & Polivy, 2003). This area of research provides a particularly powerful and salient demonstration of the importance of "the actual, imagined, or implied presence of others" (Allport, 1985, p. 3) on behavior. For instance, Herman, Fitzgerald, and Polivy (2003) found that hunger ratings of research participants were significantly altered by the availability of information regarding ratings of fictitious others. Similarly, Mills and colleagues (2002) observed that the impact of thin media images on individuals prone to restrained eating was mediated by self-enhancement motives—providing a much more powerful model (in both conceptual and statistical terms) for the impact of distal factors like media depictions of ideal body types on eating behavior.

Disordered Eating as a Coping Mechanism

Beginning with the hypothesis proposed by Heatherton and Baumeister (1991) that certain individuals seek to "escape" from a state of aversive self-awareness through binge eating, there has been an enormous amount of research on the causes, nature, and consequences of disordered eating. In addition to continued development of that and related eating-as-coping models, research in this area has documented that the same social-cognitive stresses implicated in stress reactivity also predict disordered eating (Bennett & Cooper, 1999) and that sociocultural messages (carried by family, peers, and media) combine with self-evaluation, expectancies, and related cognitive processes to create and maintain maladaptive eating patterns (Stice, 1994).

The study of disordered eating has become an integrative enterprise. Individual-differences variables such as coping (Freeman & Gil, 2004), attitudes (Garcia-Grau, Fuste, Miro, Saldana, & Bados, 2002), self-perception (Vohs, Heatherton, & Herrin, 2001), and ego involvement/strength (Heatherton, Herman, & Polivy, 1992;

Vohs & Heatherton, 2000) all appear to mediate/moderate the impact of situational cues on eating. In turn, situational factors such as modeling (Rotenberg, Carte, & Speirs, 2005), behavioral norms (Thome & Espelage, 2004), and peer pressure (Fischer, Anderson, & Smith, 2004) all exert influences on eating behavior. Of particular interest is the possibility that the particular mechanisms underlying disordered eating may vary substantially across different ethnic and racial groups (e.g., Vander Wal & Thomas, 2004).

Summary

In sharp contrast to the apparent misconception that social psychological research is not critical to the nation's public health, social psychology has made broad and deep contributions to the understanding of several highly prevalent forms of psychopathology. The contributions reviewed in this section represent trends in social psychology that the researchers themselves see as critical for the science and practice of clinical psychology. As we commented in the introduction to this chapter, social psychological principles are both powerful concepts for understanding psychopathology in their own right and are, when translated into the language of individual differences, often critical mediating variables in the link between environmental and cultural forces and distress. It is worth noting that the application of social psychology to clinical psychology has progressed far beyond simple translation. Rather, clinical investigators have extended and combined individual models to construct more complex and realistic models of etiology that take both distal and proximal contributory factors into account. In turn, constructing such models allows for more creative translation into therapeutic and preventive intervention (Cuthbert, 2002; Strauman & Merrill, 2004).

CONTRIBUTIONS TO CLINICAL PSYCHOLOGY: THE CLINICIAN'S PERSPECTIVE

It may surprise social psychologists to know that theories regarding social influence, social cognition, and other social processes have been incorporated into research on the etiology and treatment of disorders as diverse as schizophrenia, substance use, and generalized anxiety. Our impression is that clinical researchers and practitioners who work with these disorders are much more aware of the contributions of social psychology to these literatures. In some cases, the social psychological models are so integral to an understanding of the disorders that they may no longer be recognized as having those origins (at least among social psychologists); nonetheless, the contributions are impressive and we summarize a subset of them below. Due to space limitations, we focus on two disorders (schizophrenia and substance abuse) and refer the reader to excellent available summaries of social psychological processes in generalized anxiety disorder and social anxiety disorder (for reviews, see Alden & Taylor, 2004; Coles & Heimberg, 2002; Leary, 2001; Stanley & Beck, 2000). As in the previous section, we find

it useful to contemplate these literatures simultaneously in terms of the social/interpersonal forces involved and the ways in which those models have been translated into the language of individual differences in order to account for vulnerability to psychopathology.

Social Information Processing in Schizophrenia

One of the refreshing aspects of the recent emphasis in psychopathology research on understanding brain circuitry and its abnormalities is the recognition that there can be no real understanding of the neural basis of a psychological process without sufficient attention to the process itself (Adolphs, 2003). In fact, impairments in social functioning are among the defining characteristics of schizophrenia according to DSM-IV (American Psychiatric Association, 1994), and so schizophrenia researchers have borrowed liberally from models of social information processing to explore the psychological and neural mechanisms that account for deficits in social functioning (Lee, Farrow, Spence, & Woodruff, 2004). Most recently, a group of schizophrenia researchers is working to develop a standard battery of measures of cognitive functioning that includes assessment of social cognition (Green et al., 2004).

It has been argued that deficits in social cognition are at least as central to thought disorder syndromes such as schizophrenia as are deficits in non–social cognitive processes, in part because social cognition deficits are better predictors of symptoms and functional impairment in schizophrenic patients (Penn, Corrigan, Bentall, Racenstein, & Newman, 1997). Research on the development of social cognition also is of value in schizophrenia research, because it allows investigators to determine which processing deficits are characteristic of vulnerability to the disorder (i.e., are necessary but not sufficient precursors that may reflect a genetic, epigenetic, or environmental risk factor) and which reflect progression into clinical disorder (Carter & Flesher, 1995; Walker, 1994). We review several of the most influential lines of research in the schizophrenia literature that reflect social psychological models of information processing. The reader should note, however, that numerous other phenomena studied by social psychologists are relevant to understanding the cognitive and interpersonal deficits that characterize psychosis proneness and schizophrenia spectrum disorders (e.g., autobiographical memory; see Klein, German, Cosmides, & Gabriel, 2004).

Theory of Mind

Beginning with nonhuman primate research (Premack & Woodruff, 1978) and extending into human studies (Wimmer & Perner, 1983), behavioral scientists have been intrigued by the problem of understanding how people explain the behaviors of others. "Theory of mind" (ToM) refers to the observation that people believe other people have minds like theirs and that they can understand the behaviors of others in terms of the same kinds of mental contents they use to understand their own behavior, including knowledge, beliefs, and desires (Baron-Cohen, Tager-Flusberg, & Cohen, 2000). Frith (2004) noted that ToM had been applied to the study of autism, leading to the observation that performance on ToM tasks could be impaired even though other seemingly related social and intellectual abilities remained intact. Because impairment in social interaction (particularly a tendency to withdraw from others into oneself) is also characteristic of schizophrenia, Frith and colleagues hypothesized that a similar social-cognitive deficit was characteristic of schizophrenia (Frith & Frith, 1991).

Recent studies of ToM in schizophrenia have had a substantial impact on understanding of the disorder as well as on the larger research question of how human social interactions communicate knowledge, beliefs, and intent. Consistent with the observation that discrete areas of the brain appear to underlie ToM (e.g., Saxe & Kanwisher, 2003), impairments in other aspects of cognitive functioning (such as executive function and emotion recognition) found in schizophrenia do not account for deficits in ToM (Brune, 2005). Researchers taking a developmental perspective have been investigating whether ToM-type deficits can be detected prior to the onset of schizophrenia spectrum disorders. Schiffman and colleagues (2004) reported deficits in perspective-taking skills among children who later developed schizophrenia as compared to controls, indicating that such deficits may contribute to the development of such disorders. In a related investigation, Kelemen, Keri, Must, Benedek, and Janka (2004) found no evidence for impaired ToM skills in unaffected first-degree relatives of schizophrenia patients. As such, ToM is unlikely to be a necessary characteristic of an inherited vulnerability to schizophrenia.

Social Perception

Even the earliest descriptions of schizophrenia noted the inaccurate and often bizarre ways in which such individuals interpreted the behaviors of others, including paranoid delusions (Kraepelin, 1893, as cited in Ban, 2004). The study of social perception abnormalities in schizophrenia has proceeded in both bottom-up (Insel & Fernald, 2004) and top-down (Pinkham, Penn, Perkins, & Lieberman, 2003) directions. Just as important, insights gleaned from applying social information-processing models to the social perception and interpersonal abnormalities characteristic of schizophrenia spectrum disorders suggest avenues for adjunctive and perhaps even preventive interventions, as described below (Penn et al., 2004).

A number of investigations suggest that social cognition deficits may mediate the relationship between neurocognitive and behavioral/experiential aspects of schizophrenia. Wynn, Sergi, Dawson, Schell, and Green (2005) found that social perception deficits in schizophrenia were associated with a measure of neurocognitive impairment, sensorimotor gating. The same research group reported that social perception deficits were linked with visual processing abnormalities (Sergi & Green, 2002). Combs and Penn (2004) observed that college-age individuals manifesting subclinical paranoia (and who therefore were at risk for developing

schizophrenia spectrum disorders) performed significantly worse than matched controls on an *in vivo* social perception task and sat further away from the examiner during the task. Based on this emerging literature, it appears that the deficits in social cognition associated with schizophrenia spectrum disorders are central rather than peripheral aspects of vulnerability and clinical manifestations (Higgins & Moretti, 1988; Lancaster, Evans, Bond, & Lysaker, 2003).

Cognitive Remediation

The reliable association between social-cognitive processing abnormalities and other signs and symptoms of schizophrenia raises the obvious question of whether those processing deficits could be reduced or eliminated through treatment. It may surprise social psychologists to learn that models and methods developed in the social cognition literature are now routinely used in the process of antipsychotic drug development (Green & Braff, 2001). Although until recently treatments for schizophrenia were targeted almost exclusively at symptoms such as hallucinations and delusions, a consensus has emerged that treatments also need to reverse the cognitive and social-cognitive deficits summarized previously (Harvey, Green, Keefe, & Velligan, 2004).

Penn and colleagues (2004) have suggested that, at minimum, treatment for schizophrenia needs to include efforts at remediation of, or at least compensation for, social information-processing deficits and abnormalities. Although such interventions are too new to be in widespread use, there already is solid evidence that cognitive-behavioral techniques can effectively reduce deficits in executive function, attention, and memory. Kurtz, Moberg, Gur, and Gur (2001) reported a meta-analysis of intervention studies testing cognitive remediation strategies in schizophrenia and found a large average effect size ($d = 0.96$), with particularly robust effects observed in studies using encoding elaboration strategies and vigilance training to enhance social cue recognition accuracy.

Given the success of cognitive-behavioral interventions for treatment of depression, it was inevitable that such techniques would be adapted for use in adjunctive treatment of schizophrenia (Turkington, Dudley, Warman, & Beck, 2004). However, cognitive interventions in schizophrenia were intended primarily to reverse deficits in social and nonsocial information processing, rather than symptom reduction (Hogarty & Flesher, 1999). In the past 5 years alone, more than a dozen controlled clinical trials of cognitive-behavioral therapy for schizophrenia have been published, providing compelling evidence that interventions to improve social cognition can provide substantial benefits for patients over and above the effects of antipsychotic medications (which do not show reliable effects on social information processing) (e.g., Hogarty et al., 2004). Similarly, combining cognitive-behavioral therapy with social skills training leads to clinically and statistically significant improvement in interpersonal functioning among individuals with chronic schizophrenia (Granholm et al., 2005). Additional bene-

fits of these interventions include enhancing compliance with medication regimens, improving tolerance of hallucinations and other altered perceptions, and decreased likelihood of relapse (Trower et al., 2004). And perhaps the most exciting work in this field has just begun—intervention trials to determine whether targeting social information-processing abnormalities in at-risk individuals can prevent the onset of the disorder itself (e.g., Newton et al., 2005).

The Social Psychology of Substance Use

For most individuals who use illicit substances, taking drugs is a social event. Their knowledge of and exposure to drugs comes from their social contexts, their initial drug use occurs with peers, and their drug-use decision making occurs with reference to peer norms and feedback (Dishion, Capaldi, Spracklen, & Li, 1995). Association with drug-using peers is the single most potent predictor of drug-use initiation (Warr, 1996). Furthermore, diffusion of drug use across neighborhoods and communities may well follow contagious processes, with one critical pathogenic agent being peer interaction. Unfortunately, current drug-use prevention programs and practices do not benefit sufficiently from basic science knowledge of processes in social decision making, peer influence, and social diffusion (Miller-Johnson & Costanzo, in press). Ironically, recent findings indicate that many current policies and practices may even inadvertently exacerbate drug use by aggregating high-risk youth with deviant peers who provide a drug-use training ground. When peers do receive the attention of practitioners, they are the "enemy" that is the focus of peer-resistance skills training programs, rather than an ally.

In this section, we focus on two of the many contributions of social psychology to the understanding, treatment, and prevention of substance use. First, we briefly discuss peer norms and their influences on use during adolescence, with particular attention to the implications of these processes for intervention. Second, we examine the role of social information processing in laboratory and applied studies of "choosing to use." There are many other perspectives on substance use that could be included here (e.g., Gibbons, Gerrard, & Lane, 2003). We have selected these two simply in order to illustrate how critical social psychological variables are for understanding addiction—as demonstrated in the growth of social psychology research within the funding portfolio of the National Institute of Drug Abuse (NIDA; 1999).

Peer Influences on Substance Use

Recent theoretical accounts, empirical studies, and prevention trials have made clear that social influences are the primary factors in initiation of drug-using and -abusing behaviors (Donaldson et al., 1996; Flom, Friedman, Jose, & Curtis, 2001; Harrison, Fulkerson, & Park, 2000; Kandel & Adler, 1982; Olds & Thombs, 2001). Among the multiple sources of social influence, peer influence appears to have the most profound and lasting effects on beliefs about substance abuse outcomes and on the drug-

use behaviors that ensue from these beliefs. For example, in a 15-year, follow-back retrospective study, Flom and colleagues (2001) found that drug abuse in adulthood was strongly related to peer approval norms associated with drug use at age 15. Similarly, Olds and Thombs (2001), in a survey-based study of over 2,000 7th- to 12th-grade students, found that perceived "close friend" peer norms with regard to cigarette and alcohol consumption accounted for six to eight times the variance in substance use, as did supportive parent involvement in the life of the adolescent. Research by Seiving, Perry, and Williams (2000) convincingly demonstrates that the strong effect of peer influence on substance use is not simply an artifact of selecting similar peers as friends. Instead, their results reveal that substance use follows directly from the influence of peer norms, and that relationships and cliques solidify around such norms subsequent to friendship formation.

Given the critical role of peer norms in the onset of substance use during adolescence, numerous preventive interventions have targeted peer processes in middle and high school. However, the need to focus social influence-based programming on issues of adolescent social choice that range beyond narrow conceptions of peer pressure resistance is illustrated by Lynam and colleagues' (1999) comprehensive longitudinal evaluation of the D.A.R.E. program, the most frequently employed model of peer resistance training. Lynam and colleagues found that school-based peer resistance programming had *no* effect on either drug attitudes or illicit drug use assessed 10 years after the program. Many programs appear to primarily or exclusively target "refusal" or resistance skills, whereas the social development literature suggests the ubiquity of peer influences and the potential merits of using peer leaders in prevention rather than asking individuals to defy them (Perry, Williams, Veblen-Mortenson, & Toomey, 1996; Price, Gioci, Penner, & Trautlein, 1993).

A number of investigators have begun to examine the implications of peer leadership among adolescents for substance use attitudes and behaviors. Luthar and McMahon (1996), in their sample of inner-city teens, found two contrasting patterns of admired, well-liked teens. The first group was characterized by conventionally valued behavior, such as prosocial behavior and academic achievement. By comparison, the second group was characterized by disruptive/aggressive behaviors and poor school performance. Similarly, Rodkin, Farmer, Pearl, and Van Acker (2000) found two subtypes of well-liked peers: popular-prosocial ("model") and popular-antisocial ("tough") boys. The "tough" group rated themselves as being aggressive, "cool," and athletic, whereas the "model" group described themselves as also being cool and athletic but at the same time non-aggressive and academically competent. Both groups of these well-liked boys were highly central in their peer groups. These findings suggest that youth who are admired among peers are a heterogeneous group, and highly aggressive youth may be among the most influential and prominent among their peers.

Social Information Processing and Vulnerability to Substance Use

Social psychologists have been integral in the development and validation of conceptual models for understanding how substance use behaviors develop and are maintained. Using the language and concepts of social cognition, two main questions are being pursued within this broader framework: what types of social situations most commonly lead to drug use, and what social-cognitive processes mediate decision to use. Each question leads directly to translational research for preventive intervention, including school-based programs intended to prevent or delay first use.

Substance use is particularly likely when individuals are seeking relief from life stress and the tensions and struggles of living (e.g., Greeley & Oei, 1999). As a result, it has been hypothesized that risk for first use (and development of substance abuse behavior) will be greatest in situations characterized by high levels of distress. In this way, emotional distress serves as a key mediating variable in the prediction of substance use, and the substance use itself is seen as a concrete manifestation of inadequate coping skills. A number of preventive interventions have been developed that not only target behaviors directly associated with drug and alcohol use but also provide instruction in anger management and stress reduction techniques (e.g., Webb, Scudder, Kaminer, & Kaden, 2002).

The search for psychological variables that act as mediators in the onset and maintenance of substance use has led to adoption of social information-processing models developed for the study of other kinds of behavioral problems. Consider a hypothetical 14-year-old boy who is with a group of friends who bring out a marijuana cigarette and try to cajole the boy into trying it. Theory and research in antisocial behavior indicates that this decision is highly influenced by the individual's relationship with peers, schemas, and scripts about drugs and peer interaction and the manner in which the stimulus is processed online (Dodge & Pettit, 2003). Online processing can be conceptualized as occurring in sequential steps of encoding, interpretation, response accessing, response decision, and response enactment (Crick & Dodge, 1994). Patterns in processing can be measured reliably and have been found to predict growth in aggressive behavior across development (Dodge et al., 2003). Furthermore, cognitive-emotional processes have been found to mediate the effects of other distal risk factors in maladjustment outcomes (Graziano, Jenson-Campbell, & Finch, 1997). The intervention potential of this line of research is implicated in findings that parental and school supports for individual self-regulation and appropriate autonomy are correlated with successful adolescent outcomes (Eccles, Early, Frasier, Belansky, & McCarthy, 1997). A small but growing literature has suggested that individual differences in social cognition are correlated with drug-use behavior (Schmid, 2001; Wagner, Myers, & McIninch, 1999). A large prospective study indicates that social-cognitive processes in elementary school

predict initial drug use in middle school (Kaplow, Curran, Dodge, & Conduct Problems Research Group, 2002).

Summary

We presume that our convictions regarding the importance of the social/clinical psychology interface are evident to the reader by now. Nonetheless, we wish to emphasize that the contributions of social psychology are particularly evident within the context of disorders such as schizophrenia and substance abuse that are widely viewed as biologically based. In contrast to the popular belief that psychological theory and practice are increasingly irrelevant to treatment and prevention of mental disorders (Flanagan & Blashfield, 2000), we suggest that researchers and clinicians working in these two fields are as convinced of the importance of the social psychological level of analysis as we are (Holtzman, 2003). That is, even (and we are tempted to say *particularly*) in disorders with a clearly established neurobiological basis, social processes are implicated in the onset, persistence, and consequences of those disorders.

We asserted at the start of this section that the contributions summarized here are better known among clinical psychologists than their social psychologist colleagues. In a critical respect, that is indeed the case; a quick examination of the journals in which the aforementioned studies appeared makes clear the distinction between the first two "lenses" of this chapter. Whereas most of the social psychological research on depression (for instance) appears in social psychology journals, social information processing studies of schizophrenia are almost exclusively found in the journals of clinical psychology and psychiatry. However, there is no doubt that social psychologists would recognize the phenomena, theories, and methods that our clinical colleagues have borrowed. Likewise, both the clinical psychologists engaged in this research and the social psychologists viewing them from a distance would agree that combining a focus on social/interpersonal factors with an individual differences approach provides a critical level of analysis for understanding the etiology, treatment, and prevention of mental illness.

CONTRIBUTIONS TO CLINICAL PSYCHOLOGY: A GLIMPSE INTO THE FUTURE

What developments at the interface of social and clinical psychology are likely to be influential in the near future? This final lens requires us to stick our collective necks out a bit further, but we hope to convince the reader that there are indeed several research areas in social psychology that have clear, and we believe imminent, implications for clinical psychology. As we have argued throughout, this influence will likely occur through translating social psychological research into the language of individual differences. For reasons of convenience and familiarity, we focus on research related to depression, but we suggest that their ultimate impact on clinical research and practice is likely to extend beyond that particular disorder.

Self-Regulation, Repetitive Thought, and Depression

Current theories of self-regulation provide robust and thoughtful accounts of how individuals pursue personal goals and the kinds of motivational and emotional states they experience when they see themselves as attaining, or failing to attain, such goals (Carver & Scheier, 1999). A number of investigators have attempted to link problems in self-regulation with vulnerability to depression and other disorders (e.g., Karoly, 1993; Strauman, 2002). However, self-regulation theories have not yet adequately articulated how the experience of *acute* emotional distress following failure to attain important goals becomes the *chronic* emotional distress characteristic of mood and anxiety disorders. Because only a subset of individuals experiencing chronic difficulties attaining personal goals ever become depressed, there may be other factors that determine whether a particular individual responds adaptively in the face of continued failure feedback or becomes mired in a downward spiral of negative self-evaluation, doubt, and distress. One candidate for such a factor is maladaptive repetitive thought, also known as rumination.

Repetitive thought (RT) can be defined as "the process of thinking attentively, repetitively, or frequently about oneself and one's world" (Segerstrom, Stanton, Alden, & Shortridge, 2003, p. 909). Both theory and research suggest that the tendency to engage in specific types of RT in response to goal blockage predicts the likelihood of self-regulatory failure (Carver & Scheier, 1990; Pyszczynski & Greenberg, 1987). In this context, self-regulatory failure refers to an inability to make progress toward important goals and a subsequent inability to reprioritize or disengage these goals from the regulatory system (Emmons, King, & Sheldon, 1993). RT can lead to continued activation of a perceived discrepancy between the current state of being and the chosen goal, causing the goal to remain salient and distressing.

Martin and Tesser (1996) proposed a theory of rumination, which they defined as a "class of conscious thoughts that revolve around a common instrumental theme and that recur in the absence of immediate environmental demands requiring the thoughts" (p. 7). They posited that rumination is instigated by a failure to make progress toward a desired goal, and that attaining or disengaging from the blocked goal will terminate the ruminative process. However, redirecting thoughts away from the distressing content or reducing negative affect associated with the goal blockage is likely to only temporarily halt the process, because the continual cuing of goal-related thoughts by features of the social environment make continued distraction difficult. Other theorists have proposed that an unfavorable assessment of the expectancy of succeeding at goal-directed action instigates the process of rumination—especially if the blocked goal is central to the self (Carver, 1996).

While Martin and Tesser's theory of rumination is elegant, there is a critical distinction between problem solving and rumination. Problem solving refers to the process of recognizing that a problem exists, defining and representing the problem mentally, developing a solution strategy, organizing current knowledge about the problem, allocating mental and physical resources needed to solve the problem, monitoring progress toward the goal, and evaluating the quality of the solution (Pretz, Naples, & Sternberg, 2003). In contrast, rumination is better conceptualized in this context as "repetition of a theme in thoughts, without progression toward choice of a solution and a commitment to that solution" (Nolen-Hoeksema, 1996, p. 137).

Nolen-Hoeksema has studied rumination as a coping mechanism in depression (e.g., Nolen-Hoeksema, 2000). She defines ruminative coping "behaviors or thoughts that focus an individual's attention [on] the possible causes and consequences of that mood" (Nolen-Hoeksema, Morrow, & Fredrickson, 1993, p. 20). In this account, rumination is not instigated by a self-regulatory failure but is instead a maladaptive response to dysphoric mood. Indeed, dysphorics who ruminate demonstrate lower problem solving and less ability to generate effective solutions to interpersonal problems when compared to nondysphorics and dysphorics who use distraction (Lyubomirsky, Tucker, Caldwell, & Berg, 1999). We suggest that these findings could be integrated within the self-regulatory perspective taken by Martin and Tesser (1989) to create a more elaborated theory accounting for how acute failure feedback (which everyone experiences) becomes transformed over time into chronic failure feedback (which many people experience) and ultimately into a clinically depressed state (which most people receiving failure feedback do *not* experience).

Consider the following preliminary example of such an integrative approach. In the ongoing process of self-regulation, the individual evaluates his or her progress by monitoring the magnitude of perceived discrepancy between actual behavior and a particular goal or standard (Higgins, 1987, 1997) and the perceived rate of discrepancy reduction or enlargement (Carver & Scheier, 1990). Specifically, when using promotion (Higgins, 1997) as a means of self-regulating behavior, dysphoric emotions occur when the individual detects a discrepancy and subsequently evaluates that insufficient progress is being made. The intensity of the emotional discomfort resulting from this monitoring process is directly related to the magnitude, accessibility, and contextual relevance of the detected self-discrepancy, as well as the centrality of the discrepancy to a person's self-concept. A large self-discrepancy that is highly accessible, contextually relevant, and central to a person's self-concept is associated with a more intense emotional response.

The monitoring also generates a sense of expectancy (optimism or doubt) that influences subsequent motivation (Carver & Scheier, 1990). Under conditions that signal difficulties in goal attainment, a sense of optimism or doubt is generated that causes the person to initially respond by increasing the degree of persistence and effort of the initial instrumental behavior (Martin & Tesser,

1996). Detection of a large discrepancy generates doubt, whereas detection of a small discrepancy generates optimism that determines the degree of effort placed into repeating the instrumental behavior. At this stage, the evaluation that insufficient progress is being made causes the individual to disengage from the initial instrumental behavior to engage in a full evaluation of the likelihood of success (Carver & Scheier, 1998). In parallel with the expectancy evaluation, an individual may engage in problem solving: defining and representing the problem mentally, developing a solution strategy, organizing current knowledge about the problem, and allocating mental and physical resources needed to solve the problem (Pretz et al., 2003). Problem solving may entail attempts to specify alternate pathways that lead to attainment of the higher-order goal, redefining the goal at lower level of specification, or engaging in a process of redefinition of the self.

One of the most critical issues in developing this kind of integrative model is, how does acute failure/distress evolve over time into chronic failure and distress? Existing theories help to specify some of the conditions for such an evolution. Disengagement from a goal may not be possible due to the centrality of the goal to a person's self-concept; that is, the goal may reside at high levels within a person's goal hierarchy. This inability to disengage would lead to further attempts to attain a goal that are imbued with doubt, creating the conditions for further failure experiences as well as maladaptive rumination. In turn, maladaptive rumination results in continued activation of a perceived discrepancy, causing the goal to remain salient (Martin & Tesser, 1996). This begins a cyclical process where continued indirect and direct cuing of the discrepancy generate more intense negative affect and increasing levels of maladaptive rumination. The rumination prolongs existing negative affect and generates more negative affect as individuals begin to engage in secondary appraisals that their rumination is uncontrollable, harmful, and likely to produce detrimental interpersonal as well as social consequences (Papageorgiou & Wells, 2004). In this state, the individual is no longer actively producing solution strategies but is likely to be thinking about the goal object and his or her feelings regarding the object. This process can cause polarization of existing affect, which intensifies the negative experience (Martin & Tesser, 1989).

Combining this cognitive perspective on self-evaluation with the affective and motivational mechanisms proposed in regulatory focus theory suggests that a particular sequence of events would be discriminantly associated with depression. Over time, the repeated failure to attain promotion goals could result in disengagement from attempts to pursue them, that is, a downregulation in the active pursuit of concerns related to advancement and growth (Kasch, Rottenberg, Arnow, & Gotlib, 2002). Furthermore, appraising maladaptive rumination as aversive and uncontrollable could create a prevention goal of stopping the rumination (i.e., keeping bad things from happening), causing an upregulation in prevention concerns. As individuals enter a prevention-focused state, the strategic inclination to use vigilance to monitor for potential mistakes may generate a need to continue

searching the environment for ways to reduce the discrepancy or understand why they keep failing. Paradoxically, this vigilance also would make a person more likely to confront cues in the environment associated with existing discrepancies related to both promotion and prevention concerns that trigger further maladaptive rumination (Wenzlaff & Wegner, 2000).

Integrating models of self-regulation with models of repetitive thought would enhance the current understanding of cognitive processes underlying self-regulatory failure. In addition, this kind of integrative model suggests possible targets for cognitive, behavioral, and interpersonal interventions in the treatment of depression. To the extent that perceived goal failure, individual differences in regulatory focus, and rumination contribute to the onset and maintenance of a depressive syndrome, they could be addressed directly via existing therapeutic techniques (such as are common in cognitive therapy for depression [Beck, 1995]).

Biased Information Processing in Depression

In the past 15 years, social neuroscience has emerged as a distinct branch of social psychology that has far-reaching implications for clinical psychological research (Cacioppo & Berntson, 2004). Social neuroscience incorporates methodological advances in social psychology, cognitive psychology, and cognitive neuroscience to study social processes from a range of perspectives. The methods of social neuroscience, in combination with social psychological theory, can shed important new light on the nature of depression. In particular, we suggest that using functional imaging techniques to test theories about changes in information processing associated with depression could help to increase our knowledge about how depression emerges and could ultimately provide an empirical basis for treatment selection. We illustrate this proposal by discussing a current controversy in the depression literature: the nature of the cognitive changes observed as part of a depressive episode.

As stated originally in Akiskal and McKinney's (1973) "unified hypothesis" and observed in subsequent research that has combined biological and behavioral approaches to depressive disorders, depression emerges from a psychobiological "final common pathway" that produces an array of different depressive symptoms, such as memory impairment, negative cognitions, and depressed mood (Akiskal & McKinney, 1973, p. 290). There is a large but inconsistent literature documenting changes in performance on memory tasks among clinically depressed individuals (e.g., Barry, Naus, & Rehm, 2004; Blaney, 1986; Williams & Scott, 1988). Much of that literature is based on an implicit (and occasionally explicit) assumption that depression is an illness that leads to irreversible structural changes in the brain (e.g., Rompre, Stip, & Trudeau, 2004). However, Akiskal and McKinney (1973) had hypothesized that the core of depression was a "reversible functional derangement of the mechanisms of reinforcement" (p. 291), so that at least in early episodes, depression was more accurately characterized as a downregulated motivational state rather than

as an "illness." Consistent with Akiskal and McKinney's hypothesis, a recent review of relevant neurophysiological and cognitive research proposed that biased information processing is the primary cause of the memory deficits observed in depression (McLean & Strauman, 2006). From a clinical standpoint, determining whether cognitive changes associated with depression are functional (i.e., biased versions of "normal" processing) or structural (i.e., actual changes to the central nervous system) has important implications, because cognitive biases, unlike structural deficits, can at least potentially be altered to restore the individual's capacity for more adaptive information processing strategies. Consequently, strategies designed to reinstate premorbid social information processing may be a key component in successful treatment for depression.

Given the nature of this research question—does depression involve functional or structural changes in the central nervous system (CNS)—combining cognitive, behavioral, and functional neuroimaging techniques can provide essential information regarding the psychobiological changes that have been hypothesized to underlie depression. There are at least two sets of research findings that are relevant to this question: how emotional states affect memory encoding in depressed individuals (see Davidson, 1999, for a review), and how self-referential information is initially processed and/or recalled (Macrae, Moran, Heatherton, Banfield, & Kelley, 2004). To date, there are no neuroimaging studies that have explored the effects of combined negative emotional and self-referential processing in nondepressed individuals on memory (by examining encoding or retrieval), much less compared these findings with those of depressed individuals. This is unfortunate because insight into the cognitive and neurophysiological mechanisms that underlie a preference toward negative and negative self-referential information in depression may suggest new and more effective treatments for the disorder.

Despite the apparently inconsistent findings regarding memory in depressed individuals, a large body of research has shown that negative information is remembered better than positive and neutral information in depressed individuals, and that negative information is remembered especially well when it refers to the self (see Symons & Johnson, 1997, for review). Other research has shown that negative cognitions and negative self-referential information are remembered better in depressed compared to nondepressed individuals, which is hypothesized to create and maintain over time a bias toward negative self-referential information (Dent & Teasdale, 1988). This information processing bias is significant because an overemphasis on negative cognitions and negative self-referential information can lead to or maintain symptoms that in turn perpetuate depression (Williams, 1997). The persistence of negatively distorted cognitions maintains depressed mood and thereby predicts episode severity and duration as well as likelihood of relapse (Ingram & Holle, 1992).

Based on the available literature, we hypothesize that depressed individuals remember negative and negative

self-referential information more frequently than other kinds of information because they have developed a bias toward this information. We also predict that this bias is a manifestation of a functional change (i.e., a down-regulated motivational and emotional state), not an indication of a structural change in the CNS. While acknowledging the possibility that chronic recurrent depression could indeed lead to permanent changes in cortical and subcortical structures, we believe that this is the exception rather than the rule. We note in particular that our prediction implies that under certain conditions, depressed individuals will show better memory performance than nondepressed individuals—a finding that would be inconsistent with a structural or "disease" model. However, we also acknowledge that the mechanisms underlying this bias remain unknown.

One way to test this hypothesis would be to use neuro-imaging techniques to observe the activation patterns associated with negative versus positive and self-referential versus non-self-referential social information. In particular, patterns of activation in the anterior cingulate cortex (ACC) are likely to be informative, given the emerging literature implicating the ACC as underlying the processing of both emotionally valenced and self-referential information (Drevets, 2000). We would predict that depressed and nondepressed groups (when appropriately matched on age and other variables that can influence performance on cognitive tasks) would not manifest an overall difference in memory performance. Rather, whereas nondepressed individuals would show better recall or recognition for positively valenced self-referential stimuli, their depressed counterparts would better remember self-referential stimuli that were negatively valenced. In turn, these behavioral differences would likely be associated with detectable differences in location and intensity of cortical activation within the ACC. This kind of research illustrates how social psychological models, translated into an individual differences framework, can be applied to studies of the etiology, diagnosis, and treatment of major mental disorders.

Social Psychology and Psychotherapy

We close with a topic that may represent wishful thinking more than a definitive prediction. As researchers with identities as both social and clinical psychologists, we find it disappointing that contemporary social psychology has not been more closely involved with the study of psychotherapy outcome and process. With occasional exceptions (e.g., Cacioppo, Claiborn, Petty, & Heesacker, 1991; MacDonald, Nail, & Levy, 2004), social psychologists have not concerned themselves with studying how psychotherapy works—despite the fact that in previous eras, social psychological knowledge was considered essential to a thorough understanding of patient–therapist interactions (Frank & Frank, 1993). We propose that the efficacy of psychotherapy (particularly, but not limited to, cognitive-behavioral interventions) can be enhanced by taking into consideration the tools and models of social cognition (Merrill & Strauman, 2004). Several examples relating to depression are offered below to illustrate

how social psychologists can enhance both the efficacy of psychotherapy and clinician's understanding of how it works.

From the perspective of a therapist, what kinds of models or theories of behavior are most likely to be useful? Certainly one that incorporates individual differences, integrates the various components of an individual's personality system, takes into account situational variation, provides explanatory/predictive power for an individual's behavior, and, ultimately, provides information about the individual that could be applied within the framework of the therapy (e.g., identifying targets for change or compensation, contributing to better understanding of change processes, and predicting short-term and longer-term treatment outcome). Trait-based theories have not provided a useful framework for selecting or designing psychotherapeutic interventions. However, knowing what beliefs and attitudes typically influence that individual's interpretation of everyday life situations and trigger problematic conclusions or behaviors, or how affectively laden memories can influence change processes, *would* be of value in psychotherapy (Cacioppo & Berntson, 1992; Libby, Eibach, & Gilovich, 2005).

Social psychology has a long tradition of research in the area of *cognitive style*, and although several investigators have discussed the parallel development of cognitive theories in social, cognitive, and clinical psychology (e.g., Hollon & Garber, 1990), there has been relatively little crossover from these areas. One potential mechanism underlying poor outcomes seen in patients with personality disorders is an overly simplistic and rigid cognitive style. Social psychologists have explored constructs such as *need for cognition* (Cacioppo & Petty, 1982) that address stylistic differences in information processing. Understanding the interplay between the individual's cognitive style and other components of the personality system and the social environment should improve the therapist's understanding of the patient's problems as well as difficulties that arise in therapy. Individual differences in cognitive style also may influence the choice of therapeutic interventions and provide a basis for treatment matching.

Another example of an important aspect of social cognition applicable to psychotherapy is the individual's characteristic self-regulation. People differ in their customary styles of self-regulation, the ongoing cognitive process of evaluating oneself in relation to one's goals, and changing behaviors in an effort to attain goals (Carver & Scheier, 1990). One model of depression proposes that self-regulation plays an important role in the development and maintenance of the disorder (Strauman, 2002). Underlying an individual's self-regulatory processes are two hypothetical cognitive-motivational systems, which are focused on attaining positive outcomes or "making good things happen" (promotion) and on avoiding negative outcomes or "keeping bad things from happening" (prevention), respectively (Higgins, 1997). The prevention and promotion systems are hypothesized to have neural underpinnings associated with the behavioral inhibition system (BIS) and the behavioral approach system (BAS), respectively. Depres-

sion is thought to be characterized by decreased goal-directed behaviors, coupled with low levels of positive affect, which are associated with a deficit in the BAS (Watson, Weise, Vaidya, & Tellegen, 1999). Studies using brain-imaging techniques have shown that approach motivation (a BAS function) and positive affect may be associated with left frontal activation (Tomarken & Keener, 1998). Application of a self-regulation framework could offer a new means of assessing how treatments lead to symptom improvement and behavior change (Merrill & Strauman, 2004).

The inclusion of a self-regulation focus in therapy is the topic of ongoing investigation. Strauman and colleagues have developed self-system therapy (SST) (Vieth et al., 2003) a brief structured treatment for depression derived from theories of individual differences in self-regulation. In an initial randomized clinical trial, SST was more effective than a standard treatment for the subset of depressed individuals with significant problems in self-regulation such as high levels of self-discrepancy (Strauman et al., 2006). Future studies in this area aim to determine whether self-regulation may be a useful factor for treatment matching, by developing methods to identify which depressed individuals have significant problems in self-regulation.

A final area of overlap between social psychology and psychotherapy research is *person × situation interactions*. Both researchers and therapists share an interest in understanding the situational specificity of affective responses and behavioral tendencies. The interaction between situation and person is likely to be complex, dynamic, and mediated by a number of social-cognitive variables. The situation itself is a much less powerful predictor of clinically relevant behavior than the interaction between the situation and within-person variables such as beliefs, expectancies, and attitudes (Cervone, 1997). This notion of a dynamic process linking situational, cognitive, affective, and behavioral factors is consistent with the clinical literature. For example, the importance of cognitive variables in mediating the association between life events and depression has been acknowledged (Robins & Block, 1988).

CONCLUSION

Contrary to the unfortunately narrow and reductionistic worldview that dominates current headlines, social psychology is alive and well, particularly as it enriches clinical psychology. We hope that this chapter has illustrated both the critical translational principle by which social psychology influences the study of psychological disorders and the individual-differences pathway by which that influence often occurs. Clearly no single "lens" or perspective can do justice to this active research frontier; clinical psychologists will continue to pick and choose from the social psychological literature, and social psychologists will continue to develop models of the etiology, treatment, and prevention of mental disorders. If our brief glimpse of the future has any validity, there is much to be excited about.

The principles of social psychology that are summarized in this volume are all, in theory, both conceptually and empirically sound bases for understanding human distress. We have presented here three perspectives on how social psychology has been applied to clinical issues. All three perspectives translate the "common person" framework that characterizes social psychological research into an individual-differences structure to explain why some people, under certain circumstances, will experience significant distress. Clinical psychologists are to be congratulated on their willingness to borrow liberally from the social literature—or even simply for their opportunistic appropriation of interesting models. In turn, social psychology should be recognized for the extensive contributions it has made to the nation's public health. Individually and together, social and clinical psychologists are sure to continue their productive and occasionally surprising exploration of human behavior.

REFERENCES

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory based subtype of depression. *Psychological Review*, *96*, 358–372.

Adolphs, R. (2003). Investigating the cognitive neuroscience of social behavior. *Neuropsychologia*, *41*, 119–126.

Akiskal, H. S., & McKinney, W. T., Jr. (1973). Depressive disorder: Toward a unified hypothesis. *Science*, *12*, 20–29.

Alden, L. E., & Taylor, C. T. (2004). Interpersonal processes in social phobia. *Clinical Psychology Review*, *24*, 857–882.

Allport, G. W. (1985). The historical background of social psychology. In G. Lindzey & E. Arsonson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 1–46). New York: Random House.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, C. A., Krull, D. S., & Weiner, B. (1996). Explanations: Processes and consequences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 271–296). New York: Guilford Press.

Andrews, B., & Brown, G. W. (1993). Self-esteem and vulnerability to depression: The concurrent validity of interview and questionnaire measures. *Journal of Abnormal Psychology*, *102*, 565–572.

Andrews, B., & Brown, G. W. (1995). Stability and change in low self-esteem: The role of psychosocial factors. *Psychological Medicine*, *25*, 23–32.

Ban, T. A. (2004). Neuropsychopharmacology and the genetics of schizophrenia: A history of the diagnosis of schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *28*, 753–762.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. New York: Oxford University Press.

Barry, E. S., Maus, M. J., & Rehm, L. P. (2004). Depression and implicit memory: Understanding mood-congruent memory bias. *Cognitive Therapy and Research*, *28*, 387–414.

Baumeister, R. F. (1998). The self. In S. T. Fiske & D. T. Gilbert (Ed.), *The handbook of social psychology* (Vol. 1, 4th ed., pp. 680–740). New York: McGraw-Hill.

Beck, A. T. (2002). Prisoners of hate. *Behaviour Research and Therapy*, *40*, 209–216.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York: Guilford Press.

Beevers, C. G., & Meyer, B. (2004). Thought suppression and depression risk. *Cognition and Emotion*, *18*, 859–867.

Bell-Dolan, D., & Anderson, C. A. (1999). Attributional processes:

An integration of social and clinical psychology. In M. R. Leary & R. M. Kowalski (Eds.), *The social psychology of emotional and behavioral problems: Interfaces of social and clinical psychology* (pp. 37–67). Washington, DC: American Psychological Association.

Bennett, D. A., & Cooper, C. L. (1999). Eating disturbance as a manifestation of the stress process: A review of the literature. *Stress Medicine*, *15*, 167–182.

Beutler, L. E., Alomohamed, S., Moleiro, C., & Romanelli, R. (2002). Systemic treatment selection and prescriptive therapy. In F. W. Kaslow (Ed.), *Comprehensive handbook of psychotherapy: Integrative/eclectic* (Vol. 4, pp. 255–271). New York: Wiley.

Bitran, S., & Barlow, D. H. (2004). Etiology and treatment of social anxiety: A commentary. *Journal of Clinical Psychology*, *60*, 881–886.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*, 229–246.

Boon, B., Stroebe, W., Schut, H., & Ijntema, R. (2002). Ironic processes in the eating behaviour of restrained eaters. *British Journal of Health Psychology*, *7*, 1–10.

Brown, G. W., Andrews, B., Harris, T., Adler, Z., & Bridge, L. (1986). Social support, self-esteem, and depression. *Psychological Medicine*, *16*, 813–831.

Brown, G. W., & Harris, T. (1978). *Social origins of depression*. London: Tavistock.

Brownell, K. D. (1991). Dieting and the search for the perfect body: Where physiology and culture collide. *Behavior Therapy*, *22*, 1–12.

Brune, M. (2005). Emotion recognition, "theory of mind," and social behavior in schizophrenia. *Psychiatry Research*, *133*, 135–147.

Cacioppo, J. T., & Berntson, G. G. (2004). Social neuroscience. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed., pp. 977–985). Cambridge, MA: MIT Press.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multi-level integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin, 126*, 829–843.

Cacioppo, J. T., Claiborn, C. D., Petty, R. E., & Heesacker, M. (1991). General framework for the study of attitude change in psychotherapy. In C. R. Snyder & D. R. Forsyth (Eds.), *Handbook of social and clinical psychology: The health perspective* (pp. 523–539). New York: Pergamon Press.

Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual Review of Psychology, 50*, 191–214.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.

Carter, M., & Flesher, S. (1995). The neurosociology of schizophrenia: Vulnerability and functional disability. *Psychiatry*, *58*, 209–224.

Carver, C. S. (1996). Goal engagement and the human experience. In R. S. Wyer, Jr. (Ed.), *Ruminative thoughts: Vol. 11. Advances in social cognition* (pp. 49–61). Mahwah, NJ: Erlbaum.

Carver, C. S., & Scheier, M. R. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Carver, C. S., & Scheier, M. R. (1998). *On the self-regulation of behavior.* New York: Cambridge University Press.

Carver, C. S., & Scheier, M. R. (1999). Themes and issues in the self-regulation of behavior. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition.* Mahwah, NJ: Erlbaum.

Cash, T. F. (2004). Body image: Past, present, and future. *Body Image*, *1*, 1–5.

Cash, T. F., Therieault, J., & Annis, N. M. (2004). Body image in an interpersonal context: Adult attachment, fear of intimacy, and social anxiety. *Journal of Social and Clinical Psychology*, *23*, 89–103.

Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology*, *78*, 158–172.

Cervone, D. (1997). Social-cognitive mechanisms and personality coherence: Self-knowledge, situational beliefs, and cross-situational coherence in perceived self-efficacy. *Psychological Science*, *8*(1), 43–50.

Cheavens, J. S., Rosenthal, M. Z., Daughters, S. B., Nowak, J., Kosson, D., Lynch, T. R., et al. (2005). An analogue investigation of the relationships among perceived parental criticism, negative affect, and borderline personality disorder features: The role of thought suppression. *Behaviour Research and Therapy*, *43*, 257–268.

Clark, D. A., Beck, A. T., & Alford, B. A. (1999). *Scientific foundations of cognitive theory and therapy of depression*. New York: Wiley.

Coles, M. E., & Heimberg, R. G. (2002). Memory biases in the anxiety disorders: Current status. *Clinical Psychology Review*, *22*, 587–627.

Combs, D. R., & Penn, D. L. (2004). The role of subclinical paranoia on social perception and behavior. *Schizophrenia Research*, *69*, 93–104.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74–101.

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593–623.

Currin, L., Schmidt, U., Treasure, J., & Jick, H. (2005). Time trends in eating disorder incidence. *British Journal of Psychiatry*, *186*, 132–135.

Cuthbert, B. N. (2002). Social anxiety disorder: Trends and translational research. *Biological Psychiatry*, *51*, 4–10.

Davidson, R. J. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, *3*(1), 11–21.

Dawes, R. M. (1994). Psychological measurement. *Psychological Review, 101*, 278–281.

Dent, J., & Teasdale, J. D. (1988). Negative cognition and persistence of depression. *Journal of Abnormal Psychology*, *97*, 29–34.

Dishion, T. J., Capaldi, D., Spracklen, K. M., & Li, F. (1995). Peer ecology of male adolescent drug use. *Development and Psychopathology*, *7*, 803–824.

Dodge, K. A. (1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Review of Psychology*, *44*, 559–584.

Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., et al. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, *74*, 374–393.

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology*, *39*, 349–371.

Donaldson, S. I., Sussman, S., MacKinnon D. P., Severson, H. H., Glynn, T., Murray, D. M., et al. (1996). Drug abuse prevention programming: Do we know what works? *American Behavioral Scientist*, *39*, 868–883.

Drevets, W. C. (2000). Neuroimaging studies of mood disorders. *Biological Psychiatry*, *48*, 813–829.

Dweck, C. S., Higgins, E. T., & Grant-Pillow, H. (2003). Self-systems give unique meaning to self variables. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 239–252). New York: Guilford Press.

Eccles, J. S., Early, D., Frasier, K., Belansky, E., & McCarthy, K. (1997). The relation of connection, regulation, and support for autonomy to adolescents' functioning. *Journal of Adolescent Research*, *12*, 263–286.

Emmons, R. A., King, L. A., & Sheldon, K. M. (1993). Goal conflict and the self-regulation of action. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control* (pp. 528–551). Englewood Cliffs, NJ: Prentice-Hall.

Fischer, S., Anderson, K. G., & Smith, G. T. (2004). Coping with distress by eating or drinking: Role of trait urgency and expectancies. *Psychology of Addictive Behaviors*, *18*, 269–274.

Fister, S. M., & Smith, G. T. (2004). Media effects on expectancies: Exposure to realistic female images as a protective factor. *Psychology of Addictive Behaviors*, *18*, 394–397.

Flanagan, E. H., & Blashfield, R. K. (2000). Essentialism and a folk-

taxonomic approach to the classification of psychopathology. *Philosophy, Psychiatry, and Psychology*, 7, 183–189.

Flom, P. L., Friedman, S. R., Jose, B., & Curtis, R. (2001). Peer norms regarding drug use and drug selling among household youth in a low-income "drug-supermarket" urban neighborhood. *Drugs: Education, Prevention, and Policy*, 8, 219–232.

Förster, J. (2003). The influence of approach and avoidance motor actions on food intake. *European Journal of Social Psychology*, 33, 339–350.

Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Baltimore: Johns Hopkins University Press.

Frederickson, B. L., Roberts, T., Noll, S. M., Quinn, D. M., & Twenge, J. M. (1998). That swimsuit becomes you: Sex differences in self-objectification, restrained eating, and math performance. *Journal of Personality and Social Psychology*, 75, 269–284.

Freeman, L. M. Y., & Gil, K. M. (2004). Daily stress, coping, and dietary restraint in binge eating. *International Journal of Eating Disorders*, 36, 204–212.

Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, 34, 385–389.

Frith, C. D., & Frith, U. (1991). Elective affinities in schizophrenia and childhood autism. In P. E. Bebbington (Ed.), *Social psychiatry: Theory, methodology, and practice* (pp. 65–88). New Brunswick, NJ: Transaction.

Frost, J., & McKelvie, S. (2004). Self-esteem and body satisfaction in male and female elementary school, high school, and university students. *Sex Roles*, 51, 45–54.

Fuchs, T. (2004). Neurobiology and psychotherapy: An emerging dialogue. *Current Opinion in Psychiatry*, 17, 479–485.

Garcia-Grau, E., Fuste, A., Miro, A., Saldana, C., & Bados, A. (2002). Coping style and disturbed eating attitudes in adolescent girls. *International Journal of Eating Disorders*, 32, 116–120.

Gibb, B. E., Alloy, L. B., Abramson, L. Y., Beevers, C. G., & Miller, I. W. (2004). Cognitive vulnerability to depression: A taxometric analysis. *Journal of Abnormal Psychology*, 113, 81–89.

Gibbons, F. X., Gerrard, M., & Lane, D. J. (2003). A social reaction model of adolescent health risk. In J. M. Suls & K. Wallston (Eds.), *Social psychological foundations of health and illness* (pp. 107–136). Oxford, UK: Blackwell.

Gramzow, R. H., Sedikides, C., Panter, A. T., & Insko, C. A. (2000). Aspects of self-regulation and self-structure as predictors of perceived emotional distress. *Personality and Social Psychology Bulletin*, 26, 188–205.

Granholm, E., McQuaid, J. R., McClure, F. S., Auslander, L. A., Perivoliotis, D., Pedrelli, P., et al. (2005). A randomized, controlled trial of cognitive behavioral social skills training for middle-aged and older outpatients with chronic schizophrenia. *American Journal of Psychiatry*, 162, 520–529.

Graziano, W. G., Jensen-Campbell, L. A., & Finch, J. F. (1997). The self as a mediator between personality and adjustment. *Journal of Personality and Social Psychology*, 73, 392–404.

Greeley, J., & Oei, T. (1999). Alcohol and tension reduction. In K. E. Leonard & H. T. Blane (Eds.), *Psychological theories of drinking and alcoholism* (2nd ed., pp. 14–53). New York: Guilford Press.

Green, M. F., & Braff, D. L. (2001). Translating the basic and clinical cognitive neuroscience of schizophrenia to drug development and clinical trials of antipsychotic medications. *Biological Psychiatry*, 49, 374–384.

Green, M. F., Nuechterlein, K. H., Gold, J. M., Barch, D. M., Cohen, J., Essock, S., et al. (2004). Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH–MATRICS conference to select cognitive domains and test criteria. *Biological Psychiatry*, 56, 301–307.

Hardy, A., & Brewin, C. R. (2005). The role of thought suppression in the development of obsessions. *Behavioural and Cognitive Psychotherapy*, 33, 61–69.

Harrison, P. A., Fulkerson, J. A., & Park, E. (2000). The relative importance of social versus commercial sources in youth access to tobacco, alcohol, and other drugs. *Preventive Medicine: An International Devoted to Practice and Theory*, 31, 39–48.

Harvey, P. D., Green, M. F., Keefe, R. S. E., & Velligan, D. I. (2004). Cognitive functioning in schizophrenia: A consensus statement on its role in the definition and evaluation of effective treatments for the illness. *Journal of Clinical Psychiatry*, 65, 361–372.

Hayes, A. M., Harris, M. S., & Carver, C. S. (2004). Predictors of self-esteem variability. *Cognitive Therapy and Research*, 28, 369–385.

Heatherton, T. F., & Baumeister, R. F. (1991). Binge eating as escape from self-awareness. *Psychological Bulletin*, 110, 86–108.

Heatherton, T. F., Herman, C. P., & Polivy, J. (1992). Effects of distress on eating: The importance of ego-involvement. *Journal of Personality and Social Psychology*, 62, 801–803.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Herman, C. P., Fitzgerald, N. E., & Polivy, J. (2003). The influence of social norms on hunger ratings and eating. *Appetite*, 41, 15–20.

Herman, C. P., Roth, D. A., & Polivy, J. (2003). Effects of the presence of others on food intake: A normative interpretation. *Psychological Bulletin*, 129, 873–886.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52, 1280–1300.

Higgins, E. T. (2000). Social cognition: Learning about what matters in the social world. *European Journal of Social Psychology*, 30, 3–39.

Higgins, E. T., & Moretti, M. M. (1988). Standard utilization and the social-evaluative process: Vulnerability to types of aberrant beliefs. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs*. New York: Wiley.

Hogarty, G. E., & Flesher, S. (1999). A developmental theory for a Cognitive Enhancement Therapy of schizophrenia. *Schizophrenia Bulletin*, 25, 677–692.

Hogarty, G. E., Flesher, S., Ulrich, R., Carter, M., Greenwald, D., Pogue-Geile, M., et al. (2004). Cognitive enhancement therapy for schizophrenia: Effects of a 2-year randomized trial on cognition and behavior. *Archives of General Psychiatry*, 61, 866–876.

Holden, C. (2004). NIMH takes a new tack, upsetting behavioral researchers. *Science*, 306, 602.

Hollon, S. D., & Garber, J. (1990). Cognitive therapy for depression: A social cognitive perspective. *Personality and Social Psychology Bulletin*, 16(1), 58–73.

Holtzman, P. S. (2003). Less is truly more: Psychopathology research in the 21st century. In J. Hooley & M. Lenzenweger (Eds.), *Principles of experimental psychopathology: Essays in honor of Brendan A. Maher* (pp. 175–193). Washington, DC: American Psychological Association.

Hoyle, R. H., Kernis, M. H., Leary, M. R., & Baldwin, M. W. (1999). *Selfhood: Identity, esteem, regulation*. Boulder, CO: Westview Press.

Ingram, R. E. & Holle, C. (1992). Cognitive science of depression. In D. Stein & J. Young (Eds.), *Cognitive science and clinical disorders* (pp. 187–209). New York: Academic Press.

Kandel, D. B., & Adler, I. (1982). Socialization into marijuana use among French adolescents: A cross-cultural comparison with the United States *Journal of Health and Social Behavior*, 23, 295–309.

Kaplow, J. B., Curran, P. J., Dodge, K. A., & Conduct Problems Research Group. (2002). Child, parent, and peer predictors of early-onset substance use: A multisite longitudinal study. *Journal of Abnormal Child Psychology*, 30, 199–216.

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology*, 44, 23–52.

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology*, 111, 589–597.

Kelemen, O., Keri, S., Must, A., Benedek, G., & Janka, Z. (2004). No evidence for impaired "theory of mind" in unaffected first-degree relatives of schizophrenia patients. *Acta Psychiatrica Scandinavica*, 110, 146–149.

Kemeny, M. E., Gruenewald, T. L., & Dickerson, S. S. (2004).

Shame as the emotional response to threat to the social self: Implications for behavior, physiology, and health. *Psychological Inquiry, 15,* 153–160.

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry, 159,* 1133–1139.

Kernis, M. H., Whisenhunt, C. R., Waschull, S. B., Greenier, K. D., Berry, A. J., Herlocker, C. E., et al. (1998). Multiple facets of self-esteem and their relations to depressive symptoms. *Personality and Social Psychology Bulletin, 24,* 657–669.

Klein, S. B., German, T. P., Cosmides, L., & Gabriel, R. (2004). A theory of autobiographical memory: Necessary components and disorders resulting from their loss. *Social Cognition, 22,* 460–490.

Kurtz, M. M., Moberg, P. J., Gur, R. C., & Gur, R. E. (2001). Approaches to cognitive remediation of neuropsychological deficits in schizophrenia: A review and meta-analysis. *Neuropsychology Review, 11,* 197–224.

Lancaster, R. S., Evans, J. D., Bond, G. R., & Lysaker, P. H. (2003). Social cognition and neurocognitive deficits in schizophrenia. *Journal of Nervous and Mental Disease, 191,* 295–299.

Leary, M. R. (2001). Social anxiety as an early warning system: A refinement and extension of the self-presentation theory of social anxiety. In P. M. DiBartolo & S. G. Hofmann (Eds.), *From social anxiety to social phobia: Multiple perspectives* (pp. 321–334). Needham Heights, MA: Allyn & Bacon.

Leary, M. R., & Kowalski, R. M. (Eds.). (1999). *The social psychology of emotional and behavioral problems: Interfaces of social and clinical psychology.* Washington, DC: American Psychological Association.

Lee, K. H., Farrow, T. F., Spence, S. A., & Woodruff, P. W. (2004). Social cognition, brain networks, and schizophrenia. *Psychological Medicine, 34,* 391–400.

Libby, L. K., Eibach, R. P., & Gilovich, T. (2005). Here's looking at me: The effect of memory perspective on assessments of personal change. *Journal of Personality and Social Psychology, 88,* 50–62.

Linehan, M. M. (2000). Commentary on innovations in dialectical behavior therapy. *Cognitive and Behavioral Practice, 7,* 478–481.

Luthar, S. S., & McMahon, T. J. (1996). Peer reputation among inner-city adolescents: Structure and correlates. *Journal of Research on Adolescence, 6,* 581–603.

Lynam, D. R., Milich, R., Zimmerman, R., Novak, S. P., Logan, T. K., Martin, C., et al. (1999). Project DARE: No effects at 10-year follow-up. *Journal of Consulting and Clinical Psychology, 67,* 590–593.

Lyubomirsky, S., Tucker, K. L., Caldwell, N. D., & Berg, K. (1999). Why ruminators are poor problem solvers: Clues from the phenomenology of dysphoric rumination. *Journal of Personality and Social Psychology, 77,* 1041–1060.

MacDonald, G., Nail, P. R., & Levy, D. A. (2004). Expanding the scope of the social response context model. *Basic and Applied Social Psychology, 26,* 77–92.

Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., & Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex, 14,* 647–654.

Mann, T., & Ward, A. (2001). Forbidden fruit: Does thinking about a prohibited food lead to its consumption? *International Journal of Eating Disorders, 29,* 319–327.

Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer, Jr. (Ed.), *Ruminative thoughts: Vol. 11, Advances in social cognition* (pp. 1–47). Mahwah, NJ: Erlbaum.

Martin, L. L., & Tesser, A. (1989). Toward a motivational and structural theory of ruminative thought. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 306–326). New York: Guilford Press.

McLean, A. N., & Strauman, T. J. (2006). *An integrative review of memory deficits in depression: Do observed deficits contribute to depressive cognitive biases?* Manuscript under review.

Merrill, K. A., & Strauman, T. J. (2004). The role of personality in cognitive-behavioral therapies. *Behavior Therapy, 35,* 131–146.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004).

Is there a universal positivity bias in attributions?: A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin, 130,* 711–747.

Miller-Johnson, S., & Costanzo, P. (in press). If you can't beat 'em . . . induce them to join you: Peer-based interventions during adolescence. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's peer relations: From development to intervention to policy: A festschrift in honor of John D. Coie.* Washington, DC: American Psychological Association.

Millon, T. (2004). Biomedical explanations of psychopathology, the DSM, and essential necessities for advancement of diagnostic classification: An essay review. *Journal of Psychotherapy Integration, 14,* 106–119.

Mills, J. S., Polivy, J., Herman, C. P., & Tiggemann, M. (2002). Effects of exposure to thin media images: Evidence of self-enhancement among restrained eaters. *Personality and Social Psychology Bulletin, 28,* 1687–1699.

Moretti, M. M., Higgins, E. T., & Feldman, L. (1990). The self-system in depression: Conceptualization and treatment. In C. D. McCann & N. S. Endler (Eds.), *Depression: New directions in research, theory, and practice* (pp. 127–156). Toronto: Wall & Thompson.

Morrison, T. G., Kalin, R., & Morrison, M. A. (2004). Body-image evaluation and body-image investment among adolescents: A test of sociocultural and social comparison theories. *Adolescence, 39,* 259–270.

Muehrer, P. R., Salovey, P., Afifi, A. A., Coyne, J. C., Kring, A. M., Merson, M. H., et al. (2002). Overcoming barriers to collaboration between basic behavioral scientists and public health scientists in research on mental disorders. *Journal of Clinical Psychology in Medical Settings, 9,* 253–259.

Muth, J. L., & Cash, T. F. (1997). Body-image attitudes: What difference does gender make? *Journal of Applied Social Psychology, 16,* 1438–1452.

National Advisory Mental Health Council. (1988). *Approaching the 21st century: Opportunities for NIMH neuroscience research. Report to Congress on the decade of the brain* (DHHS Publication No. ADM 89-1580). Washington, DC: U.S. Government Printing Office.

National Institute of Drug Abuse. (1999). *Drug abuse and addiction research: 25 years of discovery to advance the health of the public* (6th triennial report to Congress). Washington, DC: U.S. Secretary of Health and Human Services.

National Institute of Mental Health. (1999). *Translating behavioral science into action: Report of the National Advisory Mental Health Council Behavioral Science Workgroup.* Washington, DC: Author.

Newton, E., Landau, S., Smith, P., Monks, P., Shergill, S., & Wykes, T. (2005). Early psychological intervention for auditory hallucinations: An exploratory study of young people's voices groups. *Journal of Nervous and Mental Disease, 193,* 58–66.

Nolen-Hoeksema, S. (1996). Chewing the cud and other ruminations. In R. S. Wyer, Jr. (Ed.), *Ruminative thoughts: Advances in social cognition* (Vol. 11, pp. 135–144). Mahwah, NJ: Erlbaum.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109,* 504–511.

Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and the duration of episodes of depressed mood. *Journal of Abnormal Psychology, 102,* 20–28.

Olds, R. S., & Thombs, D. L. (2001). The relationship of adolescent perceptions of peer norms and parent involvement to cigarette and alcohol use. *Journal of School Health, 71,* 223–228.

Papageorgiou, C., & Wells, A. (2004). Nature, functions, and beliefs about depressive rumination. In C. Papageorgiou & A. Wells (Eds.), *Depressive rumination: Nature, theory, and treatment* (pp. 10–40). Chichester, UK: Wiley.

Penn, D. L., Corrigan, P. W., Bentall, R. P., Racenstein, J. M., & Newman, L. (1997). Social cognition in schizophrenia. *Psychological Bulletin, 121,* 114–132.

Penn, D. L., Mueser, K. T., Tarrier, N., Gloege, A., Cather, C., Serrano, D., et al. (2004). Supportive therapy for schizophrenia:

Possible mechanisms and implications for adjunctive psychosocial treatments. *Schizophrenia Bulletin*, *30*, 101–112.

Perry, C. L., Williams, C. L., Veblen-Mortenson, S., & Toomey, T. L. (1996). Project Northland: Outcomes of a community wide alcohol use prevention program during early adolescence. *American Journal of Public Health*, *86*, 956–965.

Peterson, C., & Seligman, M. E. (1987). Explanatory style and illness. *Journal of Personality*, *55*, 237–265.

Pinkham, A. E., Penn, D. L., Perkins, D. O., & Lieberman, J. (2003). Implications of the neural basis of social cognition for the study of schizophrenia. *American Journal of Psychiatry*, *160*, 815–824.

Polivy, J., & Herman, C. P. (2002). Causes of eating disorders. *Annual Review of Psychology*, *53*, 187–213.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515–526.

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and Representing Problems. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 3–29). New York: Cambridge University Press.

Price, R. H., Gioci, M., Penner, W., & Trautlein, B. (1993). Webs of influence: School and community programs that enhance adolescent health and education. In R. Takanishi (Ed.), *Adolescence in the 1990's: Risk and opportunity* (pp. 29–63). New York: Teachers College Press.

Purdon, C., Rowa, K., & Antony, M. M. (2005). Thought suppression and its effects on thought frequency, appraisal and mood state in individuals with obsessive–compulsive disorder. *Behaviour Research and Therapy*, *43*, 93–108.

Pyszczynski T., & Greenberg, J. (1987). Self-regulatory preservation and the depressive focusing style: A self-awareness theory or reactive depression. *Psychological Bulletin*, *102*, 122–138.

Reischer, E., & Koo, K. S. (2004). The body beautiful: Symbolism and agency in the social world. *Annual Review of Anthropology*, *33*, 297–317.

Roberts, J. E., & Gamble, S. A. (2001). Current mood state and past depression as predictors of self-esteem and dysfunctional attitudes among adolescents. *Personality and Individual Differences*, *30*, 1023–1037.

Roberts, J. E., & Gotlib, I. H. (1997). Temporal variability in global self-esteem and specific self-evaluation as prospective predictors of emotional distress: Specificity in predictors and outcome. *Journal of Abnormal Psychology*, *106*, 521–529.

Roberts, J. E., & Monroe, S. M. (1994). A multidimensional model of self-esteem in depression. *Clinical Psychology Review*, *14*, 161–181.

Roberts, J. E., & Monroe, S. M. (1999). Vulnerable self-esteem and social processes in depression: Toward an interpersonal model of self-esteem regulation. In J. C. Coyne & T. Joiner (Eds.), *The interactional nature of depression: Advances in interpersonal approaches* (pp. 149–187). Washington, DC: American Psychological Association Press.

Roberts, S. B., & Kendler, K. S. (1999). Neuroticism and self-esteem as indices of vulnerability to major depression in women. *Psychological Medicine*, *29*, 1101–1109.

Robins, C. J. (2002). Zen principles and mindfulness practice in dialectical behavior therapy. *Cognitive and Behavioral Practice*, *9*, 50–57.

Robins, C. J., & Block, P. (1988). Personal vulnerability, life events, and depressive symptoms: A test of a specific interactional model . *Journal of Personality and Social Psychology*, *54*(5), 847–852.

Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2000). Heterogeneity of popular boys: Antisocial and prosocial configurations. *Developmental Psychology*, *36*, 14–24.

Roemer, L., & Salters, K. (2004). A preliminary study of the effects of directed suppression of rape-related material among rape survivors using unobtrusive measures. *Behavioural and Cognitive Psychotherapy*, *32*, 149–164.

Rompre, P.-P., Stip, E., & Trudeau, L.-E. (2004). Neurobiology of severe mental disorders: From cell to bedside. *Journal of Psychiatry and Neuroscience*, *29*, 167–169.

Rotenberg, K. J., Carte, L., & Speirs, A. (2005). The effects of modeling dietary restraint on food consumption: Do restrained models promote restrained eating? *Eating Behaviors*, *6*, 75–84.

Ruble, D. N., Costanzo, P. R., & Higgins, E. T. (1992). Social psychological foundations of mental health. In D. N. Ruble, P. R. Costanzo, & M. E. Oliveri (Eds.), *The social psychology of mental health: Basic mechanisms and applications* (pp. 1–23). New York: Guilford Press.

Ruble, D. N., Costanzo, P. R., & Oliveri, M. E. (Eds.). (1992). *The social psychology of mental health: Basic mechanisms and applications*. New York: Guilford Press.

Rude, S. S., & McCarthy, C. T. (2003). Emotional functioning in depressed and depression-vulnerable college students. *Cognition and Emotion*, *17*, 799–806.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*, 1835–1842.

Schiffman, J., Lam, C. W., Jiwatram, T., Ekstrom, M., Sorensen, H., & Mednick, S. (2004). Perspective-taking deficits in people with schizophrenia-spectrum disorders: A prospective investigation. *Psychological Medicine*, *34*, 1581–1586.

Schmid, H. (2001). Predictors of cigarette smoking by young adults and readiness to change. *Substance Use and Misuse*, *36*, 1519–1542.

Segerstrom, S. C., Stanton, A. L., Alden, L. E., & Shortridge, B. E. (2003). A multidimensional structure for repetitive thought: What's on your mind, and how, and how much? *Journal of Personality and Social Psychology*, *85*, 909–921.

Seiving, R. E., Perry, C. L., & Williams, C. L. (2000). Do friendships change behaviors, or do behaviors change friendships? Examining paths of influence in young adolescents' alcohol use. *Journal of Adolescent Health*, *26*, 27–35.

Sergi, M. J., & Green, M. F. (2002). Social perception and early visual processing in schizophrenia. *Schizophrenia Research*, *59*, 233–241.

Showers, C. J., Limke, A., & Zeigler-Hill, V. (2004). Self-structure and self-change: Applications to psychological treatment. *Behavior Therapy*, *35*, 167–184.

Siegfried, Z., Berry, E. M., Hao, S., & Avraham, Y. (2003). Animal models in the investigation of anorexia. *Physiology and Behavior*, *79*, 39–45.

Snyder, C. R., Tennen, H., Affleck, G., & Cheavens, J. (2000). Social, personality, clinical, and health psychology tributaries: The merging of a scholarly "river of dreams." *Personality and Social Psychology Review*, *4*, 16–29.

Southall, D., & Roberts, J. E. (2002). Attributional style and self-esteem in vulnerability to adolescent depressive symptoms following life stress: A 14-week prospective study. *Cognitive Therapy and Research*, *26*, 563–579.

Stanley, M. A., & Beck, J. G. (2000). Anxiety disorders. *Clinical Psychology Review*, *20*, 731–754.

Steptoe, A., Perkins-Porras, L., McKay, C., Rink, E., Hilton, S., & Cappuccio, F. P. (2003). Psychological factors associated with fruit and vegetable intake and with biomarkers in adults from a low-income neighborhood. *Health Psychology*, *22*, 148–155.

Stice, E. (1994). Review of the evidence for a sociocultural model of bulimia nervosa and an exploration of the mechanisms of action. *Clinical Psychology Review*, *14*, 633–661.

Stice, E., & Shaw, H. E. (2002). Role of body dissatisfaction in the onset and maintenance of eating pathology: A synthesis of research findings. *Journal of Psychosomatic Research*, *53*, 985–993.

Stice, E., & Whitenton, K. (2002). Risk factors for body dissatisfaction in adolescent girls: A longitudinal investigation. *Developmental Psychology*, *38*, 669–678.

Strauman, T. J. (2002). Self-regulation and depression. *Self and Identity*, *1*, 151–157.

Strauman, T. J., & Kolden, G. G. (1997). The self in depression: Research trends and clinical implications. *In Session*, *3*, 5–21.

Strauman, T. J., & Merrill, K. A. (2004). The basic science/clinical science interface and treatment development. *Clinical Psychology: Science and Practice*, *11*, 263–266.

Strauman, T. J., & Segal, Z. V. (2001). The cognitive self in basic science, psychopathology, and psychotherapy. In J. C. Muran (Ed.), *Self-relations in the psychotherapy process* (pp. 241–266). Washington, DC: American Psychological Association.

Strauman, T. J., Vieth, A. Z., Merrill, K. A., Woods, T. E., Kolden, G. G., Klein, M. H., et al. (2006). Self-system therapy as an intervention for self-regulatory dysfunction in depression: A randomized comparison with cognitive therapy. *Journal of Consulting and Clinical Psychology, 74*, 367–376.

Strobe, W. (2000). *Social psychology and health*. Buckingham, UK: Open University Press.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*, 371–394.

Tennen, H., & Affleck, G. (2002). The challenge of capturing daily processes at the interface of social and clinical psychology. *Journal of Social and Clinical Psychology*, *21*, 610–627.

Thome, J., & Espelage, D. L. (2004). Relations among exercise, coping, disordered eating, and psychological health among college students. *Eating Behaviors*, *5*, 337–351.

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). Future directions: Integrative theories, multidimensional assessments, and multicomponent interventions. In L. J. Heinberg & J. K. Thompson (Eds.), *Exacting beauty: Theory, assessment, and treatment of body image disturbance* (pp. 311–332). Washington, DC: American Psychological Association.

Tomarken, A. J., & Keener, A. D. (1998). Frontal brain asymmetry and depression: A self-regulatory perspective. *Cognition and Emotion*, *12*, 387–420.

Trower, P., Birchwood, M., Meaden, A., Byrne, S., Nelson, A., & Ross, K. (2004). Cognitive therapy for command hallucinations: Randomised controlled trial. *British Journal of Psychiatry*, *184*, 312–320.

Turkington, D., Dudley, R., Warman, D. M., & Beck, A. T. (2004). Cognitive-behavioral therapy for schizophrenia: A review. *Journal of Psychiatric Practice*, *10*, 5–16.

van den Berg, P., Thompson, J. K., Obremski-Brandon, K., & Coovert, M. (2002). The Tripartite Influence model of body image and eating disturbance: A covariance structure modeling investigation testing the mediational role of appearance comparison. *Journal of Psychosomatic Research*, *53*, 1007–1020.

Vander Wal, J. S., & Thomas, N. (2004). Predictors of body image dissatisfaction and disturbed eating attitudes and behaviors in African-American and Hispanic girls. *Eating Behaviors*, *5*, 291–301.

Van Wolputte, S. (2004). Hang on to your self: Of bodies, embodiment, and selves. *Annual Review of Anthropology*, *33*, 251–269.

Vieth, A., Strauman, T. J., Kolden, G., Woods, T., Michels, J., & Klein, M. H. (2003). Self-system therapy: A theory-based psycho-therapy for depression. *Clinical Psychology: Science and Practice*, *10*, 245–268.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249–255.

Vohs, K. D., Heatherton, T. F., & Herrin, M. (2001). Disordered eating and the transition to college: A prospective study. *International Journal of Eating Disorders*, *29*, 280–288.

Wagner, E. R., Myers, M. G., & McIninch, J. L. (1999). Stress-coping and temptation-coping as predictors of adolescent substance use. *Addictive Behaviors*, *24*, 769–779.

Warr, M. (1996). Organization and instigation in delinquent groups. *Criminology*, *34*, 11–37.

Watson, D., Weise, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838.

Webb, C., Scudder, M., Kaminer, Y., & Kadden, R. (2002). *The motivational enhancement therapy and cognitive behavioral therapy supplement: 7 sessions of cognitive behavioral therapy for adolescent cannabis users* (Cannabis Youth Treatment (CYT) Series, Volume 2). Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health Services Administration.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5–13.

Wegner, D. M., & Wenzlaff, R. M. (1996). Mental control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 466–492). New York: Guilford Press.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*, 54873.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology*, *51*, 59–91.

Widiger, T. A. (2001). What can be learned from taxometric analyses? *Clinical Psychology: Science and Practice*, *8*, 2001–2005.

Widiger, T. A. (2005). Classification and diagnosis: Historical development and contemporary issues. In B. A. Winstead & J. E. Maddux (Eds.), *Psychopathology: Foundations for a contemporary understanding* (pp. 63–83). Mahwah, NJ: Erlbaum.

Williams, J. M. G. (1997). Cognitive impairment, depression, and the specificity of autobiographical memory in the elderly. *British Journal of Clinical Psychology*, *36*, 341–347.

Williams, J. M. G., & Scott, J. (1988). Autobiographical memory in depression. *Psychological Medicine*, *18*(3), 689–695.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Wynn, J. K., Sergi, M. J., Dawson, M. E., Schell, A. M., & Green, M. F. (2005). Sensorimotor gating, orienting, and social perception in schizophrenia. *Schizophrenia Research*, *73*, 319–325.

Consumer Behavior and Marketing

Eric J. Johnson
Michel Tuan Pham
Gita Venkataramani Johar

Whenever we make a decision to buy—or not to buy—something, there are short-term and often long-term consequences. In the short-term, we may consume a sandwich, go to a movie, or buy some clothing. We also have made a short-term decision to allocate resources, such as time and money, to that acquisition. That purchase may have long-term benefits and costs: That sandwich may contribute to being overweight; the movie may change the way we look at some aspect of the world; and the clothing may earn us a compliment or help us get a new job. Less obvious is that this decision involves trade-offs. For example, the resources required to make the purchase will prevent us from pursuing other opportunities with the same resources, what economists call *opportunity costs.*

These kinds of decisions are at the heart of consumer behavior. Because they are not made in isolation, we must also be concerned how that environment is influenced by firms and marketing managers. Although this chapter concentrates on *consumers*, we also talk about *managers* because they help design the products, allocate the advertising budget, and shape the messages that influence consumers. Their lay and expert theories of consumer choice help establish the environment in which consumers behave. In addition, one important role of consumer behavior research is to help managers develop better theories and make better decisions. Finally, many of the principles we discuss have relevance to public pol-

icy and, in fact, social marketing, a field that would like to employ these principles in not-for-profit arenas.

CONSUMER BEHAVIOR AS AN ACADEMIC FIELD

Consumer research is located at the nexus of theory and practice and at the intersection of several different research traditions. Two of the primary source disciplines are psychology and economics, although methodology, particularly quantitative modeling from statistics and management science, also plays an important role.

One way of portraying the field and its relationship to social psychology is to look at the pattern of citations among major journals. Figure 38.1 shows the pattern of cross citation for three psychology journals and four of the leading journals in marketing. The numbers are the percentage of all citations in that journal that refer to papers published in another. For example, 7% of the citations in the *Journal of Consumer Research* (JCR) are to papers in the *Journal of Personality and Social Psychology* (JPSP), while 20% of JCR's citations are to its own papers. For clarity, we eliminated cross citations with frequencies of less than 2%.[1]

One strong conclusion that can be drawn is that the central journal in consumer research is a large consumer of social psychology: Seven out of every 100 citations are
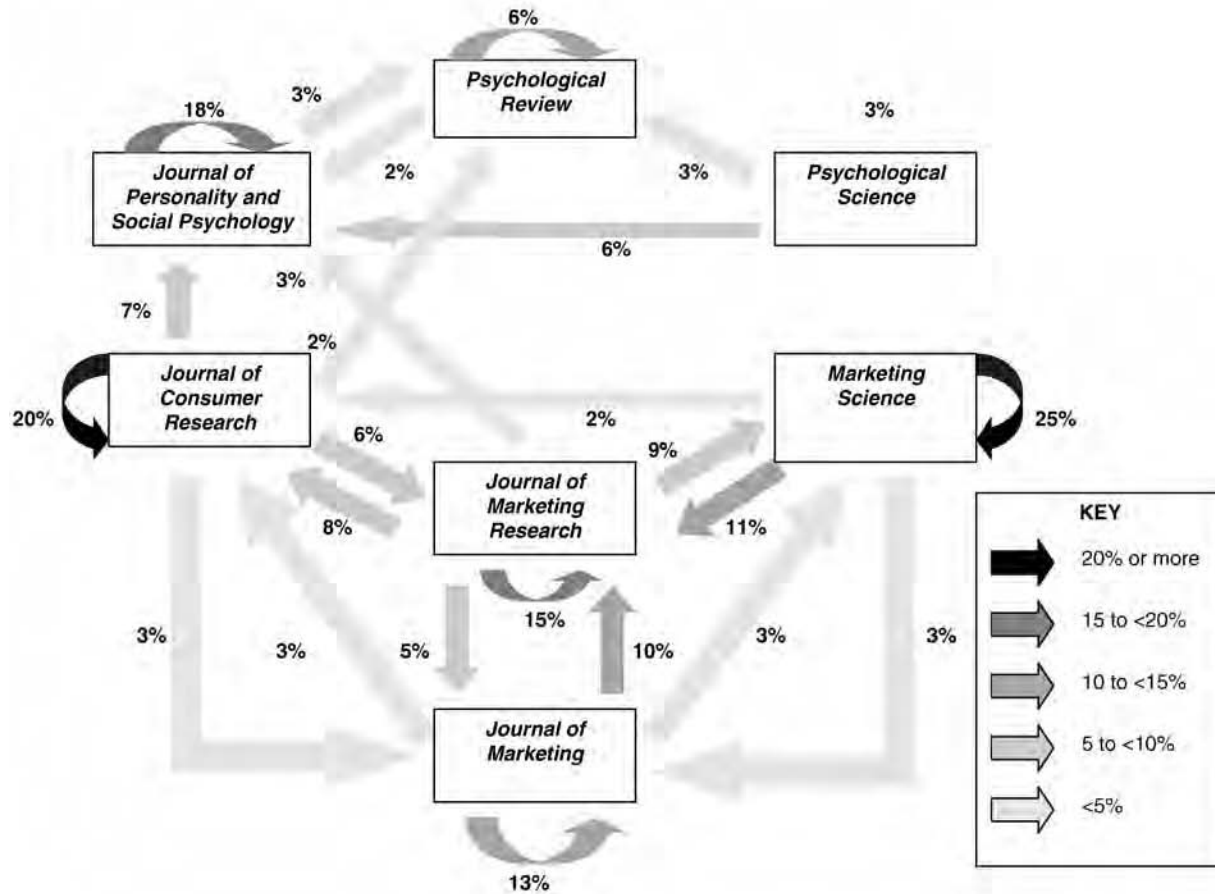
FIGURE 38.1. Cross-citations of journals, psychology and marketing, based on *ISI Web of Knowledge* data, 2004. Arrows indicate citations from one journal to papers published in another. Any cross-citations less than 2% have been deleted for clarity.

to JPSP. That is even more than to any of the marketing journals other than itself. A second conclusion is that authors in JPSP find little to cite in any marketing journal. It is as if the phenomena and findings of the world of commerce are of little relevance to social psychology. While that may be the fate of any applied discipline, we are convinced that this should not be the case. Many of the questions in consumer research are, in fact, of great relevance to social psychology, and many important empirical results should be of interest to students of fields such as persuasion and decision making. Consumer behavior provides an ideal laboratory, not just for the application of theories but also for the extending existing theoretical ideas and the development of new ones.

A third conclusion that might be drawn from this diagram is that consumer research and marketing have their own schisms. Roughly speaking, there is a limited flow of citations to and from the more quantitative journals, such as *Marketing Science*, and one might conclude that consumer research is fairly distant from either *Marketing Science* or the *Journal of Marketing*, the latter often seen as the most applied journal in the discipline. In sum, Figure 38.1 shows the reader that what may appear a monolithic field is, in reality, like many outgroups, less homogeneous than it appears, and that social psychology is an important source for, but not a user of, consumer-oriented research.

GOALS OF THE CHAPTER

Our goal in this chapter is to review for social psychologists some of the interesting research done in consumer behavior and marketing. This review must be incomplete in the usual sense, that is, many good papers and many worthwhile research areas are not covered because of space constraints and the lack of fit to a relatively arbitrary structure. However, it is incomplete in a more intentional and strategic way. Because many excellent chapters in this handbook describe the basic research underlying consumer and marketing research, we do not review those basic ideas and results here. Our goal, instead, is to review the results of the application of psychological ideas to consumer behavior and marketing and to highlight data and ideas that should inform social-psychological research. The emphasis is on research published outside the traditional mainstream of social psychology but within the mainstream of consumer research

and marketing, and it addresses concepts, theories, and methods relevant to social and cognitive psychology.

To organize this literature, and to give the reader a sense of the last few decades of the field, we have chosen to concentrate on four topics: two representing classic areas of inquiry, and two representing developing ones. Our "old school" classics are topics that have been central to consumer research for at least the last 50 years: decision making and persuasion. This is not to say that they are not hotbeds of current research but rather that they have a longer traditional history and more established methods. Our "new school" topics are affect and implicit processing, areas that have grown rapidly in consumer research, paralleling trends in social psychology.

DECISION PROCESSES

Consumer Decision Making: A View from Economics

Imagine a consumer in front of a typical American supermarket aisle, choosing a breakfast cereal. There are, by most counts, at least 120 different options in a typical American supermarket. How does a consumer choose? We start by introducing an "as-if" model, termed "value maximization," from economics (see Deaton & Muellbauer, 1980, for a classic perspective; Tversky & Kahneman, 1991, and Tversky & Simonson, 1993, for a psychological view; and Varian, 1992, for a more recent view). We do this both because (1) it stands as a normative benchmark for other decision processes, and (2) it is the source, in spirit at least, of many of the models that managers use to predict consumer choice and design new products. In this model, consumers behave as if they were examining all relevant attributes for all the cereals. That is, they choose as if they knew all available relevant information. Not only do they pick the "best" cereal according to these criteria from this large set, but they are also maximizing their choice across all possible consumption choices they are making, both now and given their best predictions about what will happen in the future. This maximization is not limited to the supermarket, but to *all possible* consumption decisions!

However unrealistic this may sound, the basic idea of value maximization has generated many useful tools such as discrete choice modeling, which managers use to design, price, and position products. Most attempts to make these models more realistic involve modifications to the idea of value maximization rather than the use of alternative frameworks that are based on other principles or assumptions.

Choice Heuristics and Representations

Heuristics for Choice

In response to the apparently unrealistic demands of value maximization, many scholars have developed descriptions of choice processes that simplify the decision process. These simplifications from value maximization, or *choice heuristics*, try to maintain the ability to make

good choices at substantial savings of effort. One example of a choice heuristic in our cereal example would be if the consumer were to eliminate any cereal with added sugar. This corresponds to an elimination by aspects heuristic (Tversky, 1972), and it simplifies the choice because the brands that are eliminated are not examined further. Heuristics save effort by ignoring information. This savings comes with a potential cost because it is possible, if unlikely, that one of the eliminated cereals is so much better on the remaining attributes (e.g., lots of vitamins and very inexpensive) that it is better than the one finally chosen.

Many heuristics have been described (Svenson, 1979), including those that are based on comparisons of the alternatives, such as the lexicographic and additive differences heuristics, and those, like elimination by aspects (Tversky, 1972), that are based on the comparison of the alternatives to a standard, which are related to Simon's notion of satisficing (Simon, 1955). Do people use these heuristics? An extensive literature examines how people make choices, many using process tracing methods, such as talk-aloud protocols (Bettman, 1970), recording of consumers' head and eye movements (Russo & Leclerc, 1994; Russo & Rosen, 1974), and other ways of observing information acquisitions (Bettman & Kakkar, 1977; Jacoby, 1975; Jacoby et al., 1994). This literature clearly indicates that a large number of heuristics are used, that they depend on the characteristics of the choice, and that people switch heuristics, even in the course of a single decision. In fact, it has been argued that the research on choice processes should be conducted at a finer level of analysis, and that the concept of heuristics is too broad to use to understand consumer choice (Payne, Bettman, & Johnson, 1991) and that a lower level of analysis is appropriate. This plethora of potential strategies, while descriptively more accurate, poses a challenge for modeling consumer behavior: How can we try to predict what a consumer would want when we do not know how they will make a choice? This represents an active area of research in consumer choice modeling.

Mental Accounting

Recall that our shopper not only maximized, according to theory, across all breakfast cereals but across all choices. Thus, according to theory, shoppers are deciding between buying the brand-name corn flakes and the store-brand flakes in light of its implications for other purchases that may occur years in the future, such as buying a retirement home. A more psychologically realistic view is that people have much more restricted "mental" accounts in which trade-offs are made (Thaler, 1984, 1999; Thaler & Johnson, 1990). For our hypothetical cereal consumer, they may be making trade-offs within their mental accounts for breakfasts, trying to find the best possible combinations *within* that category, including hot breakfasts, the bagel bought at a corner kiosk, and so on.

While this psychologically plausible assumption markedly simplifies the consumer's task, it is not without its costs. One major concern is that consumers might be

maximizing within that account, but that this leads to over- or underconsumption in other accounts and that the quality of consumer's mental accounting depends on their ability to remember to "post" expenditures into each account and whether they compare expenditures in different accounts (Heath & Soll, 1996). One particularly nice application of this idea is studying how consumers make trade-offs between time and money (Leclerc, Schmitt, & Dube, 1995; Soman, 2001).

Another way in which mental accounting affects consumer choices is that mental accounting, like "real" accounting, occurs over time. That is, inflows of consumption and outflows of resources (like money) can occur at different times. We pay for cars over many years after we initially purchase them but usually pay for vacations before we take them. An interesting growth area in consumer research tries to understand how consumers would want these transactions framed, and how they might be manipulated by frames. Thus, studies have examined whether decision makers like to have the costs and benefits of decisions put together (integrated) or kept apart (segregated) (Prelec & Loewenstein, 1998; Read, Loewenstein, & Rabin, 1999; Thaler & Johnson, 1990).

An important application of this kind of mental accounting research is to look at how the way people pay for transactions affects their spending. The concept of payment decoupling (Soman & Gourville, 2001) suggests that if expenditures are in different accounts than consumption, consumers will spend more. This has been demonstrated in recent studies (Prelec & Simester, 2001; Soman, 2003) that show, for example, that people are willing to pay up to 100% more to purchase tickets to a basketball game when using a charge card than when paying in cash, and that people remember how much they spend better when they pay with cash than with a credit card. Soman and Cheema (2002) look at how mental accounting affects consumers' use of a line of credit.

Recent findings by Zhou and Pham (2004) suggest that nonprofessional consumer investors use two separate mental accounts to manage their investments: one account, associated with promotion-focused regulation (Higgins, 1997), is used to for the achievement of financial gains, and another account, associated with prevention-focused regulation, is used to for the prevention of financial losses. Because consumers learn to associate various financial products with either the achievement of gains or the prevention of losses, they tend to evaluate different investment products using different criteria, which violate standard finance and economics principles.

A major challenge in introducing the idea of mental accounts is identifying the boundaries of mental accounts. Does our hypothetical consumer include the bagel bought at work in the same account as the one for breakfast cereal, or maybe in an account of things bought to get to work early? Some initial work has looked at using natural categories as a starting point (Henderson & Peterson, 1992; McGraw, Tetlock, & Kristel, 2003), but much work remains to be done, particularly if mental accounting is to influence quantitative models of consumer choice.

Reference Dependence and Loss Aversion

Once consumers are seen as making decisions in more narrow accounts, it seems natural to adopt a different view of how they interpret attributes. Under value maximization, when our cereal consumer was judging calories, he or she was making trade-offs against all possible consumption decisions. For example, the calorie consumption of bagel might be compared to a crème brûlée the consumer might consume 2 weeks from Sunday. A more natural and cognitively economical assumption is to see the consumer as comparing the calories of a potential new cereal relative to the calories accompanying his or her current favorite brand.

This basic idea of a reference point is reflected most famously as the value function of prospect theory (Tversky & Kahneman, 1991). Along with the idea of loss aversion, the idea of reference dependence has had a major impact on consumer research. For example, imagine that our cereal consumer now sees that his favorite cereal has had a price increase (to him, a loss) of $0.10. How does that compare to a price decrease (to them a gain) of $0.10? According to the value function of prospect theory, the loss (price increase) will have much more impact; on average twice as much of an impact as would the same size price increase. In terms of economics, we would expect the *elasticity* to change: The decrease in consumption caused by the price increase should be twice the size of the increase in consumption caused by a price decrease.

Early work in studying consumer choice first examined, like our example, whether reactions to price were reference dependent (Kalyanaram & Little, 1994; Mayhew & Winer, 1992; Winer, 1986) and whether they exhibited loss aversion (Hardie, Johnson, & Fader, 1993; Putler, 1992). They provided evidence that reference dependence and loss aversion were useful in explaining the effect of price in consumer choice (see Mazumdar, Raj, & Sinha, 2005, for a recent review). Following the multiattribute extension of the prospect theory value function, Hardie and colleagues (1993) showed that there was loss aversion for both price and, to an even greater extent, quality (see also Bell & Lattin, 2000; Heath et al., 2000; Sen & Johnson, 1997). Subsequent work has looked at the amount of loss aversion for various attributes and whether there are systematic differences across individuals (Chernev, 2004; Erdem, Mayhew, & Sun, 2001; Klapper, Ebling, & Temme, 2005; Van Dijk & Van Knippenberg, 2005). Research has also examined how the pattern of purchases affects the reference point and preference (Sood, Rottenstreich, & Brenner, 2004; Wathieu, 2004). One interesting application examines reference effects and loss aversion in online auctions (Dholakia & Simonson, 2005).

Context Effects

Perhaps the best evidence against value maximization is the existence of context effects. A context effect occurs when adding alternatives to a choice set changes what is chosen, even when the added option seems largely irrelevant. To explore this, consider a website offering two air-

fares to the same city. One is a direct, more expensive flight, and the second a flight requiring a change of planes but significantly cheaper. Depending on one's tastes, one flight or the other might be chosen. Consider what happens if that site added another flight—one that is almost identical to the one-stop in every way but more expensive. Should this change peoples' choices?

Both common sense and value maximization say that choices should not change, but empirically they do. Adding the option that is the same in every way but more expensive typically increases the share of the better one-stop flight. This observation, termed "asymmetric dominance," or the *attraction effect* (Huber, Payne, & Puto, 1982), is one of the findings in consumer choice with the greatest impact throughout the social sciences. Other context effects exist, most notably a *compromise* effect. Here the option between two extreme options gets a greater share of choices than would be predicted by value maximization (Simonson, 1989). In our example, adding a fast, more expensive flight to our Web site would increase the share of the now less expensive nonstop flight. People have speculated whether these techniques could be used to manipulate consumers (Hamilton, 2003; Stewart, Chater, Stott, & Reimers, 2003).

What causes these rather bizarre effects? Answering this question seems to be particularly challenging as these effects have been found in other species such as honeybees and gray jays (Shafir, Bechar, & Weber, 2003; Shafir, Waite, & Smith, 2002). One common set of explanations makes use of loss aversion (Tversky & Simonson, 1993), suggesting that unfavorable comparisons have more impact on choice. Thus, the middle or compromise option has smaller disadvantages, relative to the two options. This leads to the middle option being more attractive, because its losses are smaller. In a very provocative paper, Kivetz, Netzer, and Srinivasan (2004a) develop several different models for the compromise effect and show one in which the middle option serving as a reference point provides a better account for the observed data. Importantly, they also generalize the result, extending it to choices where there are more than two alternatives and two attributes (Kivetz, Netzer, & Srinivasan, 2004b). While this is a major step forward, in part because it presents a managerially useful model of this particular context effect that can be estimated, the real challenge for further research is to develop theoretical mechanisms that unite a growing number of effects.

Task Effects

While context effects refer to the kinds of options presented to the consumer, task effects refer to the way in which the consumer must make these choices. One example might be the time pressure in the supermarket caused by an impatient young child; another might be the need to justify the decision to one's supervisor. The options remain the same, but the conditions of evaluation are different.

The classic task effect in psychology is the preference reversal between pairs of options, often two gambles. Here respondents will choose one option but be willing to pay more for the other, violating economic ideas of stable preferences (Lichtenstein & Slovic, 1971). Its close relative in consumer choice is the difference between joint and separate evaluations (Hsee, 1996; Hsee & Leclerc, 1998). In joint evaluation, two options are judged at the same time, while in separate evaluation, they are presented and judged one at a time. However, the mechanisms involved in both joint and separate evaluations are different than those involved in the classic reversal among gambles. For this type of reversal, the critical variable is the ease with which an attribute can be evaluated. Joint evaluation can make variables that are hard to evaluate easy to compare, thereby increasing their weight. One important question raised in this research is which gives better predictions of future experience? Hsee and Zhang (2004) argue that because most options are experienced in isolation, joint evaluation can yield to errors.

Another task factor that has a significant effect is the need to justify one's choice to others. This can change attribute weights for some attributes (Okada, 2005) and affects the frequency of some kinds of context effects and decision errors (Simonson, 1989, 1992; Simonson & Nye, 1992). Similarly, and somewhat counterintuitively, there are conditions in which being more thoughtful can increase a context effect (Priester, Dholakia, & Fleming, 2004). Thus, working harder does not mean working smarter.

WHERE DO VALUES COME FROM?: INFERENCE AND MEMORY

Until now, we have rather blithely assumed that consumers get the information needed to make a decision directly and immediately from the world around them, for example, from store shelves, or from websites. Of course, this is a gross simplification. In this section, we look at two very important sources of information about the options chosen by consumers—memory and inference. A third important source of value results from the fit between the decision maker and the choice process. We do not discuss this in depth here, because it is presented in the excellent review in this volume

Memory-Based versus External Search

It is clear that the use of the external environment can be expensive, in terms of effort, compared to retrieving information from memory. This has led consumer researchers to distinguish between information that comes from internal and external search (Alba & Hutchinson, 1987; Biehal & Chakravarti, 1986a; Johnson & Russo, 1984; Sanbonmatsu & Fazio, 1990)—a distinction that has close relations in the social cognition literature (Hastie & Park, 1986; Srull & Wyer, 1989). A significant literature in consumer research parallels similar queries in the social-psychological literature on person memory and social cognition more generally (e.g., Fiske & Taylor, 1991; Higgins, Kuiper, & Olson, 1981). Among the significant results in this area are the ideas that consumer knowl-

edge is organized mostly around brands (Biehal & Chakravarti, 1986b; Johnson & Russo, 1978) rather than attributes; that task goals play an important role in the encoding of information (Biehal & Chakravarti, 1982; Sujan, 1985); and that consumers tend to remember the results of decisions rather than the data on which they are based (Johnson & Russo, 1978; Park & Hastak, 1994). Research has also asked a question about the form of consumer's recall of chosen products: Is recall biased in favor of the chosen alternative? The data seem clear that such biases exist, and they follow from both cognitive and, perhaps, motivational factors. Because consumers often use heuristics like elimination by aspects, which focus on a few alternatives (Grether & Wilde, 1984; Wright, 1975), recall will naturally be better for chosen alternatives (Biehal & Chakravarti, 1986; Costley & Brucks, 1992; Dick, Chakravarti, & Biehal, 1990; Johnson & Russo, 1984). However, as it appears that choice often creates distortions in the valuation of the chosen alternatives (Posavac, Sanbonmatsu, Kardes, & Fitzsimons, 2004; Russo, Meloy, & Medvec, 1998; Simon, Krawczyk, & Holyoak, 2004) , memory is subsequently distorted as well (see Ross & Sicoly, 1979, for an analogy in the social psychological literature).

Category Inference

What happens when information needed to make a choice is not available, either in the environment or in memory? This can occur, for example, when one encounters a new product—for example, a new sport utility vehicle from a known manufacturer, say Saab. One stream of research in consumer behavior examines whether consumers use inferences in these situations. Consumers seem quite comfortable inferring characteristics of the product based on their categorization of the product. Early work in this area concentrated on potentially ambiguous new products, which could be characterized as belonging to two or more categories: Is a new fruit-flavored carbonated children's drink a healthy juice, or a soft drink? This work (Sujan & Dekleva, 1987) showed that consumers used two stages in inference, the first to determine how to categorize the new product ("It's a soft-drink") and the second to infer values ("The fruit flavor is probably artificial"). Similarity between the product and category seems to be an important mediator in determining how a product is categorized (Park, Milberg, & Lawson, 1991; Viswanathan & Childers, 1999), and it differs across age groups (John & Sujan, 1990). Sometimes these inference processes can lead us astray, when, for example, marketers add features that differentiate a product in a way that is meaningless to product performance but changes the way products are categorized (Carpenter, Glazer, & Nakamoto, 1994).

Brand Inference

A special type of inference that has attracted much attention is how consumers think about brands. Because brands are an important organizing principle in consumers' product knowledge, they would also seem to play an important role in product inference. In fact, a major theme of the last two decades of consumer and marketing research concerns the values of brands, termed "brand equity." Much work attempts to assess the economic value to the firm of consumers' awareness and associations with a brand (Aaker & Keller, 1990; Keller, 1993).

Brand Personality

One approach to conceptualizing consumers' knowledge about brands employs the notion of a brand personality, adapting many concepts, such as the Big 5 typology of personality, and applying them to brands (Aaker, 1997). Research here has also been concerned with people's relationships with brands (Fournier, 1998), how brand personalities might differ across cultures (Aaker, Benet-Martinez, & Garolera, 2001; Sung & Tinkham, 2005), and how they are updated (Aaker, Fournier, & Brasel, 2004; Johar, Sengupta, & Aaker, 2005). While this area of research remains a provocative metaphor, an open question is how closely processing of brand information, which evolutionarily must be a relatively recent event, resembles the processing of person information. A very recent functional magnetic resonance imaging (fMRI) study indicates that the areas usually associated with the identification of individuals are not used in identifying brands (Yoon, Gutchess, Feinberg, & Polk, 2006).

Brand Extensions

A specific type of inference is required when a known brand introduces a product in a new product class, as in our Saab SUV example earlier. These *brand extensions* are particularly important to managers because they can reduce the cost of a new product introduction (Aaker & Keller, 1990; Boush & Loken, 1991). However, a major theme in this literature is that there must be a "fit" between the two. In fact, a bad brand extension not only fails but can hurt the core brand (John, Loken, & Joiner, 1998; Loken & John, 1993). The accessibility of the brand and the category also determine the success of a brand extension (Meyvis & Janiszewski, 2004).

The area of consumer inference and categorization continues to grow, and readers who want to pursue this area are directed to the excellent recent reviews by Kardes, Posavac, and Cronley (2004) and Loken (2006).

PERSUASION

Persuasion research has been a central concern to consumer and marketing research for decades. There seems to be a natural application of ideas from attitude research to advertising, and research in the area actively adapts paradigms and concepts from social psychology (for a review, see Eagly & Chaiken, 1993) to questions that are raised by advertising as a marketing instrument. Typical of these are questions about frequency of advertising, the kind of message, execution, and medium that advertisers should use, and how to deal with competitive

advertising. The fact that the persuasive intent of advertising is known to consumers suggests that advertisements may be processed differently from other types of messages typically studied in social psychology; this "schemer schema" or "marketplace metacognition" has formed the focus of a compelling body of literature in consumer behavior (Friestad & Wright, 1995).

Most research on persuasion in marketing has used a dual-process approach, such as the elaboration likelihood model (Petty, Cacioppo, & Schumann, 1983), the MODE (Fazio, 1990), or the heuristic–systemic model (Chaiken & Maheswaran, 1994; Ratneshwar & Chaiken, 1991). Consumer research has challenged the prevailing view that peripheral cues are used blindly in the absence of motivation and ability and has shown that consumers sometimes (e.g., under high arousal) perform a check on the cue and use it if it is found to be diagnostic in the decision-making context (Pham, 1996).

The critical question addressed by persuasion research that is most relevant to consumer behavior concerns whether a less thorough and effortful process or a more thorough and effortful process will be used to process the message and the consequences of such processing. Consequences such as attitude persistence and resistance to future persuasion are especially meaningful in a consumer behavior context given the competitive nature of the marketplace. Because this volume contains excellent reviews of these theoretical issues, we concentrate on the application of these ideas to consumer and marketing settings. In essence, we concentrate on variables that determine the kind and level of processing an incoming communication will receive. These particularly relevant set of antecedents and their consequences are the focus of our review.

Antecedents of Message Processing

Clearly consumers face many decisions, some important ("which car to buy"), some truly trivial ("how would you like that burger?"). In some areas they may be more able to process information than others. But how do we measure and conceptualize motivation and ability, two important antecedents of our reactions to persuasive messages? Most research has organized around two themes: what message characteristics lead to more elaboration, and what person characteristics engender elaboration?

Message Characteristics

There is a long list of message characteristics that influence processing. For example, Ahluwalia and Burnkrant (2004) have examined the role of rhetorical questions on elaboration. Priester, Godek, Nayakankuppum, and Park (2004) have shown conditions under which comparative advertising (ads that compare one product to another) can lead to elaboration. The effect of message framing, emphasizing the product's advantages or the competing product's disadvantages, interacts with the amount of elaboration (Meyers-Levy & Maheswaran, 2004). Advertisements can mention the country of origin of a product, and Gurhan-Canli and Maheswaran (2000a) have examined how these effects might differ across different cultures. Similarly these authors (Gurhan-Canli & Maheswaran, 2000b) have examined how country of origin for products may change the consumer's motivation to process.

Person Characteristics

Among other factors that have been hypothesized to affect motivation to process and elaboration are a person's commitment to prior positions (Ahluwalia, 2000), involvement in the brand or product category (Johar, 1995), and his or her cognitive capacity (Johar & Simmons, 2000). Relating the communication to one's self can increase elaboration (Burnkrant & Unnava, 1995), but such effects are moderated by other variables that increase elaboration. Maheswaran and Sternthal (1990) examine the role of both knowledge and message type in the type of processing performed by a consumer.

Other person characteristic research examines how persuasive messages may be processed differently by those who are bilinguals (Luna & Peracchio, 2001), and gender differences appear to exist in processing strategies (Meyers-Levy, 1988; Meyers-Levy & Maheswaran, 1991). Given the evidence of individual differences, one might wonder whether standard dual-process theories generalize across cultures. However, an alternative perspective is that the theory itself may well be predictive but there may be differences in the perceptions of consumers about the inputs to the process. Most research suggests that dual-process models of persuasion are robust, once differences in perceptions are included. For example, Aaker and Maheswaran (1997) showed that advertising appeals that are compatible with the self-constructs (e.g., independent vs. interdependent) that are chronically accessible in a given culture are more effective. Advertising may also be effective if it reflects self-constructs that are made temporarily accessible via priming, as long as consumers have low levels of commitment to the brand (Agrawal & Maheswaran, 2005).

Message and Person Factors

A combination of person and message characteristics, such as the fit between the message and the current orientation of the consumer, can also influence persuasion: increasing persuasion when message recipients have positive thoughts about a message and decreasing persuasion when they have negative thoughts (Cesario, Grant, & Higgins, 2004). Appeals that urge consumers to imagine the product experience can similarly increase persuasion when consumers are high on imagery ability but can decrease persuasion when consumers are low on imagery ability (Petrova & Cialdini, 2005).

Competitive Effects

One situation that marks advertising as a unique environment for persuasion is the vast quantity of messages and the low level of attention in which exposure occurs. Researchers have suggested that these ingredients may lead

to circumstances in which ads will interfere with one another, leading to reduced recall for the entire product class. Nice empirical demonstrations of these effects (Burke & Srull, 1988) have shown that it is not just the number of ads but the share of advertising for a product that increases recall. Keller (1987) provides demonstrations that unique elements of advertising can be used to generate recall of ad content at the point of purchase. Under some conditions, however, such interference in memory can be helpful to some brand (Jewell & Unnava, 2003) by making some brands *relatively* easier to recall. Meyvis and Janiszewski (2004) examine the role of competitive interference in brand extending brand names to new products.

Consequences of Message Processing

Persuasion research in consumer behavior has typically examined effects of message processing on attitude valence and extremity. More recent research examining the role of consumers' goals has shown that messages may be selectively processed in the service of ego-defensive and impression-management goals (Agrawal & Maheswaran, 2005; Ahluwalia, 2002; Jain & Maheswaran, 2000; Sengupta & Johar, 2001). Focus on motivated processing has resulted in the study of elaboration quality as a consequence of processing, rather than simply the amount or quantity of processing emphasized by early versions of the dual-process models.

Attitude Strength: Attitude–Behavior Correspondence

Sengupta and Johar (2002) showed that consumers who have the goal of forming an integrated attitude at the time of exposure to inconsistent information form strong attitudes that are predictive of later product choices. On the other hand, the goal of minimizing later embarrassment decreases attitude strength in terms of attitude-choice correspondence by increasing attitudinal ambivalence. Importantly, both goals result in the same high amount of elaboration; only elaboration quality differs.

Outcomes such as attitudinal ambivalence are important because they can in turn guide future information processing. For example, Zemborain and Johar (in press) found that consumers with ambivalent attitudes tend to be more susceptible to interpersonal influence (even from potentially unreliable sources) than less ambivalent consumers. This finding is particularly relevant given the plethora of information regarding others' opinion and attitudes that consumers have access to in today's marketing environment. The variety of recommendations and gratuitous advice on the Internet is just one example.

Attitude Persistence and Resistance

Other aspects of attitude include how long-lasting the attitude is and how resistant it is to future attack (Sengupta, Goodstein, & Boninger, 1997). These aspects are particularly relevant in the competitive marketing environment where devices such as comparative advertising are used to change consumers' attitudes positively toward the advertised brand and negatively toward the comparison brand. For example, Ahluwalia (2002) demonstrated that negative information about a brand has less impact on consumers committed to that brand than others. Sengupta and Fitzsimmons (2004) found that thinking about reasons why a consumer likes or dislikes a brand does not necessarily disrupt the formation of attitudes and reduce attitude persistence and predictive power, as has been suggested in the past. Instead, attitudes formed by a reasons analysis can be strong in the sense of being persistent if the cues underlying attitude formation are also present at a later point in time.

Summary

One observation that can be made about the consumer literature on persuasion is that it is theoretically quite sophisticated, with most studies typically showing interactions between the variables of interest and their combined impact on attitude extremity as well as strength. A shared theoretical framework has allowed researchers to produce predictions that are highly interactive in nature, suggesting that multiple factors *must* be taken into account to understand the effectiveness of a persuasive marketing communication. However, from a manager's perspective, this complexity may be a bit daunting. Research in this area seldom generates simple advice and suggests instead that answers will depend on the interaction of message, person and competitive factors.

Another observation concerns the separation of work in persuasion and decision making. Interestingly, these two established research areas in consumer research have not been tightly integrated. As much as the decision making and attitudinal literatures have remained apart in social psychology, so have they in consumer research. There are some exceptions (Priester, Nayakankuppam, Fleming, & Godek, 2004), but this separation in an applied field such as consumer research is unfortunate.

AFFECT

Research on affect and consumer behavior has grown dramatically. A search of the ISI Web of Knowledge (a social science database) for the terms "consumer" and "affect" returned 136 articles for the 1985–1994 period and 841 articles for the 1995–2004 period. This interest in affect, of course, parallels the one observed in social psychology. It also reflects the fact that marketing and consumption stimuli (e.g., products, services, and TV commercials) are often emotionally rich. For example, over half of the advertisements appearing on American television contain no facts at all about the product advertised (Resnik & Stern, 1977). Advertisers instead often rely on emotionally arousing cues (attractive models, pleasant music, powerful imagery, etc.). Not surprisingly, consumers' descriptions of their consumption experiences often reveal a substantial degree of emotional richness (Derbaix & Pham, 1991; Havlena & Holbrook, 1986; Richins, 1997).

Types of Affect

It is useful to distinguish among three types of affect in consumer judgment and decision making: integral affect, incidental affect, and task affect (Bodenhausen, 1993; Cohen, Pham, & Andrade, in press). *Integral affect* refers to affective responses that are directly linked to the object of judgment or decision. These include momentary feelings experienced through direct exposure to the object itself (e.g., the pleasant feeling of tasting a fine wine) and those experienced in response to some representation of the object—representation that may be externally provided (e.g., viewing a TV commercial for a product) or internally generated (e.g., thinking about a product). *Incidental affect* refers to affective experiences whose source is clearly unrelated to the object to be evaluated. These include mood states, emotional dispositions (e.g., chronic anxiety or depression), and contextual stimuli that are affect eliciting (e.g., background music and pleasant scent). Finally, *task affect* refers to affective responses that are elicited by the judgment or decision task itself. For example, a choice between two integrally pleasant alternatives, such as two attractive vacation destinations, may generate negative task affect because one of the options must be forgone.

Integral Affect

INTEGRAL AFFECT INFLUENCES JUDGMENT AND BEHAVIOR

A major theme of research in social psychology has been that integral affective responses to various objects (e.g., people, issues, and messages) predict judgment, choice, and behavior toward these objects over and above more descriptive ("cognitive") bases of judgments such as beliefs, stereotypes, base rates, prior attitudes, and so on (e.g., Abelson, Kinder, Peters, & Fiske, 1982; Breckler & Wiggins, 1989). In marketing and consumer research, this theme was pursued most extensively in the advertising domain. A large number of studies have documented that affective responses to advertisements have direct effects on consumers' attitudes toward the ad (A_{ad}) and at least indirect effects on consumers' attitudes toward the brand (A_b) through the effects on A_{ad} (e.g., Aaker, Stayman, & Hagerty, 1986; Brown, Homer, & Inman, 1998; Edell & Burke, 1987; Holbrook & Batra, 1987). Some studies suggest that integral affective responses to the ad may also influence A_b directly, independently of A_{ad} (Burke & Edell, 1989; Derbaix, 1995; Stayman & Aaker, 1988).

Conceptually related results have been obtained in a variety of other consumption domains. For instance, Bodur, Brinberg, and Coupey (2000) found that affect toward various AIDS prevention behaviors such as abstinence or condom usage predicted attitudes and intentions toward these behaviors over and above personal beliefs about these behaviors. Similarly, MacGregor, Slovic, Dreman, and Berry (2000) found that investment banking students' feelings toward various industry sectors (e.g., electronics and managed health care) were strongly predictive of their intentions to invest in these sectors, independent of the sectors' financial fundamen-

tals. Integral affective responses have also been found to have direct effects on product satisfaction (Oliver, 1993), blood donation behavior (Allen, Machleit, & Kleine, 1992), and pricing of gambles (Peters, Slovic, & Gregory, 2003).

Two main explanations have been proposed for these direct effects of integral affect. The first is simple evaluative conditioning (De Houwer, Thomas, & Baeyens, 2001). A close proximity between a target and an integral feeling experience may result in the evaluative meaning of the feelings (mostly their valence) being carried over to the target—a mechanism sometimes called "affect transfer" in consumer research and marketing (e.g., Mackenzie, 1986). The second mechanism is an affect-as-information process (Schwarz & Clore, 1983, 1996). A number of studies (Pham, 1998; Pham, Cohen, Pracejus, & Hughes, 2001) indicate that consumers often evaluate objects by monitoring and interpreting their integral feeling responses to these objects, a process known as the "how-do-I-feel-about-it?" heuristic (Schwarz & Clore, 1988). Regardless of the actual explanation, there is substantial evidence that integral affect has a marked influence on judgments and decisions (see also Finucane, Alhakami, Slovic, & Johnson, 2000). For example, the mere fact of being lightly touched on the arm by a waiter or waitress in a restaurant results in dramatic increases in tipping and restaurant satisfaction, presumably because being touched makes the patron feel good toward the waiter or waitress and restaurant (Hornik, 1992).

AFFECT-BASED JUDGMENTS ARE DIFFERENT

Judgments based on integral affect seem to differ from judgments based on descriptive inputs in systematic ways. Because integral affective responses often arise quickly (LeDoux, 1996; Zajonc, 1980), can enter evaluations through simple associations, and usually have unambiguous interpretations (Strack, 1992), judgment and decisions based on integral affect tend to be reached more rapidly, whether online (Pham et al., 2001) or memory based (Verplanken, Hofstee, & Janssen, 1998). Similarly, because affect-based judgments and decisions seem to require less processing resources, consumers tend to rely on integral affect more when their processing resources are constrained (Pham et al., 2001). For example, when given a choice between a tempting piece of chocolate cake (an affectively attractive option) and a healthier fruit salad (a "cognitively" attractive option), consumers whose cognitive resources were not constrained tended to choose the healthier fruit salad. However, when cognitive resources were constrained, consumers tended to choose the more tempting cake, presumably because affective drivers of preference still operated while the more cognitive drivers could not (Shiv & Fedorikhin, 1999).

Judgments and decisions based on integral affect also seem to be myopic in that immediate affective rewards and punishments are weighted much more heavily, compared to delayed affective consequences (see Loewenstein, 1996). This property is very obvious in impulse control situations where people have to trade off

the immediate hedonic consequences of an option (e.g., the pleasure of eating junk food or the pain of visiting the dentist) against its long-term consequences (e.g., high cholesterol and obesity; healthy teeth and gums). According to Loewenstein (1996), the myopia of affect-based judgments and decisions is caused by the differential accessibility of current and delayed affective states. Whereas the experience of immediate integral affect has strong drive properties, it is very difficult to vividly picture future affective states. Consistent with this proposition, recent brain-imaging studies indicate that preferences for immediate rewards are associated with greater activation in parts of the limbic system associated with affect (McClure, Laibson, Loewenstein, & Cohen, 2004).

Contrary to popular beliefs that affect is highly subjective, a growing body of evidence suggests that affective judgments are, in fact, quite consensual, sometimes even more so than cognitive judgments. For instance, it has been found that consumers are more likely to agree on their affective responses to various stimuli (e.g., magazine pictures and television commercials) than they are to agree on their reason-based assessments of the same stimuli (Pham et al., 2001). According to Pham and colleagues (2001), affect-based judgments will be highly consensual whenever they are based on hardwired programs involved in bioregulation or emotional schemas acquired through socialization. The inherent consensuality of affective responses explains why juries can agree strongly on how outraged they feel in response to legal cases even when they disagree widely on the amount of punitive damages they are willing to award (Kahneman, Schkade, & Sunstein, 1998).

Affect-based judgments and decision also seem to be insensitive to quantity. In an interesting study, Hsee and Rottenstreich (2004) manipulated the number of pandas that might be saved by donations to a rescue effort and how this number was represented. In one condition, the number of saved pandas was simply represented by one or four dots; in the other, it was represented by one or four cute pictures of pandas. As predicted, respondents' donations were much more sensitive to the number of pandas saved in the affect-poor (dot) condition than in the affect-rich (picture) condition. In a similar vein, it has been observed that affect-based evaluations are insensitive to probabilities (Loewenstein, Weber, Hsee, & Welch, 2001). For example, it was found that people were willing to pay much more to avoid a high probability of losing a certain amount of money than to avoid a low probability of losing the same amount, consistent with economic theory. However, people were *not* willing to pay much more to avoid a high probability of receiving an electric shock (a prospect rich in negative affect) than to avoid a low probability of receiving the same shock (Rottenstreich & Hsee, 2001).

A final property of evaluations and decisions based on integral affect is that they tend to have a *high degree of internal coherence* (Pham, 2004). This is because integral affective responses to a target, which are often immediate and highly accessible, usually trigger a confirmatory search for information that supports or helps explain the initial feelings (Pham et al., 2001; Yeung & Wyer, 2004).

This confirmatory search results in a strong correlation between the immediate affective response elicited by a target and the spontaneous thoughts that people associate with the target.

Incidental Affect

As in social psychology (e.g., Isen, Shalker, Clark, & Karp, 1978; Johnson & Tversky, 1983), numerous consumer research and marketing studies have shown that mood states and other forms of incidental affect generally have assimilative (affect-congruent) influences on evaluations, decisions, and behaviors (e.g., Adaval, 2001; Fedorikhin & Cole, 2004; Gorn, Goldberg, & Basu, 1993; Miniard, Bhatla, & Sirdeshmukh, 1992; Pham, 1998; Yi, 1990a). In fact, some of the earliest demonstrations of this phenomenon appeared in marketing. For example, Axelrod (1963) found that consumers who had viewed a depressing television documentary gave more negative evaluations to a variety of products than they had prior to seeing the documentary. Similarly, Dommermuth and Millard (1967) showed that viewing a pleasant or unpleasant movie later produced product ratings that were consistent with the mood induced by the film.

Consumer and marketing researchers have been particularly interested in studying marketplace implications of this phenomenon. There has been significant interest in assessing how consumers' responses to advertisements are influenced by the affective tone of the media context in which the ads appear (e.g., TV programs or magazines). It is generally found that incidental affect elicited by the media context has a congruent influence on evaluations of the ad but less influence on evaluations of the advertised brand (e.g., Goldberg & Gorn, 1987; Mathur & Chattopadhyay, 1991; Murry & Dacin, 1996). It has also been found that gift wrapping can enhance the recipient's evaluation of the gift by elevating the recipient's mood (Howard, 1992). Even browsing a series of attractive products may elevate a consumer's mood, which may become assimilated into subsequent evaluations (Raghunathan & Irwin, 2001). It has also been found that if a product that consumers find disgusting (e.g., hygienic napkins) incidentally touches another product in a shopping cart, consumers' attitudes toward the latter product become more unfavorable, even if the products are in their original, unopened packages with no real chance of physical contamination (Morales & Fitzsimons, in press).

The explanations offered for these assimilative effects of incidental affect are essentially the same as those offered to explain the direct effects of integral affect: evaluative conditioning (Gorn, 1982) and reliance on the "how-do-I-feel-about-it?" heuristic. Pham (1998) proposed that the "how-do-I-feel-about-it?" heuristic is a widely used decision strategy among consumers. His findings indicate that this strategy is more likely to be used when the consumers have experiential (hedonic) motives than when they have instrumental (utilitarian) motives. It is also more prevalent among consumers who favor a visual or sensory style of processing as opposed to a more verbal or propositional style. Pham and Avnet

(2004) observed that reliance on the "how-do-I-feel-about-it?" heuristic appears to be greater when consumers are under regulatory states of promotion focus as opposed to states of prevention focus (Higgins, 1997). Other studies suggest that incidental-affect-congruent evaluations are more likely when the target is evaluative as ambiguous and does not elicit strong integral affect (Miniard et al., 1992).

Because affective states are characterized not only by their valence but also by their arousal, consumer researchers have also examined the effects of incidental arousal on evaluations and decisions. When valence and emotional content are held constant, intense incidental arousal seems to increase risk seeking in decision making (Leith & Baumeister, 1996; Mano, 1992, 1994). Intense incidental arousal also increases consumers' reliance on diagnostic information in evaluations (Pham, 1996). Milder residual arousal has been found to interact with the integral affective tone of advertisements to amplify evaluations of these ads (Gorn, Pham, & Sin, 2001), a finding consistent with the excitation transfer hypothesis (Zillmann, 1971). It has also been found that consumers' states of relaxation versus activation are very sensitive to various retail environmental factors, which can then influence their judgments and behaviors. For example, fast-paced background music in a supermarket was found to increase in-store traffic flow and sales volumes compared to slow-paced music or no music, presumably because the fast-paced music increases the shopper's level of activation (Milliman, 1982). On the other hand, *slower*-paced music in a restaurant increased the amount of time patrons spent at the dinner table and how much they ordered from the bar, presumably because slower-paced music induces feelings of relaxation (Milliman, 1986). It has also been found that the use of blue colors in Web pages produces feelings of relaxation that decrease perceived download time, whereas the use of red colors produces feelings of tension that increase perceived download time (Gorn, Chattopadhyay, Sengupta, & Tripathi, 2004).

A recent theme both in consumer research and in social psychology has been to examine the differential effects of specific incidental emotions (e.g., Raghunathan & Pham, 1999; Tiedens & Linton, 2001). Raghunathan and Pham (1999) found that, in choices between high-risk/high-reward and low-risk/low-reward options, sad individuals consistently favor the former, whereas anxious individuals consistently favor the latter. This is presumably because, even though their states are incidental, sad individuals tend to infer that they have lost something of value (a typical cause of sadness), which activates a goal of reward acquisition that shifts preferences toward high-reward options. In contrast, anxious individuals tend to infer that the situation is uncertain and beyond control (typical causes of anxiety), which activates a goal of risk avoidance that shifts preferences toward low-risk options. Similarly, Lerner, Small, and Loewenstein (2004) found that incidental states of sadness amplify the endowment effect—the tendency to overvalue one's possessions—whereas incidental states of disgust attenuate this effect. This is presumably because sadness, even

if incidental, triggers an impulse to hold on to sources of rewards, whereas disgust triggers an impulse to expulse sources of discomfort. Raghunathan, Pham, and Corfman (2006) recently showed that the effects of specific incidental emotions are more pronounced if there is a surface similarity between the true source of the incidental emotion and the target decision—a phenomenon they call displaced coping. Finally, Mukhopadhyay and Johar (in press) found that feelings of pride or happiness engendered by recent shopping history (i.e., giving in vs. restraining from temptations) carry over to influence the effectiveness of advertising appeals viewed subsequently so that pride (happiness) appeals are more effective after restraint (giving in).

Task-Related Affect

Finally, the process of making a decision may itself induce affect. For example, difficult choices are often those that involve trade-offs on important attributes (Bettman, Johnson, Luce, & Payne, 1993), such as deciding how much quality to surrender in order to save money. One means of minimizing such negative affects is to maintain the status quo and not make an active choice (Luce, 1998; Nowlis, Kahn, & Dhar, 2002). Although one might expect consumers to work harder when faced with difficult choices, they instead seem to shift to simplifying choice heuristics that minimize task-related negative affect (Luce, Bettman, & Payne, 1997). Similar effects have been observed for another aspect of decision tasks that can generate negative affect: time pressure (Dhar & Nowlis, 1999; Payne, Bettman, & Johnson, 1988). It has also been found that the more consumers deliberate about their choices, the more they become emotionally attached to the options, which leads to decision-related discomfort once one option has been chosen (Carmon, Wertenbroch, & Zeelenberg, 2003).

A particularly important outcome of a task-related affect is the transfer of that affect to the valuation of the alternative chosen. Garbarino and Edell (1997) demonstrate that reducing the effort involved in selecting an alternative can increase the price respondents are willing to pay for that option. Similarly, Higgins and colleagues have shown that a fit between the manner in which a decision is made and the current orientation of a decision maker can produce positive task-induced feelings that increase the perceived value for a chosen object (Avnet & Higgins, 2003; Higgins, 2000; Higgins, Idson, Freitas, Spiegel, & Molden, 2003).

IMPLICIT PROCESSES

For decades, researchers have, almost by force of habit, asked participants whether they were aware of a given manipulation in an experiment. The lack of awareness of the manipulation was prima facie evidence that demand characteristics did not operate. More recently, questions surrounding consumers' awareness of the causes of their behavior have generated significant controversy and interest in consumer research.

Before proceeding, we need to clarify the meaning of "implicit processes," a term that is borrowed, at least in part, from memory research (Fazio & Olson, 2003). As our opening example suggests, one important characteristic attributed to implicit processes is a lack of awareness (Bargh, 2002), but the exact definition of awareness can be more complicated (Chartrand, 2005). Does awareness mean that a respondent is aware of the possibility that some factor may influence someone's behavior in some task, or does it require a respondent to know the exact influence in that task? A second characteristic associated with implicit processes is automaticity: requiring less effort or intentionality (Bargh & Chartrand, 1999). However, while automaticity is well defined in the cognitive literature (Shiffrin & Schneider, 1977), most studies in consumer behavior tend to assess this property more casually.

Priming

Priming is one of the earliest areas of consumer research to exploit the idea of implicit processes—an area that has received considerable attention from social psychologists (for a review, see Higgins, 1996). A particularly innovative study by Nedungadi (1990) manipulated the accessibility of brand names in memory and established an increased likelihood of choice, even when these effects were outside the awareness of the respondent. Herr (1989) showed that priming a category can change choices, while Yi (1990b) demonstrated that the priming of a product feature can increase the weight given to that feature.

Priming cultural knowledge has also been shown to affect choices. Mandel (2003) primed either independent or dependent identities by having Americans read a very simple description of others' motives changed by both hypothetical and actual purchases. Chen, Ng, and Rao (2005) demonstrated that priming one of two cultural identities in bicultural Singaporeans had a significant affect on their preference for accelerated consumption, which the authors attribute to cultural difference in regulatory focus. Johar, Moreau, and Schwarz (2003) used advertisements to prime cultural stereotypes of women as homemakers and found that these ads lead to stereotypical judgments of women encountered in a subsequent "study." Priming either cultural identity (Forehand & Deshpande, 2001) or particular goals (Strahan, Spencer, & Zanna, 2002) can also have an effect on the processing of subsequent persuasive communication. Even the presence of familiar objects can serve as a prime. Shrum, Wyer, and O'Guinn (1998) demonstrated that the mere presence of a television can prime beliefs that are consistent with most television content, such as increased perceptions of violence and crime.

Two studies demonstrated how priming can affect consumer behavior in actual consumption settings (North, Hargreaves, & McKendrick, 1997, 1999). The authors manipulated the background music in a store selling both German and French wine, predicting that pleasant music from the country of origin would increase sales of that country's products. Not only did this occur, but the background music accounted for about a quarter of the variance in choice between the two products. Similarly, Mandel and Johnson (2002) manipulated the background of a website, or wallpaper, using embedded designs shown to prime either quality attributes (safety for cars, comfort for couches) or price. They found marked differences in choice between the two products, cars and couches consistent with the prime, and that the prime affected both novices and experts in the product class, although through different mechanisms. In both studies, respondents specifically denied being influenced by the manipulation.

Mere Measurement

One application from social psychology that serves as an interesting case is the mere measurement effect. Using ideas from the social literature on the self-fulfilling nature of prediction errors (Greenwald, 1987; Sherman, 1980), Morwitz, Johnson, and Schmittlein (1993) hypothesized that merely asking questions about future behaviors could, in fact, change those behaviors. They examined data from two large commercial market research surveys and found that the mere asking of an intent question, such as "Do you intend to buy a personal computer in the next six months?" actually influenced purchases. Both those who answered yes and no behaved in ways that were more consistent with the request than groups that had not been asked. Subsequently, such effects have been shown to increase repurchases of one's current brand of car (Fitzsimons & Morwitz, 1996), and to accelerate purchases of supermarket goods for up to 3 months (Chandon, Morwitz, & Reinartz, 2004). In financial services, Dholika and Morwitz (2002) demonstrated the effect lasts up to a year when the question is asked as part of a consumer satisfaction survey. A similar effect has been shown in blood donation (Godin, Conner, Sheeran, & Germain, 2005) and can be shown to occur with mass communication requests (Spangenberg, Sprott, Grohmann, & Smith, 2003). Current explanations for these kinds of effects indicate that a general-intent question makes attitudes about specific items in the category more accessible (Morwitz & Fitzsimons, 2004), although some argue for a role of dissonance reduction (Spangenberg et al., 2003). However, because these effects last many months and occur without consumer awareness, they are unlikely to result from conscious processing of the question and its effect.

What makes this interesting as an implicit process is whether or not people are aware of the effect. Williams, Fitzsimons, and Block (2004) show that awareness of the persuasive nature of these effects can limit their effectiveness, but in experimental studies, it appears that such awareness is rare (Fitzsimons & Shiv, 2001). Thus, a particularly subtle interaction, occurring frequently as part of a market survey research encounter, seems to be able to have long-term behavioral consequences, which for the most part are unforeseen by both consumer and marketer.

Theoretical and Empirical Issues

One element of work in implicit processes has only started to impact consumer research—the idea of automatic goal activation. From a theoretical perspective, a study by Brendl, Markman, and Messner (2003) demonstrates an important point. In that study, the authors show that activating one need (e.g., the need to smoke) serves to lower the value of other objects, even when those objects are fungible, such as cash, and can be used to fulfill that need. The exact nature of goals, procedures for measuring their strength, and the relationships between goals will be an important item on the agenda for research into implicit processes (see Janiszewski & van Osselaer, 2005).

As a practical issue, in an applied area such as consumer research, much work in measurement and validation remains. Examination of the applicability of techniques such as the Implicit Attitude Test (Brunel, Tietje, & Greenwald, 2004; Maison, Greenwald, & Bruin, 2004) have started, but much more work needs to be done (Fazio & Olson, 2003).

An active, lively debate concerning the relative role of conscious and unconscious processes in consumer choice continues (see Dijksterhuis, Smith, van Baaren, & Wigboldus, 2005; Simonson, 2005, and other articles in this issue for examples).

Awareness and Subliminal Effects Reloaded

In 1980, a review of subliminal effects published in the *Journal of Marketing* (Moore, 1982) concluded, "The idea that subliminal directives can affect motives or actions is contradicted by much research and is incompatible with theoretical conceptions of perception and motivation. . . ." (p. 46). Has the research we have reviewed in this section done much to change this conclusion? Clearly, the focus has changed. The question no longer is whether stimuli are perceived or not but, rather, whether the effects of such stimuli, whether subliminal or supraliminal, are known to consumers. Awareness of the manipulations and a correct mental model of their effects seem necessary for any attempt by consumers to counteract their effects. As an applied area of research, questions do arise about the robustness and size of these effects, which, if answered, have implications for public policy and the ethics of marketing practice. Most important, a strong theoretical framework consistent with "conceptions of perception and motivation" would be useful.

CONCLUSION

Our goal in writing this chapter was to convince the reader that work in consumer behavior and marketing should be of great interest to social psychologists. In closing, it is then potentially helpful to outline some of the reasons social psychologists should be interested. The movement toward social cognition was motivated, in part, by the feeling that the same cognitive system that

categorized, learned about, and navigated the physical world also operated in the social world. Similarly, the same systems are involved in the world of transactions, goods, and services. Just as social cognition research hoped to be more than just a source of interesting examples and applications of cognitive concepts, consumer behavior and marketing should be more than a source of interesting examples of social psychological concepts. This realm should be, more fundamentally, a source of new ideas, concepts, and data. Much of what we have reviewed is research that, so far, has not been central to discussions in social psychology, but that, to us at least, seems to be quite relevant. It is our hope that this chapter will serve to change the picture in Figure 38.1 to one in which ideas flow more freely in both directions.

NOTE

1. While these data are from 2004, the picture would be similar for other recent years. Similarly, the inclusion of a more complete list of consumer research journals particularly the *Journal of Consumer Psychology*, or a more complete list of psychology journals, would only serve to make the same point.

REFERENCES

Aaker, D. A., & Keller, K. L. (1990). Consumer evaluations of brand extensions. *Journal of Marketing*, *54*(1), 27–41.

Aaker, D. A., Stayman, D. M., & Hagerty, M. R. (1986). Warmth in advertising: Measurement, impact, and sequence effects. *Journal of Consumer Research*, *12*(4), 365–381.

Aaker, J. L. (1997). Dimensions of brand personality. *Journal of Marketing Research*, *34*(3), 347–356.

Aaker, J. L., Benet-Martinez, V., & Garolera, J. (2001). Consumption symbols as carriers of culture: A study of Japanese and Spanish brand personality constructs. *Journal of Personality and Social Psychology*, *81*(3), 492–508.

Aaker, J., Fournier, S., & Brasel, S. A. (2004). When good brands do bad. *Journal of Consumer Research*, *31*(1), 1–16.

Aaker, J. L., & Maheswaran, D. (1997). The effect of cultural orientation on persuasion. *Journal of Consumer Research*, *24*(3), 315–328.

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology*, *43*, 619–630.

Adaval, R. (2001). Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information. *Journal of Consumer Research*, *28*(1), 1–17.

Agrawal, N., & Maheswaran, D. (2005). The effects of self-construal and commitment on persuasion. *Journal of Consumer Research*, *31*(4), 841–849.

Ahluwalia, R. (2000). Examination of psychological processes underlying resistance to persuasion. *Journal of Consumer Research*, *27*(2), 217–232.

Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research*, *29*, 270–279.

Ahluwalia, R., & Burnkrant, R. E. (2004). Answering questions

about questions: A persuasion knowledge perspective for understanding the effects of rhetorical questions. *Journal of Consumer Research*, *31*(1), 26–42.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13*(4), 411–454.

Allen, C. T., Machleit, K. A., & Kleine, S. S. (1992). A comparison of attitudes and emotions as predictors of behavior at diverse levels of behavioral experience. *Journal of Consumer Research*, *18*(4), 493–504.

Avnet, T., & Higgins, E. T. (2003). Locomotion, assessment, and regulatory fit: Value transfer from "how" to "what." *Journal of Experimental Social Psychology*, *39*(5), 525–530.

Axelrod, J. N. (1963). Induced moods and attitudes towards products. *Journal of Advertising Research*, *3*, 19–24.

Bargh, J. A. (2002). Losing consciousness: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, *29*(2), 280–285.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*(7), 462–479.

Bell, D. R., & Lattin, J. M. (2000). Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science*, *19*(2), 185–200.

Bettman, J. R. (1970). Information processing models of consumer behavior. *Journal of Marketing Research*, 7, 370–376.

Bettman, J. R., Johnson, E. J., Luce, M. F., & Payne, J. W. (1993). Correlation, conflict, and choice. *Journal of Experimental Psychology: Learning Memory and Cognition*, *19*(4), 931–951.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research*, *3*(4), 233–240.

Biehal, G., & Chakravarti, D. (1982). Information-presentation format and learning-goals as determinants of consumers' memory retrieval and choice processes. *Journal of Consumer Research*, *8*(4), 431–441.

Biehal, G., & Chakravarti, D. (1986). Consumers' use of memory and external information in choice: Macro and micro perspectives. *Journal of Consumer Research*, *12*(4), 382–405.

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping* (pp. 13–37). San Diego, CA: Academic Press.

Bodur, H. O., Brinberg, D., & Coupey, E. (2000). Belief, affect, and attitude: Alternative models of the determinants of attitude. *Journal of Consumer Psychology*, *9*(1), 17–28.

Boush, D. M., & Loken, B. (1991). A process-tracing study of brand extension evaluation. *Journal of Marketing Research*, *28*(1), 16–28.

Breckler, S. J., & Wiggins, E. C. (1989). Affect versus evaluation in the structure of attitudes. *Journal of Experimental Social Psychology*, *25*(3), 253–271.

Brendl, C. M., Markman, A. B., & Messner, C. (2003). The devaluation effect: Activating a need devalues unrelated objects. *Journal of Consumer Research*, *29*(4), 463–473.

Brown, S. P., Homer, P. M., & Inman, J. J. (1998). A meta-analysis of relationships between ad-evoked feelings and advertising responses. *Journal of Marketing Research*, *35*(1), 114–126.

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the implicit association test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology*, *14*(4), 385–404.

Burke, M. C., & Edell, J. A. (1989). The impact of feelings on ad-based affect and cognition. *Journal of Marketing Research*, *26*(1), 69–83.

Burke, R. R., & Srull, T. K. (1988). Competitive interference and consumer memory for advertising. *Journal of Consumer Research*, *15*(1), 55–68.

Burnkrant, R. E., & Unnava, H. R. (1995). Effects of self-referencing on persuasion. *Journal of Consumer Research*, *22*(1), 17–26.

Carmon, Z., Wertenbroch, K., & Zeelenberg, M. (2003). Option at-

tachment: When deliberating makes choosing feel like losing. *Journal of Consumer Research*, *30*(1), 15–29.

Carpenter, G. S., Glazer, R., & Nakamoto, K. (1994). Meaningful brands from meaningless differentiation—The dependence on irrelevant attributes. *Journal of Marketing Research*, *31*(3), 339–350.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right." *Journal of Personality and Social Psychology*, *86*, 388–404.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing—Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, *66*(3), 460–473.

Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2004). The short- and long-term effects of measuring intent to repurchase. *Journal of Consumer Research*, *31*(3), 566–572.

Chartrand, T. L. (2005). The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology*, *15*(3), 203–210.

Chen, H. P., Ng, S., & Rao, A. R. (2005). Cultural differences in consumer impatience. *Journal of Marketing Research*, *42*(3), 291–301.

Chernev, A. (2004). Goal orientation and consumer preference for the status quo. *Journal of Consumer Research*, *31*(3), 557–565.

Cohen, J. B., Pham, M. T., & Andrade, E. B. (in press). The role and nature of affect in consumer judgment and decision making. In C. P. Haugtvedt, P. M. Herr, & F. R. Kardes (Eds.), *Handbook of consumer psychology*. Mahwah, NJ: Erlbaum.

Costley, C. L., & Brucks, M. (1992). Selective recall and information use in consumer preferences. *Journal of Consumer Research*, *18*(4), 464–474.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869.

Deaton, A., & Muellbauer, J. (1980). *Economics and consumer behavior*. New York: Cambridge University Press.

Derbaix, C. M. (1995). The impact of affective reactions on attitudes toward the advertisement and the brand—A step toward ecological validity. *Journal of Marketing Research*, *32*(4), 470–479.

Derbaix, C., & Pham, M. T. (1991). Affective reactions to consumption situations—A pilot investigation. *Journal of Economic Psychology*, *12*(2), 325–355.

Dhar, R., & Nowlis, S. M. (1999). The effect of time pressure on consumer choice deferral. *Journal of Consumer Research*, *25*(4), 369–384.

Dholakia, U. M., & Morwitz, V. G. (2002). The scope and persistence of mere-measurement effects: Evidence from a field study of customer satisfaction measurement. *Journal of Consumer Research*, *29*(2), 159–167.

Dholakia, U. M., & Simonson, I. (2005). The effect of explicit reference points on consumer choice and online bidding behavior. *Marketing Science*, *24*(2), 206–217.

Dick, A., Chakravarti, D., & Biehal, G. (1990). Memory-based inferences during consumer choice. *Journal of Consumer Research*, *17*(1), 82–93.

Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The unconscious consumer: Effects of environment on consumer behavior. *Journal of Consumer Psychology*, *15*(3), 193–202.

Dommermuth, W. P., & Millard, W. J. (1967). Consumption coincidence in product evaluation. *Journal of Marketing Research*, *4*(4), 388–390.

Eagly, A. H., & Chaiken, S. (1983). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.

Edell, J. A., & Burke, M. C. (1987). The power of feelings in understanding advertising effects. *Journal of Consumer Research*, *14*(3), 421–433.

Erdem, T., Mayhew, G., & Sun, B. H. (2001). Understanding reference-price shoppers: A within- and cross-category analysis. *Journal of Marketing Research*, *38*(4), 445–457.

Fazio, R. H. (1990). Multiple processes by which attitudes guide

behavior—The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75–109.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Fedorikhin, A., & Cole, C. A. (2004). Mood effects on attitudes, perceived risk and choice: Moderators and mediators. *Journal of Consumer Psychology*, *14*(1–2), 2–12.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1–17.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Fitzsimons, G. J., & Morwitz, V. G. (1996). The effect of measuring intent on brand-level purchase behavior. *Journal of Consumer Research*, *23*(1), 1–11.

Fitzsimons, G. J., & Shiv, B. (2001). Nonconscious and contaminative effects of hypothetical questions on subsequent decision making. *Journal of Consumer Research*, *28*(2), 224–238.

Forehand, M. R., & Deshpande, R. (2001). What we see makes us who we are: Priming ethnic self-awareness and advertising response. *Journal of Marketing Research*, *38*(3), 336–348.

Fournier, S. (1998). Consumers and their brands: Developing relationship theory in consumer research. *Journal of Consumer Research*, *24*(4), 343–373.

Friestad, M., & Wright, P. (1995). Persuasion knowledge: Lay people's and researchers' beliefs about the psychology of advertising. *Journal of Consumer Research*, *22*(1), 62–74.

Garbarino, E. C., & Edell, J. A. (1997). Cognitive effort, affect, and choice. *Journal of Consumer Research*, *24*(2), 147–158.

Godin, G., Conner, M., Sheeran, P., & Germain, M. (2005). Asking questions changes behaviour: The mere-measurement effect in the field of blood donation. *Psychology and Health*, *20*, 96–97.

Goldberg, M. E., & Gorn, G. J. (1987). Happy and sad TV programs: How they affect reactions to commercials. *Journal of Consumer Research*, *14*(3), 387–403.

Gorn, G. J. (1982). The effects of music in advertising on choice behavior: A classical conditioning approach. *Journal of Marketing*, *46*(1), 94–101.

Gorn, G. J., Chattopadhyay, A., Sengupta, J., & Tripathi, S. (2004). Waiting for the Web: How screen color affects time perception. *Journal of Marketing Research*, *41*(2), 215–225.

Gorn, G. J., Goldberg, M. E., & Basu, K. (1993). Mood, awareness and product evaluation. *Journal of Consumer Psychology*, *2*(3), 237–256.

Gorn, G. J., Pham, M. T., & Sin, L. Y. (2001). When arousal influences ad evaluation and valence does not (and vice versa). *Journal of Consumer Psychology*, *11*(1), 43–55.

Greenwald, A. G. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology*, *72*, 315–318.

Grether, D., & Wilde, L. (1984). An analysis of conjunctive choice: Theory and experiments. *Journal of Consumer Research*, *10*(4), 373–385.

Gurhan-Canli, Z., & Maheswaran, D. (2000a). Cultural variations in country of origin effects. *Journal of Marketing Research*, *37*(3), 309–317.

Gurhan-Canli, Z., & Maheswaran, D. (2000b). Determinants of country-of-origin evaluations. *Journal of Consumer Research*, *27*(1), 96–108.

Hamilton, R. W. (2003). Why do people suggest what they do not want? Using context effects to influence others' choices. *Journal of Consumer Research*, *29*(4), 492–506.

Hardie, B. G. S., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science*, *12*(4), 378–394.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or online. *Psychological Review*, *93*(3), 258–268.

Havlena, W. J., & Holbrook, M. B. (1986). The varieties of consumption experience: Comparing two typologies of emotion in consumer behavior. *Journal of Consumer Research*, *13*(3), 394–404.

Heath, C., & Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of Consumer Research*, *23*(1), 40–52.

Heath, T. B., Ryu, G., Chatterjee, S., McCarthy, M. S., Mothersbaugh, D. L., Milberg, S., et al. (2000). Asymmetric competition in choice and the leveraging of competitive disadvantages. *Journal of Consumer Research*, *27*(3), 291–308.

Henderson, P. W., & Peterson, R. A. (1992). Mental accounting and categorization. *Organizational Behavior and Human Decision Processes*, *51*(1), 92–117.

Herr, P. M. (1989). Priming price—Prior knowledge and context effects. *Journal of Consumer Research*, *16*(1), 67–75.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168).

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*(12), 1280–1300.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, *55*(11), 1217–1230.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, *84*(6), 1140–1153.

Higgins, E. T., Kuiper, N. A., & Olson, J. M. (1981). Social cognition: A need to get personal. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1, pp. 395–420). Hillsdale, NJ: Erlbaum.

Holbrook, M. B., & Batra, R. (1987). Assessing the role of emotions as mediators of consumer responses to advertising. *Journal of Consumer Research*, *14*(3), 404–420.

Hornik, J. (1992). Tactile stimulation and consumer response. *Journal of Consumer Research*, *19*(3), 449–458.

Howard, D. J. (1992). Gift-wrapping effects on product attitudes: A mood-biasing explanation. *Journal of Consumer Psychology*, *1*(3), 197–223.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, *67*(3), 247–257.

Hsee, C. K., & Leclerc, F. (1998). Will products look more attractive when presented separately or together? *Journal of Consumer Research*, *25*(2), 175–186.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General*, *133*(1), 23–30.

Hsee, C. K., & Zhang, J. (2004). Distinction bias: Misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology*, *86*(5), 680–695.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives—Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98.

Isen, A. M., Shalker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior—Cognitive loop. *Journal of Personality and Social Psychology*, *36*(1), 1–12.

Jacoby, J. (1975). Perspectives on a consumer information-processing research—program. *Communication Research*, *2*(3), 203–215.

Jacoby, J., Jaccard, J. J., Currim, I., Kuss, A., Ansari, A., & Troutman, T. (1994). Tracing the impact of item-by-item information accessing on uncertainty reduction. *Journal of Consumer Research*, *21*(2), 291–303.

Jain, S. P., & Maheswaran, D. (2000). Motivate reasoning: A depth-of-processing perspective. *Journal of Consumer Research*, *26*, 358–371.

Janiszewski, C., & van Osselaer, S. M. J. (2005). Behavior activation is not enough. *Journal of Consumer Psychology*, *15*(3), 218–224.

Jewell, R. D., & Unnava, H. R. (2003). When competitive interfer-

ence can be beneficial. *Journal of Consumer Research*, *30*(2), 283–291.

Johar, G. V. (1995). Consumer involvement and deception from implied advertising claims. *Journal of Marketing Research*, *32*(3), 267–279.

Johar, G. V., Moreau, C. P., & Schwarz, N. (2003). Gender typed advertisements and impression formation: The role of chronic and temporary accessibility. *Journal of Consumer Psychology*, *13*, 220–229.

Johar, G. V., Sengupta, J., & Aaker, J. L. (2005). Two roads to updating brand personality impressions: Trait versus evaluative inferencing. *Journal of Marketing Research*, *42*(4), 458–469.

Johar, G. V., & Simmons, C. J. (2000). The use of concurrent disclosures to correct invalid inferences. *Journal of Consumer Research*, *26*(4), 307–322.

John, D. R., Loken, B., & Joiner, C. (1998). The negative impact of extensions: Can flagship products be diluted? *Journal of Marketing*, *62*(1), 19–32.

John, D. R., & Sujan, M. (1990). Age-differences in product categorization. *Journal of Consumer Research*, *16*(4), 452–460.

Johnson, E. J., & Russo, J. E. (1978). The organization of product information in memory identified by recall times. In H. K. Hunt (Ed.), *Advances in consumer research* (Vol. 6, pp. 79–86). Ann Arbor, MI: Association for Consumer Research.

Johnson, E. J., & Russo, J. E. (1984). Product familiarity and learning new information. *Journal of Consumer Research*, *11*(1), 542–550.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, *45*(1), 20–31.

Kahneman, D., Schkade, D., & Sunstein, C. R. (1998). Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk and Uncertainty*, *16*(1), 49–86.

Kalyanaram, G., & Little, J. D. C. (1994). An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research*, *21*(3), 408–418.

Kardes, F. R., Posavac, S. S., & Cronley, M. L. (2004). Consumer inference: A review of processes, bases, and judgment contexts. *Journal of Consumer Psychology*, *14*(3), 230–256.

Keller, K. L. (1987). Memory factors in advertising: The effect of advertising retrieval cues on brand evaluations. *Journal of Consumer Research*, *14*(3), 316–333.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, *57*(1), 1–22.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004a). Alternative models for capturing the compromise effect. *Journal of Marketing Research*, *41*(3), 237–257.

Kivetz, R., Netzer, O., & Srinivasan, V. (2004b). Extending compromise effect models to complex buying situations and other context effects. *Journal of Marketing Research*, *41*(3), 262–268.

Klapper, D., Ebling, C., & Temme, J. (2005). Another look at loss aversion in brand choice data: Can we characterize the loss averse consumer? International *Journal of Research in Marketing*, *22*(3), 239–254.

Leclerc, F., Schmitt, B. H., & Dube, L. (1995). Waiting time and decision making: Is time like money? *Journal of Consumer Research*, *22*(1), 110–119.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon & Schuster.

Leith, K. P., & Baumeister, R. F. (1996). Why do bad moods increase self-defeating behavior? Emotion, risk taking, and self-regulation. *Journal of Personality and Social Psychology*, *71*(6), 1250–1267.

Lerner, J., Small, D., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic transactions. *Psychological Science*, *15*(5), 337–341.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*(1), 46–55.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272–292.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, E. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Loken, B. (2006). Consumer psychology: Categorization, inferences, affect, and persuasion. *Annual Review of Psychology*, *57*, 453–485.

Loken, B., & John, D. R. (1993). Diluting brand beliefs—When do brand extensions have a negative impact. *Journal of Marketing*, *57*(3), 71–84.

Luce, M. F. (1998). Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research*, *24*(4), 409–433.

Luce, M. F., Bettman, J. R., & Payne, J. W. (1997). Choice processing in emotionally difficult decisions. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(2), 384–405.

Luna, D., & Peracchio, L. A. (2001). Moderators of language effects in advertising to bilinguals: A psycholinguistic approach. *Journal of Consumer Research*, *28*(2), 284–295.

MacGregor, D. G., Slovic, P., Dreman, D., & Berry, M. (2000). Imagery, affect, and financial judgment. *Journal of Psychology and Financial Markets*, *1*(2), 104–110.

Mackenzie, S. B. (1986). The role of attention in mediating the effect of advertising on attribute importance. *Journal of Consumer Research*, *13*(2), 174–195.

Maheswaran, D., & Sternthal, B. (1990). The effects of knowledge, motivation, and type of message on ad processing and product judgments. *Journal of Consumer Research*, *17*(1), 66–73.

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the implicit association test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology*, *14*(4), 405–415.

Mandel, N. (2003). Shifting selves and decision making: The effects of self-construal priming on consumer risk-taking. *Journal of Consumer Research*, *30*(1), 30–40.

Mandel, N., & Johnson, E. J. (2002). When Web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, *29*(2), 235–245.

Mano, H. (1992). Judgments under distress: Assessing the role of unpleasantness and arousal in judgment formation. *Organizational Behavior and Human Decision Processes*, *52*(2), 216–245.

Mano, H. (1994). Risk-taking, framing effects, and affect. *Organizational Behavior and Human Decision Processes*, *57*(1), 38–58.

Mathur, M., & Chattopadhyay, A. (1991). The impact of moods generated by television programs on responses to advertising. *Psychology and Marketing*, *8*(1), 59–77.

Mayhew, G. E., & Winer, R. S. (1992). An empirical analysis of internal and external reference prices using scanner data. *Journal of Consumer Research*, *19*(1), 62–70.

Mazumdar, T., Raj, S. P., & Sinha, I. (2005). Reference price research: Review and propositions. *Journal of Marketing*, *69*(4), 84–102.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503–507.

McGraw, A. P., Tetlock, P. E., & Kristel, O. V. (2003). The limits of fungibility: Relational schemata and the value of things. *Journal of Consumer Research*, *30*(2), 219–229.

Meyers-Levy, J. (1988). The influence of sex roles on judgment. *Journal of Consumer Research*, *14*(4), 522–530.

Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research*, *18*(1), 63–70.

Meyers-Levy, J., & Maheswaran, D. (2004). Exploring message framing outcomes when systematic, heuristic, or both types of processing occur. *Journal of Consumer Psychology*, *14*(1–2), 159–167.

Meyvis, T., & Janiszewski, C. (2004). When are broader brands stronger brands? An accessibility perspective on the success of brand extensions. *Journal of Consumer Research*, *31*(2), 346–357.

Milliman, R. E. (1982). Using background music to affect the behavior of supermarket shoppers. *Journal of Marketing*, *46*(3), 86–91.

Milliman, R. E. (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research*, *13*(2), 286–289.

Miniard, P. W., Bhatla, S., & Sirdeshmukh, D. (1992). Mood as a determinant of postconsumption product evaluations. *Journal of Consumer Psychology*, *1*(2), 173–195.

Moore, T. (1982). Subliminal advertising: What you see is what you get. *Journal of Marketing*, *46*(2), 38–47.

Morales, A. C., & Fitzsimons, G. J. (in press). Product contagion: Changing consumer evaluations through physical contact with "disgusting" products. *Journal of Marketing Research*.

Morwitz, V. G., & Fitzsimons, G. J. (2004). The mere-measurement effect: Why does measuring intentions change actual behavior? *Journal of Consumer Psychology*, *14*(1–2), 64–74.

Morwitz, V. G., Johnson, E., & Schmittlein, D. (1993). Does measuring intent change behavior. *Journal of Consumer Research*, *20*(1), 46–61.

Mukhopadhyay, A., & Johar, G. V. (in press). Tempted or not: The effect of recent purchase history on responses to affective advertising. *Journal of Consumer Research*.

Murry, J. P., & Dacin, P. A. (1996). Cognitive moderators of negative-emotion effects: Implications for understanding media context. *Journal of Consumer Research*, *22*(4), 439–447.

Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of Consumer Research*, *17*(3), 263–276.

North, A. C., Hargreaves, D. J., & McKendrick, J. (1997). In-store music affects product choice. *Nature*, *390*, 132.

North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). Research reports—The influence of in-store music on wine selections. *Journal of Applied Psychology*, *84*(2), 271.

Nowlis, S. M., Kahn, B. E., & Dhar, R. (2002). Coping with ambivalence: The effect of removing a neutral option on consumer attitude and preference judgments. *Journal of Consumer Research*, *29*(3), 319–334.

Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utifitarian goods. *Journal of Marketing Research*, *42*(1), 43–53.

Oliver, R. L. (1993). Cognitive, affective, and attribute bases of the satisfaction response. *Journal of Consumer Research*, *20*(3), 418–430.

Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions—The role of product feature similarity and brand concept consistency. *Journal of Consumer Research*, *18*(2), 185–193.

Park, J. W., & Hastak, M. (1994). Memory-based product Judgments—Effects of involvement at encoding and retrieval. *Journal of Consumer Research*, *21*(3), 534–547.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 534–552.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1991). *The adaptive decision-maker*. Cambridge, UK: Cambridge University Press.

Peters, E., Slovic, P., & Gregory, R. (2003). The role of affect in the WTA/WTP disparity. *Journal of Behavioral Decision Making*, *16*(4), 309–330.

Petrova, P. K., & Cialdini, R. B. (2005). Fluency of consumption imagery and the backfire effects of imagery appeals. *Journal of Consumer Research*, *32*, 442–452.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135–146.

Pham, M. T. (1996). Cue representation and selection effects of arousal on persuasion. *Journal of Consumer Research*, *22*(4), 373–387.

Pham, M. T. (1998). Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, *25*(2), 144–159.

Pham, M. T. (2004). The logic of feeling. *Journal of Consumer Psychology*, *14*(4).

Pham, M. T., & Avnet, T. (2004). Ideals and oughts and the reliance on affect versus substance in persuasion. *Journal of Consumer Research*, *30*(4), 503–518.

Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research*, *28*(2), 167–188.

Posavac, S. S., Sanbonmatsu, D. M., Kardes, F. R., & Fitzsimons, G. J. (2004). The brand positivity effect: When evaluation confers preference. *Journal of Consumer Research*, *31*(3), 643–651.

Prelec, D., & Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science*, *17*(1), 4–28.

Prelec, D., & Simester, D. (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. *Marketing Letters*, *12*(1), 5–12.

Priester, J. R., Dholakia, U. M., & Fleming, M. A. (2004). When and why the background contrast effect emerges: Thought engenders meaning by influencing the perception of applicability. *Journal of Consumer Research*, *31*(3), 491–501.

Priester, J. R., Godek, J., Nayakankuppum, D. J., & Park, K. (2004). Brand congruity and comparative advertising: When and why comparative advertisements lead to greater elaboration. *Journal of Consumer Psychology*, *14*(1–2), 115–123.

Priester, J. R., Nayakankuppam, D., Fleming, M. A., & Godek, J. (2004). The A(2)SC(2) model: The influence of attitudes and attitude strength on consideration and choice. *Journal of Consumer Research*, *30*(4), 574–587.

Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. *Marketing Science*, *11*(3), 287–309.

Raghunathan, R., & Irwin, J. R. (2001). Walking the hedonic product treadmill: Default contrast and mood-based assimilation in judgments of predicted happiness with a target product. *Journal of Consumer Research*, *28*(3), 355–368.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, *79*(1), 56–77.

Raghunathan, R., Pham, M. T., & Corfman, K. P. (2006). Informational properties of anxiety and sadness, and displaced coping. *Journal of Consumer Research*, *32*(4).

Ratneshwar, S., & Chaiken, S. (1991). Comprehension's role in persuasion: The case of its moderating effect on the persuasive impact of source cues. *Journal of Consumer Research*, *18*(1), 52–62.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, *19*(1–3), 171–197.

Resnik, A., & Stern, B. L. (1977). Analysis of information-content in television advertising. *Journal of Marketing*, *41*(1), 50–53.

Richins, M. L. (1997). Measuring emotions in the consumption experience. *Journal of Consumer Research*, *24*(2), 127–146.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, *37*, 322–336.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, *12*(3), 185–190.

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of Consumer Research*, *21*(2), 274–290.

Russo, J. E., Meloy, M. G., & Medvec, V. H. (1998). Predecisional distortion of product information. *Journal of Marketing Research*, *35*(4), 438–452.

Russo, J. E., & Rosen, L. D. (1974). An eye fixation analysis of multialternative choice. *Memory and Cognition*, *3*, 267–276.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision-making. *Journal of Personality and Social Psychology*, *59*(4), 614–622.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it?: The information function of affective states. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior: New evidence and integrative attempts* (pp. 44–63). Toronto: Hogrefe.

Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433–465). New York: Guilford Press.

Sen, S., & Johnson, E. J. (1997). Mere-possession effects without possession in consumer choice. *Journal of Consumer Research*, *24*(1), 105–117.

Sengupta, J., & Fitzsimons, G. (2000). Disruption vs. reinforcement: The effects of analyzing reasons for brand preferences. *Journal of Marketing Research*, *37*, 318–330.

Sengupta, J., & Fitzsimons, G. J. (2004). The effect of analyzing reasons on the stability of brand attitudes: A reconciliation of opposing predictions. *Journal of Consumer Research*, *31*, 705–711.

Sengupta, J., Goodstein, R. C., & Boninger, D. S. (1997). All cues are not created equal: Obtaining attitude persistence under low-involvement conditions. *Journal of Consumer Research*, *23*, 351–361.

Sengupta, J., & Johar, G. V. (2001). Contingent effects of anxiety on message elaboration and persuasion. *Personality and Social Psychology Bulletin*, *27*, 139–150.

Sengupta, J., & Johar, G. V. (2002). Effects of inconsistent attribute information on the predictive value of product attitudes: Toward a resolution of opposing perspectives. *Journal of Consumer Research*, *29*, 39–56.

Shafir, S., Bechar, A., & Weber, E. U. (2003). Cognition-mediated coevolution—Context-dependent evaluations and sensitivity of pollinators to variability in nectar rewards. *Plant Systematics and Evolution*, *238*(1–4), 195–209.

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (Apis mellifera) and gray jays (Perisoreus canadensis). *Behavioral Ecology and Sociobiology*, *51*(2), 180–187.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*(2), 211–221.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, *26*(3), 278–292.

Shrum, L. J., Wyer, R. S., Jr., & O'Guinn, T. C. (1998). The effects of television consumption on social perceptions: The use of priming procedures to investigate psychological processes. *Journal of Consumer Research*, *24*(4), 447–458.

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, *15*(5), 331–336.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *59*, 99–118.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174.

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research*, *19*(1), 105–118.

Simonson, I. (2005). In defense of consciousness: The role of conscious and unconscious inputs in consumer choice. *Journal of Consumer Psychology*, *15*(3), 211–217.

Simonson, I., & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes*, *51*(3), 416–446.

Soman, D. (2001). The mental accounting of sunk time costs: Why time is not like money. *Journal of Behavioral Decision Making*, *14*(3), 169–185.

Soman, D. (2003). The effect of payment transparency on consumption: Quasi-experiments from the field. *Marketing Letters*, *14*(3), 173–183.

Soman, D., & Cheema, A. (2002). The effect of credit on spending decisions: The role of the credit limit and credibility. *Marketing Science*, *21*(1), 32–53.

Soman, D., & Gourville, J. T. (2001). Transaction decoupling: How price bundling affects the decision to consume. *Journal of Marketing Research*, *38*(1), 30–44.

Sood, S., Rottenstreich, Y., & Brenner, L. (2004). On decisions that lead to decisions: Direct and derived evaluations of preference. *Journal of Consumer Research*, *31*(1), 17–25.

Spangenberg, E. R., Sprott, D. E., Grohmann, B., & Smith, R. J. (2003). Mass-communicated prediction requests: Practical application and a cognitive dissonance explanation for self-prophecy. *Journal of Marketing*, *67*(3), 47–62.

Srull, T. K., & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review*, *96*(1), 58–83.

Stayman, D. M., & Aaker, D. A. (1988). Are all the effects of ad-induced feelings mediated by A-sub(Ad)? *Journal of Consumer Research*, *15*(3), 368–373.

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, *132*(1), 23–46.

Strack, F. (1992). The different routes to social judgements: Experiential versus informational strategies. In L. L. M. A. Tesser (Ed.), *The construction of social judgements* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strahan, E. J., Spencer, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: Striking while the iron is hot. *Journal of Experimental Social Psychology*, *38*(6), 556–568.

Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research*, *12*(1), 31–46.

Sujan, M., & Dekleva, C. (1987). Product categorization and inference making—Some implications for comparative advertising. *Journal of Consumer Research*, *14*(3), 372–378.

Svenson, O. (1979). Process descriptions of decision making. *Organization and Human Decision Processes*, *23*, 86–112.

Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science*, *3*, 199–214.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*(3), 183–206.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even—The effects of prior outcomes on risky choice. *Management Science*, *36*(6), 643–660.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, *81*(6), 973–988.

Tversky, A. (1972). Elimination by aspects—Theory of choice. *Psychological Review*, *79*(4), 281–299.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice—A reference-dependent model. *Quarterly Journal of Economics*, *106*(4), 1039–1061.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science*, *39*(10), 1179–1189.

Van Dijk, E., & Van Knippenberg, D. (2005). Wanna trade? Product knowledge and the perceived differences between the gains and losses of trade. *European Journal of Social Psychology*, *35*(1), 23–34.

Varian, H. R. (1992). *Microeconomic analysis* (3rd ed.). New York: Norton.

Verplanken, B., Hofstee, G., & Janssen, H. J. W. (1998). Accessibility of effective versus cognitive components of attitudes. *European Journal of Social Psychology*, *28*(1), 23–35.

Viswanathan, M., & Childers, T. L. (1999). Understanding how product attributes influence product categorization: Development and validation of fuzzy set-based measures of gradedness in product categories. *Journal of Marketing Research*, *36*(1), 75–94.

Wathieu, L. (2004). Consumer habituation. *Management Science*, *50*(5), 587–596.

Williams, P., Fitzsimons, G. J., & Block, L. G. (2004). When consumers do not recognize "benign" intention questions as persuasion attempts. *Journal of Consumer Research*, *31*(3), 540–550.

Winer, R. S. (1986). A reference price model of brand choice for frequently purchased products. *Journal of Consumer Research*, *13*(2), 250–256.

Wright, P. (1975). Consumer choice strategies—Simplifying vs optimizing. *Journal of Marketing Research*, *12*(1), 60–67.

Yeung, C. W. M., & Wyer, R. S., Jr. (2004). Affect, appraisal, and consumer judgment. *Journal of Consumer Research*, *31*(2), 412–424.

Yi, Y. (1990a). Cognitive and affective priming effects of the context for print advertisements. *Journal of Advertising*, *19*(2), 40–48.

Yi, Y. J. (1990b). The effects of contextual priming in print advertisements. *Journal of Consumer Research*, *17*(2), 215–222.

Yoon, C., Gutchess, A., Feinberg, F., & Polk, T. (2006). A functional magnetic resonance imaging study of neural dissasociation between brand and person judgements. *Journal of Consumer Research*, *33*, 31–40.

Zajonc, R. B. (1980). Feeling and thinking—Preferences need no inferences. *American Psychologist, 35*, 151–175.

Zemborain, M., & Johar, G. V. (in press). Attitudinal ambivalence and openness to persuasion: A framework for interpersonal influence. *Journal of Consumer Research*.

Zhou, R., & Pham, M. T. (2004). Promotion and prevention across mental accounts: How financial products dictate consumers' investment goals. *Journal of Consumer Research*, *31*(2), 125–135.

Zillmann, D. (1971). Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology*, *7*(4), 419–439.

Psychology and Politics

The Challenges of Integrating Levels of Analysis in Social Science

PHILIP E. TETLOCK

OVERVIEW

The stakes can be high in political–psychological debates—a factor that has undoubtedly contributed to political psychology's allure as a research specialty (Herrmann, 1986; Knutson, 1973; Sears, Huddy, & Jervis, 2003). The questions researchers pose tend to draw media and government attention: Are prejudice and racism still pervasive forces in American society? Are people inherently tribal creatures whose low thresholds for drawing invidious distinctions make nasty intergroup conflict inevitable? Are conservatives inherently more simple-minded than liberals? Are high-level officials subject to the same judgmental shortcomings as ordinary mortals? Do these officials, as a result, exaggerate the zero-sum character of international conflicts and overrely on competitive, threat-based influence tactics?

Of course, posing questions is one matter; answering them, quite another. Political–psychological questions are notoriously resistant to scientific consensus—in part, because key terms resist translation into precise manipulations and measures and, in part, because partisans of both the left and right are all too eager to reach ideological closure by claiming premature empirical vindication (Kruglanski, 2004; Tetlock, 2005).

This chapter examines four research programs in political psychology that illustrate, but far from exhaust, the theoretical and methodological diversity of the field: (1) attributions of implicit prejudice that rest on controlled laboratory experiments that explore the causes and consequences of millisecond differentials in reaction times to minority versus majority group stimuli; (2) attributions of irrationality in world politics that rest on mixtures of laboratory and field methods of assessing the pervasiveness of judgmental biases among political elites as well as among the mass public; (3) attributions of belligerence in international conflict that rest on mixtures of laboratory and field methods of assessing the relative efficacy of bargaining and negotiation tactics that vary along the deterrence–reassurance dimension; and (4) attributions of mental rigidity that rest largely on psychometric methods of assessing patterns of covariation between measures of cognitive style and political ideology.

This chapter also makes the case that even if we can agree on the basic psychological principles at work in each domain, there is no guarantee of agreement on the political–psychological principles at work. It is one thing to document millisecond differentials in reaction times toward a legally protected group, quite another to characterize that phenomenon as "prejudice." And it is one thing to document a human tendency to rely on low-effort heuristics to simplify choice tasks, quite another to show that public policy at a given moment is "deeply flawed."

It is crucial here to recognize that logically distinct principles govern the descriptive–explanatory and political–normative phases of such debates. In the descriptive–explanatory phase, researchers clash over the robustness of effects (is there something worth ex-

plaining?) and over whose first-order, psychological principles of cognitive and affective functioning best account for those empirical patterns worthy of theoretical attention. This phase is well captured by the phrase "normal science." By contrast, in the political–normative phase of the debate, researchers clash over the wisdom of moving from the relatively value-neutral, descriptive–explanatory language of science to the value-laden language of politics. This phase brings into play competing second-order principles that are ontological–epistemological and moral–ethical in character and that can be used either to trumpet or trivialize the political significance of psychological patterns. The first-order principles are typically spelled out; the second-order ones are typically slipped in. The result is often widespread conceptual confusion over where psychological claims end and political ones begin.

The principal barriers to bridging psychological and political-psychological principles are twofold: (1) contending views of the level-of-analysis problem and of whether the primary burden of proof should fall on advocates of psychological explanations (traditionally organized around intrapsychic processes) or of political explanations (traditionally organized around supraindividual, institutional and systemic, processes); (2) contending views of the fact-value problem and of the degree to which researchers should heed or ignore David Hume's famous is–ought distinction, which stipulates that one can never deduce a value-laden "ought" conclusion from factual or "is" premises. These philosophical fissures run so deep, and arise so frequently in political psychology, that it is worth sketching the canonical arguments in advance.

Level-of-Analysis Disputes

Although rarely articulated quite so baldly, the goal driving most research programs in political psychology is reductionist: reducing complex macropolitical phenomena to their constituent micropsychological processes. The inspirational precedents come from the physical and biological sciences: demonstrations that "heat" is nothing but kinetic energy, that the color "red" is nothing but photon emissions of 600 nanometers, and that DNA is the chemical structure that conveys hereditary information (Hempel, 1965; Suppe, 1977). Applied to political psychology, the reductionist syllogism begins with the major premise that psychology is the science of human behavior. The minor premise is that human beings are the proximal causes of all political behavior. In their role as citizens, they decide whether to treat "outgroups" contemptuously or respectfully; in their role as leaders, they decide whether to engage in ethnic cleansing or to build pluralistic polities. The conclusion is that insofar as psychology achieves its explanatory goals, it cannot avoid shedding light on the central problems of political life (Tetlock, 1998). In the strongest variant of this view, political psychology is merely psychology with political stimulus materials. The basic *processes*—of memory or judgment or emotion—are roughly invariant. All that varies—and sometimes dramatically—is the superficial

content on which those processes operate: Aryans or Jews, Hutus or Tutsis.

There is, of course, nothing wrong with reductionism per se. It has been associated historically with the greatest advances in scientific knowledge (Pinker, 1997). Reductionism does not, however, always work. And it is dangerous to assume automatically that the reductionist model is the most fruitful one to explore in a given domain. Antireductionists in the social sciences—from Durkheim (1925) to Waltz (1979)—have argued that psychological explanations have little to contribute to the explanation of macropolitical and economic regularities. Indeed, some variants of this view argue for putting reductionism in reverse: demonstrating that we only understand the psychological parts of political processes when we understand the (social) system within which the parts function. Relevant scientific precedents would be photosynthesis—a phenomenon that can be reduced to physicochemical processes but whose functional significance can be appreciated only when we situate these processes in their evolutionary context (Suppe, 1977)—or, as Searle (1984) has argued, consciousness—a "first-person" phenomenon that is utterly dependent on the physical workings of the brain but can never be fully captured using the "third-person" objectifying methods of cognitive neuroscience and artificial intelligence.

Applied to political psychology, this argument suggests that psychologists who fail to place their work in its broader cultural and political context are fated to commit fatal errors. One cannot understand foreign policy decision making without a prior grasp of the quasi-anarchic nature of world politics—how the absence of a sovereign (world government) creates security dilemmas for states that are only imperfectly solvable in a limited number of ways (Waltz, 1979). One cannot understand contemporary debates over distributive justice without a prior grasp of how globalization—the increasingly rapid flow of capital, information, and labor across borders—severely strains traditional employment contracts and welfare-state guarantees (Friedman, 1999; Mitchell, Tetlock, Newman, & Lerner, 2003). And one cannot understand the oft-observed faddishness of political psychology without grasping the critical role that wars, social movements, technological transformations, and economic booms and busts play in shaping what researchers deem to be worthwhile problems, how they go about investigating those problems, and where they set their standards of proof for judging competing solutions to those problems (Goldgeier & Tetlock, 2000).

The brusquest dismissals of political psychology by nonpsychologists have, however, traditionally been grounded in a Realpolitik brand of reductionism quite alien to psychologists, a hardball perspective that treats all political thought as an effort to advance the cause of one or another group in the endless struggle for social dominance (Tetlock, 2005). Politics is about "who will dominate whom?"—a point on which an illustrious array of thinkers agree: Thucydides, Hobbes, Marx, Nietzsche, Pareto, and Foucault (McClelland, 1996) and, now from inside political psychology, social-dominance and system-justification theorists (Jost, Pelham, & Carvallo, 2002;

Sidanius & Pratto, 1999). All systems of political thought are ultimately justifications of bids for power. Ideas—a popular microconstruct—are epiphenomenal, smoke to the fire of the real macrodrivers of history. To adapt Clausewitz's famous definition of war as the continuation of politics by other means, psychological research—be it on sex differences or intelligence or prejudice—reduces to the continuation of politics by other means—no matter how pure the avowed epistemic intentions of the scholar.

This chapter stakes out a contextualist position on the reductionism (cf. McGuire, 1983)—a position that acknowledges that the relative usefulness of micro versus macro interpretations of political behavior shifts over time and across topics. But it is only fair to alert readers that the conclusions they draw from the literatures examined here will be unavoidably colored by philosophical assumptions that they rarely have scientific occasions to give much thought. It is not unusual in political psychology for one observer's compelling *reason* for holding a position to be another's wafer-thin *rationalization*. Indeed, in the spirit of Harold Lasswell (1948) and C. Wright Mills (1959), we could elevate this reason–rationalization tension to the status of a cornerstone (second-order, meta-observational) principle. It is safest to assume that political psychologists are every bit as much political actors as are their research subjects—and as subject to the bias of naive realism and the resulting temptation to dismiss dissonant opinions by attributing them to ugly or selfish motives, attitudes, and values (Ross & Griffin, 1991).

Fact-Value Disputes

Research programs in political psychology generally have a moral as well as epistemological agenda. Although rarely articulated so baldly, the driving moral goal is to advance values such as racial equality, international understanding, political tolerance, and peace (cf. Deutsch, 1983; White, 1984). As Shelley Taylor (1998) observed in her history of "the social being in social psychology" (p. 61), some of our discipline's classic experiments from the 1930s, 1940s, and 1950s—such as the work on leadership styles by Lewin, Lippitt, and White (1939)—were driven by concerns "over the dictatorships of Hitler and Mussolini and the goal of demonstrating that "democracy was an inherently superior form of leadership that would produce higher morale and greater productivity than an autocratic style" (p. 61). There is, of course, nothing wrong with aspiring to make the world a better place. But there is potential for confusion, and mischief, when our moral evaluations of attitudes and behavior bleed into our political–psychological descriptions of the same phenomena. In Taylor's words, the findings of Lewin and colleagues "emerged as predicted, although a close examination of the methodology of the studies suggests that, given the guidelines offered to the democratic and autocratic leaders, one could not have expected any other results" (p. 61). It is instructive to pose a turnabout thought experiment: Would this venerable classic have the same status if the researchers had stacked the methodological deck in favor of an autocratic, philosopher-

king style of leadership, with instructions that trumpeted the advantages of decisive visionary leadership committed to achieving long-term goals requiring temporarily unpopular sacrifices? Or, would fame turn into infamy?

Positivists and their descendants have traditionally been the strongest advocates of a strict partitioning of facts and values (Boring, 1950). Within this philosophy-of-science tradition, researchers run afoul of Hume's famous is–ought syllogism when they start opining on the defensibility of research participants' opinions, of pronouncing them to be prejudiced, racist, irrational, belligerent, naively simplistic, or incorrigibly rigid. The researchers' expertise is "factual," and they overstep their professional competence when they blur the distinction between their views of what is and what ought to be (Suedfeld & Tetlock, 1991). As MacCoun (1998) noted in his review of political biases in social science, researchers also run afoul of Robert Merton's (1987) constitutive norms of science when they let their moral values creep into their characterizations of fact. Those norms include "universalism" (judge scientific accomplishments by impersonal criteria—the politics of the researcher should be irrelevant), "disinterestedness" (proceed objectively, with no regard for whose ox might be gored), and "organized skepticism" (organize yourselves into communities dedicated to rigorous peer review of all claims).

Nonetheless, many social scientists and philosophers now agree that a purely descriptive data language is an impossible ideal and that the fact-value distinction is neither as absolute nor as compelling as mid-20th-century positivists implied (Proctor, 1991; Putnam, 2002). There is no value-neutral vantage point for analyzing social phenomena (no view from nowhere, to use Nagel's [1986] felicitous phrase). Even the simplest scientific acts—such as setting a significance level for accepting or rejecting hypotheses—presuppose value judgments of the relative importance of making Type I versus Type II errors. It is inevitable—and some defiantly declare desirable—that the values of scientists seep into every facet of their work—from framing of hypotheses to choice of methodology to interpretation of findings.

This chapter stakes out a middle-ground position in the fact-value debate—one that concedes the impossibility of pure value neutrality but that stresses the importance of making good-faith efforts to disentangle factual from moral judgments of social phenomena (MacCoun, 1998). But, again, it is fair to alert readers that the conclusions they draw will be unavoidably colored by foundational assumptions they rarely see scientific reason to articulate. It is not unusual in political psychology for one observer's responsible contribution to a policy debate to be another's irresponsible politicization of the scientific process (Sears, 2004; Tetlock, 1994). Indeed, we should count such occurrences as special cases of the more fundamental reason–rationalization principle.

Point of Departure

Each of the four lines of work examined here raises messy mixtures of psychological, philosophical, and

moral–political issues. Scholarly assessments of the justifiability of attributions of racism or prejudice or irrationality or belligerence hinge not just on evidence but on level-of-analysis assumptions about what constitutes a plausible alternative explanation and on moral assumptions about the defensibility of particular attitudes (and when scholars are justified in letting their own moral views shape their characterizations of the views of others). Political psychology is arguably the only domain of psychology that *routinely* needs ideological subscripts for its dependent variables—subscripts to track the shifting labels that contending factions apply to our measures. Readers can take this state of affairs as evidence that political psychology is hopelessly relativistic and "perspectival," a field of interest only to anti-social-science constructivists and postmodernists (Gergen, 1978); or they can take this state of affairs as an indication that political psychology poses greater-than-usual scientific challenges that require us to model the mindsets not just of research participants but of the researchers themselves.

ATTRIBUTIONS OF IMPLICIT PREJUDICE TO MASS PUBLICS

Cognitive psychologists have used priming techniques for decades to explore fundamental associative laws of human memory (Higgins, 1996). In a typical priming experiment, a "prime" word is presented before a target word. The participant is instructed to react to the target word quickly, usually by pressing a key or pronouncing the target. The semantic relation between the prime and the target is of central interest in this study. To the extent that the prime influences the response to the target word, the two stimuli are deemed to be associated. Some primes facilitate the response to the target word, as when the prime is "bread" and the target is "butter." This facilitation is widely theorized to be due to the automatic activation of mental representations of the target by the earlier presentation of the closely related prime.

In a pioneering study, Fazio, Sanbonmatsu, Powell, and Kardes (1986) extended this priming technique to attitudes. For example, a prime might be a negatively valenced word such as "murderer." If the subsequently presented target is evaluatively congruent, such as "evil," the evaluation of the target would be accomplished more quickly than if the target were "happy." In another pioneering study, Fazio, Jackson, Dunton, and Williams (1995) extended this procedure to include pictures of White and African American faces as primes. The targets were adjectives with either a positive (e.g., wonderful) or negative (e.g., annoying) connotation. Participants had to push either the "good" key or "bad" key as quickly as possible. Fazio and colleagues (1995) found that among the White respondents, reaction time to good words was quicker following presentation of the White faces, and reaction time to bad words was quicker following the African American faces. Just as "bread" facilitated response times to the target "butter" due to their proximity in the associative network, Fazio and colleagues concluded that the participants in their study associated positivity more

closely with Whites and negativity more closely with Blacks—and that their technique might provide a bona fide (as opposed to bogus) pipeline into racial attitudes that people would rather not admit to having.

One critical feature of the affective-priming technique is the fleeting interval between the onset of prime and target. Because the interval is often as short as 300 msec, there is a strong implication that the cognitive processes underlying affective priming must be due to automatic semantic activation beyond conscious control. Even more decisive evidence for automaticity comes from the work of Wittenbrink, Judd, and Park (1997), who demonstrated subliminal priming effects.

Since the path-breaking studies of affective priming, and the parallel development of the Implicit Association Test (IAT; Greenwald & Banaji, 1995), there have been many efforts to assess the validity of claims that implicit-associative measures are indeed tapping into a construct that can be accurately labeled implicit prejudice. Banaji (2001) has stated the reductionist logic of her research program with exemplary clarity:

> If attitude is "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly & Chaiken, 1998, p. 269), and if implicit memory is "revealed when previous experiences facilitate performance on a task that does not require conscious or intentional recollection of these experiences" (Schacter, 1987, p. 501), then implicit attitudes are introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects (Greenwald & Banaji, 1995, p. 8). (p. 119)

Working from these first-order psychological principles, it is a short inferential step to conclude that the same methods that cognitive psychologists use in studying implicit memory should also prove useful in studying implicit attitudes. It requires, though, an inferential leap to conclude that the research community is justified in equating implicit attitudes and prejudice (as psychologists use those terms) with prejudice or racism (as ordinary people use those terms). Justifying that inference requires demonstrating that implicit-associative measures are reliably and strongly correlated with compelling criterion variables that capture discriminatory treatment (Tetlock & Arkes, 2004).

This rapidly growing literature has been ably reviewed elsewhere (Fazio & Olson, 2003) so there is no point in rereviewing it. The focal point of political–psychological controversy is over whether greater implicit negative affectivity toward Black-versus-White racial cues is sufficient evidence to indict large percentages of Americans as implicitly prejudiced (Redding, 2004). For when social psychologists use the politically charged language of prejudice and racism to characterize their findings, they enter—like it or not—the larger political battle over the tenacity of prejudice: Has racism lost most of its sting in the post–Civil Rights era or has it merely taken on new more subtle forms? The answers that policy elites reach to this question prove—perhaps not surprisingly—to be a potent predictor of how supportive those elites are of govern-

ment programs especially targeted at helping African Americans (cf. Kinder, 1998; Sniderman & Carmines, 1997).

Taken at face value, the self-report survey data suggest that although old-fashioned racism has not disappeared, it has been dramatically weakened as a political force. Overt White hostility toward African Americans declined precipitously in the 1960s and by the 1990s had reached historic lows. Whereas White Americans were once deeply divided over whether African Americans should be accorded the full rights of citizenship, there is now near consensus at the level of both mass and elite opinion that *de jure* segregation and doctrines of inherent racial superiority are unacceptable (Schuman, Steeh, Bobo, & Krysan, 1997).

These stylized facts are not in dispute, but there is dispute over their significance. Observers of race relations on the political left doubt that self-report surveys provide an accurate barometer of prejudice. They insist that racism still lies "at the center, not the periphery, in the permanent not in the fleeting, in the real lives of black and white people, not in the caverns of the mind" (Bell, 1992, p. 208). Claude Steele's research on stereotype threat is often taken to reinforce this worldview: African Americans have apparently so internalized negative societal stereotypes of themselves that even trivial reminders of their racial identity impair their academic performance (Steele, Spencer, & Aronson, 2002). By contrast, observers on the right tell a radically different story. Shelby Steele's (2006) writings imply that stereotype-threat effects arise because reminding African Americans of their racial identities on elite campuses makes salient SAT-admission differentials of racial groups, thus activating an ideology of victimology that reduces the need to try hard on academic tests. In a similar vein, Thernstrom and Thernstrom (1997) maintain that although African Americans have made large socioeconomic gains, the remaining pockets of inequality are best explained not by White racism but by racial gaps in achievement values, educational attainment, and family structure rooted in forces endogenous to Black communities.

Social psychologists have largely lined up with observers who stress the lingering power of prejudice and doubt. They are skeptical of the depth of the changes in racial attitudes documented in representative-sample surveys. Many Whites, they suggest, have learned to say the right thing, but they have not truly internalized the egalitarian ideals that would justify calling them non-racist (Greenwald & Banaji, 1995; Jackman & Jackman, 1983). Framed within the broader debate, though, the intense interest among social psychologists in implicit prejudice is inevitably viewed with suspicion among conservatives, as the efforts of a largely liberal research community to find evidence that bolsters government efforts to lend special assistance to Blacks because Blacks confront special obstacles to socioeconomic advancement (Redding, 2004; Stein, 2004).

Defenders of the implicit-prejudice research program counter that the burgeoning body of construct-validity research is already sufficient to sustain the prejudice in-

terpretation. They argue that research to date has revealed a lot about how implicit measures of prejudice correlate with a variety of criterion measures (for a recent meta-analysis of 60 IAT studies, see Poehlman, Uhlman, Greenwald, & Banaji, 2005). Particularly compelling in their view is evidence that:

1. Implicit-associative measures are positively correlated with explicit measures of prejudice such as the Modern Racism Scale, especially after correcting for attenuation due to measurement error, thus providing convergent validity across very different methods for the racism interpretation (e.g., Cunningham, Preacher, & Banaji, 2001).
2. Implicit-associative measures predict avoidance and awkwardness in interpersonal encounters with African American experimental confederates (nonverbal cues, such as interpersonal distance, eye blinking and gaze aversion, and paralinguistic cues, such as speech fluency and stuttering) and a preference for sitting beside and interacting with Whites, thus providing criterion-validity support for the racism interpretation (Dovidio, Kawakami, & Gaertner, 2002; Monteith, Voils, & Ashburn-Nardo, 2001; Poehlman et al., 2005).
3. Implicit-associative measures are associated with amygdala activation (taken by some as a neurological locus of hostility and aggression), thus further underscoring the power of the racism interpretation to unify findings across diverse methodologies (Chee, Sriram, Soon, & Lee, 2000; Phelps et al., 2000).

Defenders of implicit-prejudice interpretations of implicit-associative measures do not disguise their impatience with more benign shared-cultural-stereotype interpretations of the same measures (e.g., Karpinski & Hilton, 2001): the notion that African American stimuli prime negative associations more readily not because people hold such associations themselves but rather because they are aware that such associations have a long history in America. Banaji (2001) sees a false dualism at work, a tendency to distinguish too sharply between the true internal self and the external cultural environment. Our "selves" are the product of our social learning history and a racist culture inevitably leaves its imprint on our most readily accessible mental representations of the world. In Banaji's words:

> Perhaps the struggle to find a place to point the finger, to take the burden of possession off oneself, comes from the inherently political nature of such assessments. We certainly don't see the same agitation when we can't seem to remember a list of words for which we show intact priming. Individuals are the transducers of cultural experience—they provide the physical, social, and psychological shell through which culture speaks. Yet when revealed attitudes are not palatable, the reaction is to look for an answer elsewhere and pointing to culture (not as the environment in which the attitude is learned but rather as the thing whose attitude is being measured) is perfectly understandable and perfectly wrong. (p. 138)

Banaji's argument raises profound level-of-analysis questions about individual identity—and the meaningfulness of holding individuals (as opposed to political cultures) accountable for their attitudes—be they implicit or explicit. On a humbler empirical plain, though, it is worth noting that, in the spirit of signal-detection theory (Swets, 1982), there are two distinct sets of issues on which proponents and opponents of implicit-prejudice arguments might disagree:

1. Over the facts—the degree to which the construct-validity evidence convincingly shows that unconscious negative affect drives differential response times to minority-group stimuli in implicit-prejudice research paradigms.
2. Over values—the degree to which existing evidence passes their threshold for labeling people "prejudiced" or "racist." Researchers with different worldviews may attach different values to the risks of Type 1 errors (calling nonracists racists) versus Type II errors (calling racists nonracists) and these different values may lead them to set different thresholds for making accusations of racism.

It turns out, as is so often the case in political psychology, that the feuding camps disagree on both the facts and the values at stake (cf. Jervis, 1976). In a scholarly exchange with implicit-prejudice researchers, Tetlock and Arkes (2004) challenged whether the research community had given due weight to alternative, usually more benign explanations for the patterns of covariation between implicit-prejudice measures and various construct-validational indicators. These alternative explanations included:

1. The possibility, raised by Karpinski and Hilton (2001) and Fazio and Olson (2003), that implicit measures are picking up not racial animus but, rather, awareness of unflattering, media-amplified stereotypes of African Americans (or of particular subgroups, such as young, inner-city males) that participants would themselves disavow if those associations were called to their attention. As Judd, Blair, and Chapleau (2004) observe in their study designed to sort out why African American, relative to White, faces automatically facilitate the categorization of handguns: "It is one thing to believe that the tendency to misclassify weapons in the hands of young African Americans is part of a negative, prejudicial bias on the part of police officers. It is a rather different thing to believe that the officers are influenced by a widely shared stereotype of African-Americans as both violent and athletic, unreliable and possessing a good sense of humor, lazy and street-smart" (p. 77). For Judd and colleagues, these two assessments imply empirically distinguishable views of American society; for Banaji (2001), the distinction verges on meaningless, "perfectly understandable and perfectly wrong."
2. The possibility that the target construct, implicit prejudice, may be picking up on racial animus but animus that can be readily picked up by traditional self-report measures of prejudice and animus that is concentrated among relatively small numbers of bigoted outliers. If the goal is to enter the great American racism debate—on a grand societal scale—it is essential to show that the new prejudice is not reducible to the old—and that the effects are not being driven by isolated pockets of old-fashioned prejudice (Mitchell & Tetlock, 2006).
3. The possibility that the target construct is a multidimensional one that derives from a messy mix of affective associations, including anger at the injustices of racism, shame and embarrassment at past or current White mistreatment of Blacks, and sadness at missed opportunities to achieve fuller racial equality. In the parable of the "two Jesses," Tetlock and Arkes (2004) posit two respondents with markedly different associative networks for encoding information about African Americans. One, like Jesse Jackson, is sympathetic to a liberal policy agenda: He believes that racial discrimination is an ongoing, not just a past, problem and supports aggressive affirmative action to level the playing field. The other, like Jesse Helms, is sympathetic to a conservative agenda: He believes that, just as other minority groups had to work their way up the success ladder, so should African Americans. This respondent rejects affirmative action and believes the primary cause of racial inequality is now internal to the African American community: the breakdown of personal responsibility within inner-city communities. These eponymous figures agree that crime rates are too high and test scores too low, but they make different causal attributions for these facts and experience different mixtures of sorrow and anger. Is there compelling theoretical reason for expecting the IAT or other implicit associative measures of "negative affectivity" to differentiate these ideal-type belief systems, to differentiate people who have the same amount of negative affect but whose affect can be traced to qualitatively different emotional states and causal attributions for inequality (cf. Lerner & Keltner, 2001; for evidence that sympathy for underdogs can translate into negative implicit affect, see Uhlmann, Brescoll, & Paluck, 2006)?
4. The possibility that sizable fractions of the variance in IAT or other test scores might be attributable to individual differences that are neither associationist nor attitudinal, such as familiarity (Rothermund & Wentura, 2004), test anxiety (Frantz, Cuddy, Burnett, Ray, & Hart, 2004), general processing speed (Blanton & Jaccard, 2006), or cognitive dexterity (Fiedler, Messner, & Bluemke, 2006), and that these confounding influences undercut claims of natural zero points and ratio scales (Blanton & Jaccard, 2006).
5. The possibility that some component of the negative affectivity is traceable to realistic assessments of covariations between negatively valenced attributes—crime, out-of-wedlock births, school dropout rates—and racial categorizations (cf. Jussim, 2005). Tetlock and Arkes (2004) worry that implicit-prejudice research has set the threshold for labeling people prejudiced so low that most people will qualify as long as they live in societies in which intergroup inequalities exist on dimensions to which people assign evaluative significance (e.g., pov-

erty, crime) and they recognize, consciously or unconsciously, the existence of these inequalities. This argument implies that we need to control statistically for something no one has yet controlled: political knowledge. Implicit racism may be implicit awareness of depressing inequalities—a view consistent with experimental demonstrations of the power of positive social contexts to attenuate automatic negative affect in response to African Americans (Barden, Maddux, Petty, & Brewer, 2004). This possibility led Tetlock and Arkes to ask whether Jesse Jackson (who once confessed to relief on discovering that the young men walking behind him at night were White, not Black) would "fail" the IAT? It also led them to pose the corollary question: Is "Bayesian bigot" an oxymoron or is it reasonable to issue moral indictments for doing what decision theorists would call the right thing—using base-rate data?

These questions imply a further question: Is so politically charged a debate scientifically resolvable? Here we find a spectrum of views. At one end are those, such as Banaji and colleagues (2004) and Sears (2004), who believe the answer is an emphatic "yes": The scientific method is unfolding as it should and critics are just too impatient. It takes time to build up the requisite body of construct-validational evidence. At the other end, we find a perspective so philosophically extreme that no one in the debate has yet to endorse it: depoliticize the dispute by purging political hot-potato terms from our scientific vocabulary, abandoning all pretense of relevance, and replacing value-laden talk about "prejudice" and "racism," with dry technical talk about the "dimensionality of affectivity." From a history-of-psychology perspective (Boring, 1950), it is striking though that this proposal has so few advocates. Positivist traditionalists would be puzzled: Why should social scientists, in their roles as scientists and not citizens, be in the business of making pronouncements of racism and prejudice? Why can't researchers be content to say that they have documented a pattern of implicit affect, the behavioral correlates take certain forms, and the attentive public can now draw its own moral and legal conclusions about whether the correlates rise to actionable discrimination? Prejudice is more than a scientific judgment: It requires passing judgment on which aspects of public opinion (implicit or explicit) deserve our censure.

A middle-ground option is to accept that the dispute is inherently political and that special precautions may be necessary for ensuring reasonable ideological balance in the hypothesis-testing process. One path to this end is via Bayesian reputational bets and adversarial collaborations (cf. Mellers, Hertwig, & Kahneman, 2001). This process requires researchers from competing camps to agree, *ex ante*, on what would constitute fair tests of their respective positions and to update their beliefs in accord with the results.

Recent debates suggest that adversarial collaboration would reveal that different camps have markedly different views on the diagnosticity of the evidence—be it eye blinks or amygdala activation or factor analyses of response latency data. One tough test of the viability of ad-

versarial collaboration arises when we explore clashing views on the theoretical implications of correlations between implicit and explicit attitudinal measures. Fazio and Olson (2003) note that this is a messy literature: Correlations have ranged widely and there seem to be no theoretical constraints to prevent researchers from adopting a "heads-I-win-tails-I-do-not-lose" stance in which positive correlations suggest convergent validity, zero-order correlations suggest discriminant validity, and negative correlations suggest repression or overcompensation.

Banaji and colleagues (2004) make the convergent validity argument: after correcting for unreliability, the correlations between the IAT and certain explicit measures of prejudice can rise as high as 0.5. But skeptics are unpersuaded: Tetlock and Arkes (2004) point out that one of the most popular explicit measures—the Modern Racism Scale—is itself controversial, with some viewing it as a measure of prejudice and others viewing it as a measure of a political preference for the more conservative position that African Americans collectively should rely more on themselves and less on government (see the symbolic racism debate between Sniderman and Tetlock [1986] and Kinder [1986]. Should we conclude that slower responders to the IAT are implicit racists—as Banaji and colleagues suggest? Or does it just mean that these respondents believe, rightly or wrongly, that the primary obstacles to racial inequality now reside in the African American community?

Skeptics maintain that the data are insufficiently diagnostic with respect to whether the IAT is tapping implicit prejudice. But it is fair to ask: What would induce them to change their minds? It is also fair to ask what associations between implicit and explicit measures of prejudice would induce defenders of implicit-prejudice interpretations to change their minds. The answers are not obvious. For instance, when implicit measures predict explicit indicators of prejudice, investigators can construe such findings as construct-validation. But when such measures fail as predictors, investigators can interpret such findings as evidence of how skillful respondents are at cognitively overriding prejudicial impulses. Moreover, the most reliable individual-difference moderator of tendency to override—the motivation to control prejudice scale developed by Fazio and colleagues (1995)—does not break this impasse. This carefully constructed scale is also open to alternative interpretations. "Prejudice" is—it should now be clear—a term with ever-shifting political meanings. Are high scorers on the scale reining in raw bigotry or are they displaying hypersensitivity, even to matter-of-fact racial differences—exactly the hypersensitivity that one would expect if they feared that esteemed professors would label them as racist if they displayed even trace awareness of unpleasant facts of life in our society? Indeed, as Frantz and colleagues (2004) ask: Is the sinking sense that one is failing the IAT sufficient to induce stereotype threat among whites (fear that they fit the profile of the modern racist) and hence the oft-observed interpersonal awkwardness?

From a contextualist perspective, prejudice is not the epistemological equivalent of a rock—the true nature of

which we can ascertain if we only subject it to the right tests. Researchers need to acknowledge the wide spectrum of meanings that political observers assign to this popular conceptual variable. At one end are the most restrictive definitions of prejudice that require satisfying all three of Allport's (1954/1968) criteria: attitudes based on misinformation, charged with hostility, and unresponsive to evidence. Such definitions raise the risk of "missing" sneaky racists but lower the risk of false alarms. At the other end are the least demanding definitions that presume none of the three Allportian criteria. For Banaji and colleagues (2004), correlation with any form of disparate impact suffices. Such definitions raise the opposite risks. The dispute is thus as much about values as about facts, about where we should set our standards of proof for issuing indictments for prejudice. We have reached an epistemological, not just a disciplinary, boundary.

ATTRIBUTIONS OF IRRATIONALITY TO POLITICAL ELITES

From Lasswell (1948) to Janis (1982) to Suedfeld (1992), many political psychologists have subscribed to variants of the following reductionist syllogism: Policy elites are all too human (notwithstanding occasional efforts to promote themselves to deity status). Psychology is the science of human behavior. As such, psychology must shed light on why even the high and mighty think, feel, and act as they do.

Of course, reductionists need not agree on which first-order psychological principles confer the greatest explanatory leverage. The oldest tradition is the psychodynamic, which traces its lineage from Freud, Langer, and Erikson to contemporary scholars such as Etheredge (1980), Greenstein (1975), and Winter (1993). But the greatest activity in the last three decades has concentrated on applying first-order principles from the vast literature on errors and biases in human judgment at both the individual (Gilovich, Griffin, & Kahneman, 2000; Kahneman & Tversky, 2001) and group-dynamics levels (Aldag & Fuller, 1993; t'Hart, 1994). Whether working solo or in small groups, the recurring theme is how often people stray from key benchmarks of rationality. People, it is alleged, are too slow to change their minds in response to dissonant evidence, too quick to jump to conclusions from ambiguous evidence, loath to confront tough trade-offs, unaware of past mistakes, and excessively confident in their current predictions.

Cognitive Biases

Virtually all work in this tradition builds on three premises:

1. World politics is staggeringly complex—there are too many possible causes and interactions among causes, and not nearly enough observations to test these hypotheses, forcing us to rely on speculative counterfactual thought experiments (Tetlock & Belkin, 1996).

2. People—limited capacity information processors that we are—frequently resort to simplifying strategies to deal with this otherwise unmanageable ambiguity and complexity (Axelrod, 1976; George, 1980; Herrmann & Fischerkeller, 1995; Jervis, 1976; Mintz, 1993; Vertzberger, 1990). Although there is disagreement on how to characterize these simplifying strategies (proposals include the usual suspects: heuristics, schemas, scripts, cognitive maps, operational codes, associative networks, etc.), there is consensus on their indispensability. Decision makers need to feel they have reasonably firm answers to fundamental questions about the world. What are the basic objectives of other states? What should our own objectives be? Can conflict be avoided and, if so, how? If not, what form will conflict take?

3. There is, however, a price to be paid for the cognitive and political benefits of a simple, stable worldview. The price of cognitive economy in world politics is—as in other domains of life—susceptibility to error and bias (although there is disagreement over the price that must be paid, with a minority putting it startlingly close to zero) (Gigerenzer, Todd, & the ABC Research Group, 1999).

This argument runs into strong headwinds when the time comes to operationalize bias in real-world settings. In the laboratory, claims about deviations from rationality are hardly beyond dispute, but they at least rest on well-defined normative baselines (Kruglanski, 2004). Researchers can manipulate the order in which evidence is presented (assessing primary/recency effects), the directionality of the evidence (assessing double standards), the diagnosticity of the evidence (assessing whether people are good Bayesian belief updaters who live up to reputational bets), and the precise probabilities and monetary values of outcomes (assessing whether decision makers act in accord with expected utility theory or fall prey to arbitrary faming effects). It is a much tougher challenge to defend the same normative verdicts in the real world: When should the financial masterminds of Long-Term Capital Management have recognized that their hedging strategies were failing (Loewenstein, 2000) or when should Sovietologists have recognized that the Soviet Union was imploding (Breslauer & Tetlock, 1991), or when should policy elites stay the course in Vietnam, Iraq . . . or write off sunk costs?

Given the sheer bulk of laboratory and field evidence in support of the error-and-bias portrait of human nature, many psychologists might understandably conclude that relevance to policy elites has already been established (Gilovich, Griffin, & Kahneman, 2000). Relevance is not, however, enough to clinch the scientific case. Even within psychology, there are ongoing challenges to the empirical robustness of the error-and-bias portrait. Does the portrait hold up when data are presented in more naturalistic ways or when decision makers have incentives to become more thoughtful (Camerer, 1995; Gigerenzer et al., 1999)? Within psychology, there are also challenges that maintain that too much attention has been given to conditions under which simple heuristics lead us astray and not enough to conditions under which such heuristics perform surprisingly well (Gigerenzer

et al., 1999). Outside psychology, skeptics working at more macrolevels of analysis have advanced their own external-validity grounds for doubt, including (1) selection arguments (to make it into high-level political roles, one must be a lot more mature than the typical sophomore participant in laboratory studies); (2) motivational arguments (policymakers are especially motivated to learn from their mistakes and to make rational decisions because the stakes are so high); and (3) accountability-constraint arguments (even if policymakers were prone to the same effects as laboratory participants, they work within complex systems of checks and balances, and of elite debate and competition for power, that prevent psychological biases from being translated into policy) (see Druckman, 2004).

From McGuire's (1983) contextualist perspective, each argument highlights testable boundary-condition hypotheses. One strategy of testing them is to explore whether hypothesized threats to generalizability make a difference when manipulated in experiments. The results are reassuring to psychologists, if not to the rest of the world. Increasing the stakes is not a surefire method of checking biases such as belief perseverance and overconfidence (Camerer & Hogarth, 1999); more sophisticated subject pools also appear susceptible to basic biases and errors (Gilovich et al., 2000); and the only types of accountability that regularly attenuate bias are tricky to orchestrate in the political world (predecisional accountability to well-informed audiences with unknown views; Tetlock & Lerner, 1999).

Another strategy is to explore whether we obtain similar functional relationships between the conceptual independent and dependent variables studied in laboratory and real-world settings? The methodological division of labor seems straightforward: controlled experiments gauge the internal validity of the original causal claim whereas field methods—case studies, content analysis, codifying expert consensus—gauge how well the hypothesized process holds up in the hurly-burly of world politics. And when multimethod convergence arises in political psychology, it is understandably taken as an encouraging sign that reductionism works. The guiding assumption is that claims that pass radically different tests stand a better chance of being true than do claims whose support is confined to one genre of research.

By this standard also, the error-and-bias portrait holds up well in the political domain, with the caveat that most external-validational work would not pass muster among experimentalists. The work largely consists of comparative case studies, raising obvious concerns about selective sampling of cases (Do we even know how to draw a representative sample from the ill-defined population of decision-making episodes?) and about selective interpretation of cases (Are we giving "due" weight to alternative explanations?). The next sections illustrate the strengths and weaknesses of the evidence.

Heuristic-Driven Analogical Reasoning

People try to understand novel problems by reaching for familiar concepts. Frequently, these concepts take the form of metaphors and analogies that illuminate some aspects of the problem but obscure other aspects.

Lakoff and Johnson (1980) argue that metaphors are cognitive–linguistic structures that pervade all forms of discourse, and international politics is no exception. Moreover, metaphorical preferences are correlated closely with policy preferences. Consider the "ladder of escalation" and the "slippery slope" (Jervis, 1989). The former metaphor implies that just as we can easily climb up and down a ladder one step at a time, so we can control the escalation and deescalation of conventional or even nuclear conflicts; the latter metaphor implies that once in a conflict, leaders can easily lose control. In Cold War days, backers of the "ladder" metaphor supported a war-fighting doctrine that stressed cultivating counterforce capabilities; they believed that nuclear war could, in principle, be controlled and even won. By contrast, "slippery slopers" endorsed MAD (mutual assured destruction)—both as a policy and strategic reality—and they feared that, once initiated, conflicts would inevitably escalate to all-out war. They argued that nuclear powers need to avoid crises. Managing them once they break out is too risky. As President Kennedy remarked after the Cuban Missile crisis, "One can't have too many of these" (Blight, 1990).

People also bestow meaning on new situations by drawing on salient historical precedents (Gilovich, 1981; Neustadt & May, 1986; Vertzberger, 1990). But this essential strategy, too, can be abused. One mistake is to dwell on the most obvious precedent—a pivotal event early in one's career (Goldgeier & Tetlock, 2000) or perhaps the most recent crisis (Jervis, 1976; Reiter, 1996)—rather than survey a broader set of precedents. Consider the potpourri of conflicts that American observers in the elite press compared to Vietnam between 1975 and 2006: Lebanon, Israel's Vietnam (twice over); Eritrea, Ethiopia's Vietnam; Chad, Libya's Vietnam; Angola, Cuba's Vietnam; Afghanistan, the Soviet Union's Vietnam; Bosnia, the European Community's Vietnam; Kashmir, India's Vietnam; Nicaragua, Bosnia, Somalia, and Iraq (twice over) all new American Vietnams, and, of course, Kampuchea, Vietnam's Vietnam. To be sure, there are similarities, but the differences are also marked and often slighted.

Khong (1991) reports what is still the most elegant study of analogical reasoning in foreign policy. Coding high-level deliberations in the early 1960s, he documents how American policy in Vietnam was shaped by the perceived similarity of the Vietnamese conflict to the Korean war. Once again, a Communist army from the north had attacked a pro-Western regime in the south. This diagnosis led to a series of prescriptions. The United States should resist the aggression with American troops and could expect victory, albeit with considerable bloodshed. A side-constraint lesson drawn from the Korean conflict was that the United States should avoid provoking Chinese entry into the Vietnam war and practice "graduated escalation."

This example illustrates a second pitfall in analogical foreign policy reasoning: the tendency to neglect differences between the present and the preferred precedent.

In public but also in private, policymakers too rarely engage in balanced comparisons (Neustadt & May, 1986). From a psychological viewpoint, this result is reminiscent of laboratory research on the overweighting of hypothesis-confirming information (Klayman & Ha, 1987). To reinvoke Vietnam, American policymakers concentrated on surface similarities between the Vietnamese and Korean conflicts while George Ball—alone within Johnson's inner circle—noted the differences (e.g., the conventional vs. guerrilla natures of the conflicts, the degree to which the United States could count on international support). Whereas doves complained about this analogical mismatching, hawks complained about mismatching that led decision makers to exaggerate the risk of Chinese intervention. China had less reason to intervene in Vietnam in 1965 than in Korea in 1950, preoccupied as Beijing was in the late 1960s by the internal turmoil of the Great Proletarian Cultural Revolution and the external threat of the Soviet Union, which had just announced the Brezhnev Doctrine (claiming a Soviet right to intervene in wayward socialist states).

A third mistake is to permit preconceptions to drive the conclusions one draws from history. In the United States, hawks and doves drew sharply divergent lessons from the Vietnam war (Holsti & Rosenau, 1979). Hawks learned that the Soviet Union was expansionist and that the United States should avoid graduated escalation and honor alliance commitments. Doves learned that the United States should avoid guerrilla wars, that the press is more truthful than the administration, and that civilian leaders should be wary of military advice. No lesson appeared on both the hawk and dove lists (Zimmerman & Axelrod, 1981).

Belief Perseverance

Extending laboratory demonstrations of cognitive conservatism and belief perseverance (Nisbett & Ross, 1980), students of foreign policy often find that policymakers resist changing their minds (George, 1980). They also find that cognitive mechanisms, such as selective attention to confirming evidence, source derogation, and biased assimilation, play key roles in buffering beliefs from refutation. And they find that, although foreign policy belief systems take many forms (Herrmann & Fischerkeller, 1995), one of the most resistant-to-change configurations of beliefs is the inherent bad-faith model of adversaries (Blanton, 1996; Holsti, 1967; Murray & Meyers, 2004). A state is believed to be implacably hostile: contrary indicators, that in another context might be regarded as probative, are ignored, dismissed as propaganda ploys, or interpreted as signs of weakness—for example, American and Soviet images of each in the Cold War, Israeli and Palestinian, and Indian and Pakistani views of each other up to the present (Kelman, 1983). Although such images are occasionally on the mark, they can produce missed opportunities for conflict resolution (Spillman & Spillman, 1991). More generally, belief perseverance can slow shifts from less to more successful strategies. World War I offers one of the more horrendous examples: Military strategists continued to launch infantry charges despite enormous losses, leading to the wry observation that men die more easily than beliefs (Art & Waltz, 1979, p. 13).

Recent postmortems of intelligence failures—with respect to 9/11 and Iraqi WMD (weapons of mass destruction)—highlight the conditions conducive to belief perseverance. Jervis (2006) shows that a theme running through critiques of intelligence agency estimates of Iraqi capabilities is the readiness with which experts see what they expect to see—and what they suspect their political masters want them to see. Jervis also notes a core irony of the Iraqi WMD controversy: to deter his enemies, Saddam Hussein wanted those enemies to suspect that he possessed WMD capabilities, thereby creating exactly the type of ambiguity that his powerful foe, the United States, could use to justify demolishing his regime. That said, postmortem critiques are themselves methodologically problematic. In an insightful critique of the 9/11 Presidential Commission's criticism of American preparedness of the terrorist attacks on 9/11, Posner (2005) notes how hindsight bias fosters the dangerous illusion that it should have been easy for analysts to have connected the dots and prevented the terrorist attacks.

Compelling real-world evidence for irrational belief perseverance is hard to get. It requires an unusually long investigative attention span: First, to persuade professional observers of the political scene to specify the *ex ante* conditional likelihood of each member of the set of all possible futures (conditional on their own vs. rival views of the world being correct); second, to wait several years to find out which views spawned more accurate expectations, and only then to be able to apply the Bayesian belief updating formula to identify who was willing to change his or her mind by the prescribed amounts. Tetlock (1999b, 2005) reports a batch of such studies and finds that (1) experts whose most likely futures fail to materialize rarely change as much as they should given the reputational bets those experts themselves endorsed *ex ante*, but experts whose most likely futures do materialize tend to increase their confidence to roughly the right extent; (2) experts who get it wrong resist changing their minds by invoking a formidable battery of dissonance-reduction strategies, including the close-call-counterfactual defense (what I predicted did not happen but it almost did), the off-on-timing defense (it has not happened yet but it will), and the Type I/II error defense (yes, I was wrong but I made the right mistake); and (3) experts with high needs for closure and a preference for parsimony (hedgehogs) were more likely to invoke belief system defenses and to resist changing their minds when they got it wrong. These results converge with experimental work (Kruglanski, 2004; Suedfeld & Tetlock, 2001).

It is dangerous, however, to overstate the power of belief perseverance. Policymakers do sometimes change their minds (Levy, 1994; Tetlock, 2005). The key questions are: Who changes? Under what conditions? And what forms does change take? Converging evidence from archival studies of political elites and experimental studies of judgment suggest at least four possible answers:

1. When policymakers do change their minds, they are constrained by the cognitive-consistency principle of least resistance and show a preference for tweaking cognitions with minimal connections to other cognitions (Abelson, 1959). Policymakers should thus abandon beliefs about specific tactics before giving up on a strategy and abandon strategies before jettisoning assumptions fundamental to their worldview (Breslauer & Tetlock, 1991; Spiegel, 1991).

2. Timely belief change is more likely in competitive markets (Smith, 1991) that provide quick, unequivocal feedback and opportunities for repeated play on similar problems so that base rates of experience can accumulate, thereby reducing reliance on theory-driven speculation about what would have happened if one had chosen differently. Policymakers should thus learn more quickly in currency and bond markets than, say, in the realm of nuclear deterrence (who knows what lessons we should draw from the nonoccurrence of a unique event such as nuclear war?).

3. Timely belief change is more likely when decision makers are accountable for bottom-line outcomes and are free to improvise solutions than when decision makers are bound by complex procedural–bureaucratic norms that limit latitude to improvise (Wilson, 1989).

4. Timely belief change is more likely when decision makers—for either dispositional reasons (Kruglanski, 2004) or situational reasons (Tetlock & Lerner, 1999)—display self-critical styles of thinking.

Overconfidence

The overconfidence literature is a methodological minefield. It is difficult, however, to dismiss the entire body of experimental demonstrations of overconfidence as due to regression artifacts (Dawes, 1998; Einhorn & Hogarth 1981; Fischhoff, 1991). And the same can be said for demonstrations of overconfidence in world politics. Tetlock (2005, Chapters 2 and 3) has found that (1) most political experts consistently assign too low subjective probabilities to events that occur and too high probabilities to events that fail to occur; (2) this effect is not just an artifact of regression toward the mean; and (3) this effect is most pronounced among experts with strong ideological commitments and high needs for closure who are making long-term predictions within their domains of expertise. This result fits snugly under an umbrella social cognition framework that portrays overconfidence as pumped up by both the motivation to generate reasons that one-sidedly favor certain predictions (high need for closure plus strong theoretical priors) and the ability to do so (the large knowledge base provided by expertise and unchecked by immediate reality constraints because of the long-term nature of predictions).

In the foreign policy realm, Levy (1989) has noted that overconfidence can lead decision makers to overestimate their ability to detect subtle clues to the other side's intentions and to overassimilate incoming information to their existing beliefs. Overconfident deci-sion makers in defender states are likely to misapply deterrence strategies—either by failing to respond to potential challenges because they are certain that no attack will occur (e.g., Israel in 1973) or by issuing gratuitous threats because they are certain that there will be an attack, even when no attack is planned (Levy, 1989). Overconfident aggressors are prone to exaggerate the likelihood that defenders will yield to challenges (Lebow, 1981). In addition, overconfidence can produce flawed policies when decision makers assess military and economic capabilities. For instance, the mistaken belief that one is militarily superior to a rival may generate risky policies that can lead to costly wars that no one wanted (Levy, 1989). By contrast, a mistaken belief that one is inferior to a rival can exacerbate conflict in either of two ways. First, such beliefs generate unnecessary arms races as the weaker side tries to catch up. The rival perceives this effort as a bid for superiority, matches it, and sets the stage for an action/reaction pattern of conflict spiral. Second, the weaker state will be too quick to yield to a rival's demands (Levy, 1989). At best, such capitulation produces a diplomatic defeat; at worst, it tempts aggressors to up the ante and ultimately produces wars that might have been avoided with firmer initial policies (a widely held view of Chamberlain's appeasement policy of 1938).

The Fundamental Attribution Error

People often gravitate toward dispositional explanations for others' conduct, and overlook plausible situational accounts exist (Gilbert & Malone, 1995; Jones, 1979; Ross, 1977). This proclivity can interact dangerously with the security dilemma at the heart of international relation theory (Jervis, 1976). To protect themselves in an anarchic environment (no world government), states must seek security either through costly defense programs of their own or by entering into entangling alliances that oblige others to defend them. Assessing intentions in such an environment is hard. There is usually no easy way to distinguish between defensive states responding to the competitive logic of the situation and expansionist states seeking to transform the status quo. If everyone assumes the worst, the stage is set for conflict-spiral-driven arms races that no one wanted (Downs, 1991; Kramer, 1988). The fundamental attribution error exacerbates matters by lowering the perceptual threshold for attributing hostile intentions to others. This tendency—in conjunction with the security dilemma—can lead to an inordinate number of "Type I errors" in which decision makers exaggerate the hostile intentions of defensively motivated powers. The security dilemma compels even peaceful states to arm; the fundamental attribution error then leads observers to draw incorrect dispositional inferences from decisions to arm. Both processes—combined with self-serving motivational biases (Heradstveit & Bonham, 1996)—set the stage for a self-righteous spiral of hostility in which policymakers know that they arm for defensive reasons, assume that others know this, and then conclude that those who insist on building up their

military capabilities must have aggressive designs. At best, the result is a lot of unnecessary defense spending; at worst, needless bloodshed (White, 1984).

The fundamental attribution error may encourage a second misperception in the international arena: the tendency to perceive governments as unitary causal agents rather than as complex amalgams of bureaucratic and political subsystems, each pursuing its own missions and goals (Vertzberger, 1990). Retrospective reconstructions of the Cuban missile crisis have revealed numerous junctures at which American and Soviet forces could easily have come into violent contact with each other, not as a result of following some carefully choreographed master plan plotted by top leaders but, rather, as a result of local commanders executing standard operating procedures (Blight, 1990; Sagan & Waltz, 1995). The organizational analog of the fundamental attribution error is insensitivity to the numerous points of slippage between the official policies of collectivities and the policies actually implemented.

Claims about the fundamental attribution error in world politics should, however, be subject to rigorous normative and empirical scrutiny. On the normative side, skeptics can challenge the presumption of "error." Deterrence theorists might note that setting a low threshold for making dispositional attributions can be adaptive. One may make more Type I errors (false alarms of malevolent intent) but fewer Type II errors (missing the threats posed by truly predatory powers). And theorists of international institutions note the value of pressuring states to respect the norms of the transnational trading and security regimes. One way of exerting such pressure is to communicate little tolerance for excuses for norm violations, thereby increasing the reputation costs of such conduct. A balanced appraisal of the fundamental attribution "error" hinges on our estimates of the risk of each type of error as well as on the value we place on avoiding each error—a standard signal detection problem. On the empirical side, skeptics can challenge the presumption of "fundamental" by pointing to collectivist cultures in which sensitivity to contextual constraints on conduct is common (Peng & Nisbett, 1999). Skeptics can also question what counts as a "dispositional" explanation when we shift from an interpersonal to an international level of analysis and the number of causal entities expands exponentially. One could code the domestic political system of one's adversary might be coded either as a situational constraint on leadership policy or as a reflection of the deepest dispositional aspirations of a "people."

Avoidance of Value Trade-offs

For a mutually reinforcing array of cognitive, emotional, and social reasons, politicians find value trade-offs unpleasant and frequently define issues in ways that bypass the need for such judgments (Jervis, 1976; Mintz, 1993; Payne et al., 1992). Trade-off avoidance can, however, be dangerous. Implementing a policy of deterrence, for example, raises tricky trade-offs between the need to resist exploitation and the need to avoid exacerbating the worst-case fears of adversaries. The first value calls for de-

terrence; the second, for reassurance. There are also more complex multipronged geopolitical trade-offs. Kennedy (1987) has noted how great powers over the centuries have mismanaged the trade-offs among defense spending, productive investment, and consumer spending. Although policymakers must allocate resources for defense to deter adversaries and for consumption to satisfy basic needs, too much of either type of spending cuts into the long-term investment required for the sustained economic growth that undergirds military capabilities.

Research suggests that policymakers often avoid trade-offs in a host of ways: (1) holding out hope that a dominant option (one superior on all important values) can be found; (2) resorting to dissonance reduction tactics such as bolstering (Festinger, 1964) and belief-system overkill (Jervis, 1976) that create the illusion that one(s preferred policy is superior on all relevant values to all possible alternatives; (3) buckpassing and procrastinating to diffuse responsibility or delay the day of reckoning (Janis & Mann, 1977; Tetlock & Boettger, 1994); and (4) relying on lexicographic decision rules—such as elimination by aspects (Tversky, 1972)—that initially eliminate options that fail to pass some threshold on the most important value and then screen options on less important values (Mintz, 1993; Payne et al., 1992). Whichever avoidance strategy they adopt, policymakers who fail to acknowledge the trade-off structure of their environment can get into serious trouble by slighting one or another set of affected values.

It would be a mistake though to imply that policymakers are oblivious to trade-offs. Although complex trade-off reasoning is rare in public speeches, policymakers may know more than they let be known. Acknowledging trade-offs can be embarrassing. In addition, some policymakers display an awareness of trade-offs even in public pronouncements. Content analysis of the political rhetoric of Gorbachev revealed much sensitivity to the multifaceted trade-offs that had to be made by the Soviet Union had to make if it were to survive, in Gorbachev's words, into the next century in a manner befitting a great power (Tetlock & Boettger, 1989). Gorbachev's career illustrates that holding a complex view of the trade-off structure of one's environment is no guarantee that one will traverse the terrain successfully.

It would also be a mistake to imply that trade-off avoidance invariably leads to disaster. Procrastination is sometimes prudent. An expanding economy or evolving international scene may eliminate the need for trade-offs that sensible people once considered unavoidable. Passing the buck may be an effective way to diffuse blame for policies that inevitably impose losses on constituencies. And simple lexicographic decision rules may yield decisions in many environments that are almost as good as those yielded by more exhaustive, but also exhausting, utility-maximization algorithms.

Framing Effects

Prospect theory asserts that choice is influenced by "framing" (Kahneman & Tversky, 1979). When we de-

scribe a problem as entailing a high probability of gain, people tend to be risk-averse; when we recast the same problem as entailing a high probability of loss, people tend to be risk seeking. This prediction has been supported in numerous experiments (Bazerman, 2005; Tversky & Kahneman, 1981) as well as in case studies of policy decisions (Farnham, 1992; Levy, 1992; McDermott, 1998).

Framing effects can create severe impediments in negotiations (Bazerman, 2005; Jervis, 1989). When negotiators view their own concessions as losses and concessions by their opponent as gains, the subjective value of the former will greatly outweigh that of the latter. Both sides will therefore perceive a "fair" deal to be one in which the opponent makes many more concessions—hardly conducive to reaching agreements. Reactive devaluation makes matters even worse. When both sides distrust each other, concessions by the other side are often minimized for the simple (not inherently invalid) reason that the other side made them (Maoz, Ward, Katz, & Ross, 2004; Ross & Griffin, 1991). For instance, in 1981, President Reagan unveiled his zero-option proposal calling for Soviet dismantling of intermediate-range missiles (SS-20s) in eastern Europe while the United States would refrain from deploying new missiles in western Europe. The Kremlin initially rejected this proposal, but in 1986, the new Soviet leadership embraced it. Gorbachev's concessions stunned many Western observers, who now assumed that the zero option must favor the Soviets because of their conventional superiority and urged the United States to wiggle out of the potential agreement.

As prospect theory in one or another of its variations (Kahneman & Tversky, 2001) has emerged as the leading alternative to expected utility theory, its influence has proliferated throughout international relations. Levy (1992) notes a host of real-world observations on bargaining, deterrence, and the causes of war that are consistent with the spirit of prospect theory: (1) it is easier to defend the status quo than to defend a recent gain; (2) forcing a party to do something ("compellence") is more difficult than preventing a party from doing something (deterrence); (3) conflict is more likely when a state believes that it will suffer losses if it does not fight; (4) superpower intervention will be more likely if the client state is suffering; and (5) states motivated by fear of loss are especially likely to engage in risky escalation.

Although prospect theory fits these observations, much has been lost in translation from the laboratory literature (in which researchers can manipulate the framing and likelihood of outcomes) to historical accounts of world politics (Boettcher, 1995; Tetlock, 1999a). One critical issue for nontautological applications of prospect theory is "renormalization"—the process of adjusting the reference point after losses or gains. McDermott (1998) speculates that decision makers renormalize more rapidly for gains (recent acquisitions quickly become part of their endowment) than for losses (they may grieve for centuries over their setbacks, nurturing irredentist dreams of reconquest). The pressing need remains, however, for valid research methods—such as content analysis of group discussions (Levi & Whyte, 1997)—for determin-

ing when specific actors are in a loss or gain frame of mind. Without such methods, there is a danger of retrospective data fitting, *post hoc ergo propter hoc* reasoning: Because decision makers behaved rashly, we know they must have been in the losses.

To generate predictions from prospect theory (expansively interpreted) about whether decision makers will pursue risky or cautious policies, we need to know what is presently unknowable. For example, if decision makers have recently suffered losses, have they made peace with those losses (incorporated them into new reference points), or do they nurture grudges and dreams of dramatic recovery—as, for example, Hitler did with respect to the Treaty of Versailles? All other things equal, the latter, a loss frame of mind, should be more conducive to risk taking than the former. Conversely, if decision makers have recently enjoyed a string of triumphs, have they incorporated these new acquisitions into their reference point endowment of power, wealth, and status or has the novelty of fresh acquisitions yet to wear off? All other things being equal, the latter "gain" frame of mind, should be much more conducive to caution. But all other things are never equal in the real world. Moreover, much also hinges on decision makers' perceptions of the likelihood of the consequences of alternative courses of action. The probability-weighting function of prospect theory suggests that decision makers will often insure themselves against low-probability losses (risk aversion) and gamble for low-probability gains (risk seeking).

Another critical issue concerns how the risk preferences of individuals are amplified or attenuated by group processes such as diffusion of responsibility, persuasive arguments, cultural norms, and political competition for power (Druckman, 2004; Vertzberger, 1995). This argument reminds us of the need to be vigilant to the ever-shifting dimensions of value on which people may be risk averse or risk seeking. Leaders may suddenly become recklessly obdurate when issues of honor and identity are at stake. Saddam Hussein, just prior to the annihilation of his army in Kuwait, reportedly invoked the old Arab aphorism that "it is better to be a rooster for a day than a chicken for all eternity" (Post, 1991).

Do Psychologists Have Pro-Bias?

Although researchers have emphasized the role of cognitive biases in creating conflicts that might have been avoided had decision makers sized up situations more accurately, it is worth noting that these biases can attenuate as well as exacerbate conflicts. Much depends on the geopolitical circumstances. The fundamental attribution error can alert us to predatory powers; simplistic analogies are sometimes apt; belief perseverance can stop us from abandoning veridical assessments in response to skillful "disinformation" campaigns; and high-risk policies can yield big payoffs. Indeed, efforts to eliminate these "biases" through institutional checks and balances are likely to be resisted by skeptics who argue that such cognitive tendencies are often functional. Consider overconfidence. Some psychologists have made a strong case that when this "bias" takes the form of infectious "can-

do" optimism, it promotes occupational success and mental health (Seligman, 1990; Taylor & Brown, 1988). Moreover, there is the flip-side risk that self-critical thinkers may dilute justifiably confident judgments by heeding irrelevant or specious arguments (Tetlock & Boettger, 1989)—a severe threat to good judgment in environments in which the signal-to-noise ratio is unfavorable and other parties are trying to confuse or deceive the perceiver.

Such arguments strike resonant chords among those in the policy community who stress the dangers of "analysis paralysis" (Tetlock, 2000). Advice that strikes some academic observers as obviously sound strikes some policy elites as obviously flawed. Decision analysts face an uphill battle in convincing skeptics that the benefits of their prescriptions outweigh the costs. Whether or not they acknowledge it, policymakers must decide how to decide (Payne et al., 1992) by balancing the estimated benefits of complex, self-critical analysis against the psychological and political costs. What increments in decision quality is it reasonable to expect from seeking out additional evidence? Some observers see enormous potential improvement (e.g., Herek, Janis, & Huth, 1987; Janis, 1989); others suspect that policymakers are already shrewd cognitive managers skilled at identifying when they have reached the point of diminishing analytical returns (e.g., Suedfeld, 1992). These strong conclusions rest, however, on weak evidentiary foundations. In the spirit of Robert Merton's (1987) observation that the most rapidly advancing sciences are those that are most explicitly specify their domains of ignorance, we are well advised to admit both how poorly experts predict real-world trends and how little we know about the correlates of what systematic variation there is in judgmental accuracy (Tetlock, 2005).

Motivated Biases and Errors

Assumptions of rationality can be challenged not only on "cold" cognitive grounds but also on "hot" motivational grounds. Examples of such challenges arise in the literature on crises (high-pressure settings in which decision making may be driven by the ebb and flow of human emotions), on escalating commitment to sunk costs, and on the impact of deep-rooted personality needs.

Crises

Policymakers rarely have a lot of time to think. They frequently work under stressful conditions in which they must process large amounts of inconsistent information under severe time pressure, always with the knowledge that miscalculations may have serious consequences for both their own careers and vital national interests (Holsti, 1989). This combination of an imperative demand for crucial decisions to be made quickly, with massive information overload, is a form of psychological stress likely to reduce the information-processing capacity of the individuals involved (Suedfeld & Tetlock, 1977).

Both experimental studies and content-analytic studies of archives yield support for this hypothesis. The laboratory literature has repeatedly shown that stress—beyond the inflection point on the inverted-U curve—impairs complex information processing (Streufert & Streufert, 1978; Svenson & Maule, 1994). Impairment can take many forms, including a lower likelihood of discriminating among unfamiliar stimuli, a higher likelihood of relying on simple heuristics, greater reliance on old, now inappropriate, problem-solving strategies, and reduced curiosity in new, potentially dissonant evidence (Janis & Mann, 1977; Staw, Sandelands, & Dutton, 1981).

Archival studies reinforce these pessimistic conclusions, most notably, the work of Suedfeld and colleagues on declining integrative complexity in response to international tension (Koh, Han, & Kim, 2004; Maoz & Shayer, 1987; Raphael, 1982; Suedfeld & Tetlock, 1977). These downward shifts are especially pronounced in crises that culminate in war. It is tempting here to tell a causal story in which crisis-induced stress impairs the capacity to identify viable complex compromises, thereby contributing to the violent outcome. But it is wise to resist temptation until two issues are resolved. First, falling integrative complexity may be a sign not of simplification of mental representations but rather of a self-conscious hardening of bargaining positions. Policymakers may decide to lower their integrative complexity (closing loopholes, eliminating qualifications, denying trade-offs, and disengaging from empathic role taking) as a means of communicating firmness of resolve to adversaries (Tetlock, 1985). Here, we need more studies that trace shifts in integrative complexity in both private and public documents (Guttieri, Wallace, & Suedfeld, 1995; Levi & Tetlock, 1980; Walker & Watson, 1994). Second, there are exceptions to the generalization that high stress produces cognitive simplification. Decision makers have sometimes risen to the challenge and responded to intensely stressful circumstances in thoughtful ways (Brecher, 1993). For instance, during both the Entebbe crisis (Maoz, 1981) and the Six-Day War of 1967 (Stein & Tanter, 1980), Israeli officials managed—despite great stress—to consider numerous options, assess consequences in a probabilistic manner, trade off values, and remain open to dissonant information.

The challenge is identify when crisis-induced stress triggers simplification and rigidification. One approach is to look for quantitative variables that influence where the hypothetical optimum for the effects of stress lies (Streufert & Streufert, 1978). Another approach is to look for qualitative moderator variables that activate simple or complex coping strategies. For instance, the Janis and Mann (1977) conflict model predicts simplification of thought (defensive avoidance) only when decision makers confront a dilemma in which they must choose between two equally unpleasant alternatives and are pessimistic about finding a more palatable alternative in the time available. Under these conditions, decision makers are predicted to choose and bolster one of the options, focusing on its strengths and the other option's weaknesses (thereby spreading the alternatives). By contrast, when decision makers are more optimistic about finding

an acceptable solution in the available time but still perceive serious risks, they will shift into vigilant modes of information processing in which they balance conflicting risks in a reasonably thoughtful way.

Postdecisional Bolstering

Another motivated bias has been identified by research on escalating commitment to sunk costs. Experimental work suggests (Brockner & Rubin, 1985; Staw, 1980; Tetlock, 1992), and case studies corroborate, that policymakers who feel accountable for hard-to-reverse decisions often concentrate more on justifying what they have done than on figuring out what needs to be done next. These exercises in retrospective rationality—whether viewed as dissonance reduction or impression management—tend to be especially intense to the degree that earlier decisions cast doubt on decision makers' integrity or ability.

Situations of this sort—military quagmires such as Vietnams and Afghanistans or financial quagmires such as unpromising World Bank projects with large sunk costs—are common in political life. Indeed, the primary job of opposition parties in democracies is to find fault with the government and to refute the justifications and excuses that the government offers in its defense. The psychodrama becomes politically consequential when it extends beyond verbal sparring at press conferences and begins to bias policy appraisals. After all, if one convinces oneself and others that a bad decision worked out well, it starts to seem reasonable to channel even more resources into the same cause. Such sincerity can be deadly when decision makers must choose among courses of action in international confrontations under the watchful eyes of judgmental domestic constituencies.

Personality Needs

Lasswell (1948) famously defined political behavior as a function of private motives displaced onto public objects and rationalized in terms of the common good. By mixing laboratory and naturalistic methods, researchers have built a rather convincing case for this proposition (Etheredge, 1980; Greenstein, 1975; Herrmann, 1988; Walker, 1983; Winter, 1993).

In their analyses of disagreements among makers of American foreign policy elites between 1898 and 1984 over the use of force, Etheredge (1980) and Shephard (1988) found that policy preferences were closely linked to personality variables. Working from biographical data on personal relationships, coders rated leaders on interpersonal dominance (strong need to have their way and tendency to respond angrily when thwarted) and extroversion (strong need to be in the company of others). As predicted, dominant leaders resorted more to force than their less dominant colleagues and extroverted leaders advocated more conciliatory policies than their more introverted colleagues. Laboratory work suggests causal pathways for these effects, including perceptual mediators (dominant people see high-pressure tactics as more efficacious) and motivational mediators (dominant peo-

ple try to maximize their relative gains over others, like neorealists, whereas less dominant people try to maximize either absolute gains, like neoclassical economists, or joint gains, like good team players) (see Brewer & Kramer, 1985; Sternberg & Soriano, 1984).

In a cumulative series of studies, Winter (1993) adapted the content-analysis systems for assessing motivational imagery in the semiprojective Thematic Apperception Test (TAT) to analyze the private and public statements of world leaders. He has also proposed a psychodynamic conflict-spiral model that posits that combinations of high power motivation and low affiliation motivation encourage resort to force in international relations. Winter tested this prediction against archival materials drawn from several historical crises. In each case, Winter observed the predicted correlations between motives and war versus peace. In another study, Peterson, Winter, and Doty (1994) linked their motive theory to processes of misperception hypothesized to occur in conflict spirals (cf. Kelman & Bloom, 1973). In this integrative model, international conflicts escalate to violence when three conditions are satisfied: (1) there is high power motivation in the leadership of both countries; (2) each side exaggerates the power imagery in messages from the other side; and (3) each side expresses more power motivation in response to its exaggerated perceptions of the other side's power motivation. In ingenious simulations, Peterson and colleagues used letters exchanged in an actual crisis as stimulus materials and showed that subjects with high power motivation were especially likely to see power motivation in communications from the other side and to recommend the use of force.

The motive-imagery explanation is parsimonious. The same content-analytic method yields similar relationships across experimental and archival settings. But the interpretive difficulty is the same as in integrative complexity research: the possibility of spurious multimethod convergence. Political statements cannot be taken as face-value reflections of intrapsychic processes. Leaders use such statements both to express what they think and to influence what others think they think. There are good reasons for supposing that people can strategically raise and lower their integrative complexity (Tetlock, 1981), and it would be astonishing if motivational imagery were not also responsive to shifts in impression management goals.

Group-Induced Biases and Errors

Accountability to colleagues is often trotted out as an all-purpose solution to decision errors; accountability can, however, become intellectually incestuous when policymakers expect to answer only to like-minded colleagues and constituencies. This concentration of accountability to an ingroup is a defining feature of groupthink (Janis, 1982). The combination of opinionated leadership, insulation from external critics, and intolerance of dissent often appears sufficient to amplify already dangerous tendencies in individual judgment. Groupthink decision makers are more prone to

jump to premature conclusions, to dismiss contradictory evidence, to deny trade-offs, to bolster preferred options, to suppress dissent within the group, and to display excessive optimism.

Janis's case studies are the best-known effort to apply work on group dynamics to elite political settings. The groupthink model does, however, have serious limitations (t'Hart, Stern, & Sundelius, 1995). First, the evidence—from case studies to experiments—is mixed. Close inspection of case studies underscores the ambiguity of many diagnoses of (groupthink(in foreign policy contexts. For example, comparing Berman's (1982) and Janis's (1982) accounts of Johnson's decision to intervene in Vietnam, one needs to be a mindreader to determine whether (1) a manipulative Johnson had made up his mind in advance and used group deliberations merely to justify a predetermined policy; or (2) an uncertain Johnson leaned heavily on a cliquish advisory group for cognitive and emotional support. Also mixed is the content-analytic and Q-sort evidence (Tetlock, 1979; Tetlock, Peterson, McGuire, Feld, & Chang, 1992; Walker & Watson, 1994). These studies support some aspects of the model (rigidity and self-righteousness in hypothesized cases of groupthink) but not others (there is little evidence that cohesiveness alone or in interaction with other antecedents contributes to defective decision making). And laboratory studies have been even less supportive of the hypothesized preconditions for defective decision making (Aldag & Fuller, 1993; Turner, Pratkanis, Probasco, & Love, 1992)—although defenders of the model can always invoke the external-validity argument that experimental simulations pale next to their dramatic real-life counterparts.

Second, the groupthink model oversimplifies process-outcome linkages (Tetlock, Peterson, McGuire, Feld, & Chang, 1992; t'Hart, 1994). It is easy to identify cases in which concurrence seeking has been associated with outcomes that most observers now applaud (e.g., Churchill's suppression of dissent in cabinet meetings in 1940–1941, when some ministers flirted with the now unthinkable: a negotiated peace with Hitler) and cases in which vigilant decision making has been associated with outcomes that left group members bitterly disappointed (e.g., Carter encouraged rather vigorous debate over the wisdom of the hostage-rescue mission in Iran in 1980, even though the Secretary of State was eventually shunted out of the loop). The correlation between quality of process and of outcome was perfectly positive in Janis's (1982) carefully chosen case studies and positive correlations can be engineered to hold lab settings (Postmes, Spears, & Cihangir, 2001), but it is likely to slip in representative samplings of naturally occurring decision-making episodes (Bovens & t'Hart, 1996). We need contingency theories that identify (1) the distinctive patterns of group decision making that lead, under specified circumstances, to political success or failure (t'Hart et al., 1995; Vertzberger, 1995); and (2) the diverse functions that leadership groups serve. Groups do not just exist to solve external problems; they provide symbolic arenas in which bureaucratic and political conflicts can be expressed, support for shared values reaffirmed, and potentially divisive trade-offs concealed

(t'Hart et al., 1995). As with cognitive biases (Tetlock, 2002), patterns of group decision making judged maladaptive within a functionalist framework that stresses analytic problem solving appear reasonable within frameworks that stress other imperatives, such as acting decisively, forging a united front and mobilizing external support. At the risk of sounding repetitiously relativistic, much hinges on the value perspective of the scientific observer. Does the observer deem these alternative functions "good reasons"?

ATTRIBUTIONS OF EXCESSIVE COMPETITIVENESS, EVEN BELLICOSITY

The reductionist syllogism here rests on the first-order principle that laboratory and field studies have already taught us a great deal about the factors that promote or impede successful negotiation and conflict resolution at the interpersonal level (Bazerman, 2005; Kleiboer, 1996; Pruitt, 1998; Pruitt & Rubin, 1986; Thompson, 1990). The minor premise asserts process invariance: the same basic psychological processes at work in the interpersonal sphere drive international relations. It follows that explanatory frameworks developed at the microlevel can be readily applied at the macrolevel.

Many international relations scholars are not convinced. The most influential theory within this discipline remains some variant of deterrence theory (Schelling, 1966). Although deterrence theory comes in many forms (from thoughtful prose to game-theoretic models), it rests on a clear set of alternative first-order principles—principles for which prominent political psychologists have not concealed their disdain (Deutsch, 1983; White, 1984), even characterizing them as paranoid (Mack, 1985). First, the world is posited to be a dangerous place. Opportunistic opponents will seize every chance to expand their influence at one's expense. Whenever the expected utility of defecting or attacking exceeds that of other options, the risk of defection or aggression rises to an unacceptably high level. It follows that to deter defection or aggression—be it in a trade or military dispute—one must issue retaliatory threats that convince one's opponent that the expected utility of defection or aggression is lower than that of the status quo. Deterrent threats must therefore be potent and credible. Would-be aggressors must believe that the defender possesses the resolve and capability to implement the threat.

Deterrence theorists accept these principles in the abstract but disagree vigorously over how to implement them in policy. During the Cold War, for instance, some argued that in a MAD world, nuclear weapons could deter attacks only on one's territory whereas others argued that nuclear threats could also deter attacks on allies. More recently, deterrence theorists have split over the wisdom of launching preemptive attacks on countries suspected of providing havens for terrorists or developing weapons of mass destruction.

Any serious evaluation of deterrence theory must grapple with the deep problems of determining from the historical record whether deterrence actually worked

(Danilovic, 2004). To be sure, dramatic failures of deterrence as a policy are easy to identify. One country sought to prevent an attack on another, but the attack occurred nonetheless. The historical data are, however, sufficiently ambiguous to allow endless arguments over whether individual cases also represent failures of deterrence theory (as opposed to failures of implementation). An equally imposing obstacle is presented by alleged deterrence successes. When crises do not occur, is it because of the credibility of threats (successes for deterrence theory) or because other states never intended to attack in the first place? Causal inference requires assumptions about what would have happened in the missing counterfactuals cells in the contingency table in which the defender issued no threats. Research on counterfactual reasoning shows that the dominant tendency among sophisticated observers is to plug in ideological assumptions about what would have happened into the missing data cells. Moreover, this tendency is especially pronounced among observers who have high needs for closure (Tetlock, 2005) and a preference for integratively simple reasoning (Suedfeld & Tetlock, 2001).

These issues are of more than idle intellectual interest. The same policy elites who disagreed over the role that strong deterrence threats played in ending the Cold War also disagreed over the wisdom of dislodging Iraq from Kuwait in 1991 and of invading Iraq in 2003 and of using military threats to induce North Korea to abandon its nuclear weapons program. The hawk–dove personality distinction is fundamental in all debates over the use of force (Etheredge, 1980; Herrmann, Tetlock, & Visser, 1999).

One way out of this impasse is to turn to the laboratory literature on bargaining and negotiation for guidance on the relative utility of threats versus other influence tactics. This literature suggests that deterrence theory needs to be seriously qualified. For instance, laboratory studies of bargaining showed long ago that threats can impede cooperation and lower joint outcomes by triggering anger, competitiveness, and reactance (Deutsch, 1983; Kelley, 1965). And work on negotiation processes suggests a rather pervasive tendency for people, especially those in competitive mindsets or under accountability pressure from key constituencies, to exaggerate the zero-sum character of disputes (the fixed-pie fallacy; Bazerman, 2005; Thompson, 1990) and to overlook positive-sum integrative solutions (Pruitt, 1998).

The obvious alternative to threats of punishment is promises of reward. But, since Munich gave appeasement a bad name, scholars have slighted the role of positive inducements. The most outspoken proponents of positive inducements have been conflict-spiral theorists who warn of escalating action–reaction cycles in international conflict (Deutsch, 1983; White, 1984). And even these theorists have not advocated unilateral disarmament—and for good reason: Experimental evidence indicates that in mixed-motive games, such as Prisoner's Dilemma, unconditional cooperation is ruthlessly exploited (McClintok, Stech, & Beggan, 1987). Archival work on international disputes points to similar conclu-

sion: Nations adopting an appeasement strategy managed to avoid war but almost always suffered a diplomatic defeat (Leng, 1993). Positive inducements such as financial rewards for compliance can also be very expensive if the other side complies (particularly if it quickly becomes satiated and increases its demands for compensation) and they can foster unwanted dependency and a sense of entitlement. Finally, just as deterrence theorists face difficulties in operationalizing threats, so "reward" theorists encounter problems in operationalizing positive inducements, which may be perceived as condescending, presumptuous, manipulative, or insultingly small or large (Patchen, 1987).

The most attractive options appear to be mixed-influence strategies. Inspired by Robert Axelrod's (1984) "the evolution of cooperation," many researchers have focused on firm but fair strategies of resolving conflicts. The most famous is tit-for-tat. One begins by cooperating and thereafter repeats one's opponent's previous move. Although there are boundary conditions on the efficacy of tit-for-tat (e.g., misperception of intent behind moves can wreak havoc; Bendor, 2001), considerable research demonstrates that tit-for-tat is as effective as it is simple. In Axelrod's round-robin Prisoner's Dilemma computer tournaments that pitted expert nominated strategies against one another, tit-for-tat—the simplest entrant— earned the highest average number of points. Axelrod argued that tit-for-tat works because it is nice (never defects first), perceptive (quickly discerns the intent of the other), clear (easy for the other side to recognize), easily provoked (quickly retaliates), forgiving (willing to abandon defection immediately after the other side's first cooperative act), and patient (willing to persevere). Numerous archival studies have also shown that strategies of a tit-for-tat character tend to be more effective than either pure threat or pure appeasement strategies in averting both war and diplomatic defeat (Leng, 1993).

One could also argue that one of psychology's most ambitious contributions to the policy debate, Charles Osgood's graduated and reciprocated initiatives in tension reduction (GRIT) is much in the spirit of tit-for-tat (Lindskold, 1978; Patchen, 1987). Both are designed to resist exploitation and to shift the interaction onto a mutually beneficial, cooperative plane. Unlike tit-for-tat, however, GRIT does not assume that the game has yet to begin. Rather, GRIT assumes that the parties are already trapped in a costly conflict spiral. To unwind the spiral, Osgood proposed that one side should announce its intention to reduce tensions and then back up its talk with unilateral conciliatory gestures. These actions are designed to convince the opponent of the initiator's peaceful intentions but not to weaken the military or economic position of the initiator. The opponent is then invited to respond with conciliatory gestures but warned that attempts to exploit the situation will force the initiator to return to a hardline posture. GRIT is a softer, more liberal brand of tit-for-tat: It is nicer (it cooperates even in the face of defection) and less easily provoked (it continues to cooperate even when the opponent ignores what one has done. A good deal of evidence suggests that GRIT stimulates cooperation. The most cumulative ex-

perimental program—that of Lindskold (1978)—has repeatedly shown that GRIT leads to more integrative agreements than do competitive and no-message strategies and, moreover, that GRIT elicits more cooperation when initiated from a position of strength than from one of weakness (a finding that could be invoked as support for a mixed hawk–dove strategy of defense buildup as prelude to GRIT).

In sum, any policy-relevant conclusions we draw should be complex and tentative ones. An exclusive deterrence-based emphasis on threats can trigger otherwise avoidable conflicts. But so too can calls for unilateral disarmament, albeit via a different mechanism—by tempting aggressors. Encouraging, though, is the multimethod convergence suggesting that in many situations variants of a firm but fair strategy does a good job of both protecting vital interests and preventing conflicts from spiraling out of control. The challenge is translating this abstract causal principle into normative prescriptions in specific conflicts. There will always be a big element of guesswork in figuring out the optimal policy blend for dealing with specific régimes, and this will be especially true in dealing with régimes that we find profoundly alien in their practices and values (witness the old debates over Cold War strategy (Tetlock, McGuire, & Mitchell, 1991) as well as more recent debates over dealing with Islamic suicide attacks or with the North Korean nuclear program).

ATTRIBUTIONS OF SIMPLE-MINDEDNESS, DOGMATISM, AND RIGIDITY

Since the pioneering work on the authoritarian personality (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950), psychologists have taken an active interest in treating political beliefs as dependent variables. The reductionist syllogism here rests on the first-order principle that all attitudes serve shifting blends of basic psychological functions: appraising reality, claiming social identities, preserving self-esteem motives, and expressing core values (Smith, Bruner, & White, 1956). The minor premise is that political attitudes are, by definition, attitudes. The conclusion is that functionalist psychological theories can help us to explain political attitudes.

The most influential functionalist hypotheses have posited that conservative beliefs tend to be especially rigid and dismissive of counterarguments because such beliefs help holders to (1) cope with unresolved childhood conflicts by identifying with esteemed authority figures; (2) suppress complexity and ambiguity by squeezing reality into procrustean categorization schemes; and (3) adjust to the intolerant norms of their community (cf. Altemeyer, 1996; Lane, 1973).

For over 50 years now, researchers have explored these conjectured linkages between cognitive styles and political ideology, between how people think and what people believe. This work has used a wide range of methods—ranging from self-report scales to content analyses of interviews and speeches—and has explored a wide range of samples—from college undergraduates to politi-cal activists to British parliamentarians and U.S. senators (Putnam, 1988; Stone, 1980; Tetlock, 1984).

A recent meta-analysis of this literature has sparked controversy by concluding that right-wing conservatism is related to dogmatism and intolerance of ambiguity; uncertainty avoidance; fear of threat, loss, and death; system instability; and epistemic needs to achieve order, structure, and closure as well as negatively related to openness to experience, integrative complexity, and (to a lesser extent) self-esteem (Jost, Glaser, Kruglanski, & Sulloway, 2003a, 2003b). The authors do, though, add a critical caveat: This does not mean that liberals crave uncertainty and risk. There can be rigidity of the left as well as of the right. Their claim is simply that rigidity of the left is markedly less common than that on the right.

Greenberg and Jonas (2003) challenge the meta-analysis on a mixture of empirical and conceptual grounds. They see far more exceptions to the rigidity-of-the-right hypothesis than do Jost and colleagues (2003a). For instance, Tetlock (1984)—drawing on Putnam's (1973) dataset on British parliamentarians—found support for the ideologue hypothesis: Hard-core socialists within the British Labor Party were as prone to dichotomous reasoning as hard-core conservatives. Greenberg and Jonas also point to Tetlock and Boettger (1989), who found, by content-analyzing statements by Soviet Politburo members, that more orthodox communists were less integratively complex than the reformist, pro-Gorbachev, Politburo members. These findings underscore the definitional ambiguity shrouding the concept of conservatism: Do hard-core "lefties" in the British Labor Party count as radicals or "conservative" upholders of the Party tradition? And do hard-line Soviet Communists in the late 1980s count as Bolshevik radicals or as conservative restorationists?

If we opt for the latter interpretations, it is easy to assimilate the Parliamentary and Politburo data to the Jost and colleagues (2003a) preference for the rigidity-of-the-right hypothesis (Tetlock, 1984). But Greenberg and Jonas also challenge definitions of conservatism that emphasize commitment to upholding the status quo. When the status quo is welfare-state liberalism or social democracy, conservatives can plausibly be portrayed—and sometimes are by adversaries—as radicals who want to change basic institutions: advocating the abolition of affirmative action, widespread privatization of state-controlled sectors of the economy, deregulation of business, and implementing dramatic cuts in taxation. Jost and colleagues (2003b) reply that these conservative positions should be viewed as reactionary efforts to preserve long-standing patterns of group dominance. This controversy also underscores the definitional murkiness of the ideology/cognitive style debate. The Mertonian injunction to specify scientific ignorance requires conceding that neither Jost and colleagues (2003a) nor Greenberg and Jonas (2003) have evidentiary warrant for claiming to know why "conservatives"—far from a homogeneous group—take the policy positions they do.

In a contextualist spirit, Tetlock and Mitchell (1993) proposed a taxonomy of flattering and unflattering cognitive and motivational portraits of liberalism and con-

servatism—each almost certainly true at particular times and places. Jost and colleagues (2003) favor one of the less flattering motivational possibilities: the desire to preserve the existing hierarchy of group dominance. Greenberg and Jonas (2003) flirt with the more flattering possibility that conservatives believe that the long-term well-being of society could be promoted by market mechanisms that stimulate economic growth and technological innovation.

These possibilities are, however, neither exclusive nor exhaustive. One could be a committed opponent of wealth-redistribution schemes that most liberals feel are essential for achieving equality of opportunity and a committed proponent of free-market economics—and view the two positions as mutually reinforcing. It is also worth noting that either of the two belief sets could play the role of independent or dependent variable vis-à-vis the other. And one could come to a conservative worldview via a variety of other cognitive-emotional paths, including—among other possibilities listed in Tetlock and Mitchell (1993)—a prospect-theory-grounded aversion to takings (imposing losses on those who are better off and transferring income to those worse off; Mitchell et al., 2003), a belief in the debilitating effects of generous welfare and entitlement programs, and a pessimistic view of human nature that leads one to suspect the readiness of others to exploit weakness and loopholes. Finally, there is evidence—certainly at the elite level and possibly at the level of the mass public as well—that conservatism is a multidimensional construct and that, at minimum, we need to distinguish more authoritarian from more libertarian conservatives (Kinder, 1998).

There are, however, potential methodological paths out of this scientific morass. One route is via hypothetical society paradigms that, artificial though they are, bring some causal clarity (Mitchell, Tetlock, Mellers, & Ordonez, 1993; Mitchell et al., 2003). In the real political world, there is a hopeless confounding of conservative beliefs about how meritocratic society is (answers to "why are the poor so poor and the rich so rich?") and conservative resistance to redistribution of wealth. It could be that conservative beliefs about meritocracy are merely justifications for maintaining existing hierarchies. Or, it could be that conservatives endorse existing patterns of group dominance because they honestly believe that society operates in a reasonably meritocratic fashion. Mitchell and colleagues (2003) cannot cut this Gordian Knot but they can show that when people are experimentally led to believe that socioeconomic status is determined purely by chance (zero meritocracy) or determined entirely by effort and ability (100% meritocracy), liberals and conservatives respond remarkably similarly to proposals to redistribute wealth. Both tend to support redistribution in the zero-meritocracy society and oppose it in the perfect meritocratic society. It was only in the ambiguous 50/50 world, where socioeconomic outcomes are partly determined by chance and partly by personal qualities, that pronounced ideological polarization emerged, with conservatives more opposed to redistri-

bution and liberals more supportive. Beliefs about meritocracy thus have some independent causal force.

In addition to clarifying causal relations among content components of belief systems, experiments can explore cognitive content/cognitive style linkages more effectively than the traditional reliance on psychometric approaches. On the one hand, researchers have shown that they can—via manipulations such as sociotropic threat and mortality salience—influence the conservatism of the policy preferences that people endorse (in particular, the punitiveness of their responses to norm violations (Gibson, 2006; Greenberg, Solomon, & Pyszczynski, 1991; Tetlock et al., in press). On the other hand, it is also possible to manipulate the flexibility and complexity of the reasoning that underlies policy preferences. This can be done, for example, by varying the degree to which issues bring core values into conflict, thus shifting the cognitive-dissonance-predicted preference of people for low-effort strategies of resolving intrapsychic conflict (such as denial and bolstering) toward more effort-demanding strategies (such as differentiation and integration) (Abelson, 1959; Tetlock, 1986). Neither the explanans (cognitive style) nor the explanandum (ideology) in the dispute between Jost and colleagues (2003a, 2003b) and Greenberg and Jonas (2003) is a static trait entity. Each has a dynamic component that permits it to vary across contexts, issues, and time. There are settings in which the linkages between cognitive style and ideology are powerful and in the direction posited by Jost and colleagues (e.g., McClosky, 1967; Tetlock, 1983) and other settings in which the relationships vanish (Gruenfeld, 1995; Tetlock, 1986, 2005). Much hinges on the issues under examination (some value-conflict problems "pull" more complex thinking from liberals; others "pull" more complex thinking from conservatives) (Tetlock, Peterson, & Lerner, 1996), the political balance of power (liberals cease to be more integratively complex than conservatives when liberals lose majority control of Congress) (Tetlock, Hannum, & Micheletti, 1984), and the range of views within the sample (if one goes far enough to the left, one finds a fierce contempt for alternative views that rivals anything found on the right) (Tetlock, 2005).

All that said, it is noteworthy that although the main effect revealed in the Jost and colleagues (2003a, 2003b). meta-analysis can be neutralized, it never seems to reverse direction so that, across issues, conservatives as a group score as more flexible and multidimensional than liberals. But should this result be disturbing to conservatives? Much hinges on the value spin that scientific observers put on the facts: the correlations between our measures of cognitive styles and of political orientations.

There is a common presumption in academic circles that complexity of thought is inherently superior—both as a means to ends and as an end in itself—to simplicity. Many conservatives, however, take issue with this value-laden premise: They subscribe to a different set of epistemological and moral ideals from those upheld by liberals. Tetlock (2000) found that conservative and libertarian business executives were more likely than their lib-

eral counterparts to rise to the defense of judgmental biases such as overconfidence (essential for getting things done and avoiding the analysis–paralysis trap), the fundamental attribution error (essential for communicating to employees that one has limited tolerance for situational justifications and excuses), and the severity effect (essential for motivating people to try to exert more control over the possible accidental consequences of acts of negligence; Tetlock et al., in press). Indeed, there are also settings in which liberal observers favor the integratively simple side of historical debates over the integratively complex side. For instance, Tetlock and Tyler (1996) demonstrated that Churchill used substantially less complex rhetoric to justify his antiappeasement policy stance toward Nazi Germany than did Chamberlain to justify his policy of identifying an integrative, mutually beneficial set of understandings with Hitler (although Churchill quickly loses liberal sympathy when the debate shifts to self-government for India and Churchill takes colonialist positions widely deplored as reactionary, even racist, today). Tetlock, Armor, and Peterson (1994) showed that, when one examines patterns of argumentation in antebellum America (1850s), one finds the least tolerance for trade-offs and willingness to acknowledge the legitimacy of alternative perspectives on both the extreme left (abolitionists) and right (Southern secessionists). The extreme left of that time and place are now widely portrayed as principled, not dogmatic. More generally, Tetlock (1986; Tetlock et al., 1996) showed that it is dangerous to make sweeping generalizations about immutable tendencies of one or another group to think in a particular fashion. Ideology-by-issue interactions in flexible trade-off reasoning are quite common. Much hinges on the degree to which an issue domain activates countervailing values. The value pluralism model, for instance, predicts that integrative complexity peaks when a policy problem activates two values that are approximately equally important, that point in opposite policy directions, and when the social context closes the door to buck passing and procrastination.

In this author's view, political psychology needs more disagreements of the sort represented in the exchange between Jost and colleagues and Greenberg and Jonas. The net result was clarification. Jost and colleagues (2003) close on a fittingly undogmatic note:

> To be clear, we never argued that it is intrinsically good to be tolerant of uncertainty or ambiguity, low on the need for cognitive closure, or even high in cognitive complexity. In many cases, including mass politics, "liberal" traits may be liabilities and being intolerant of ambiguity, high on the need for closure, or low in cognitive complexity might be associated with such generally valued characteristics as personal commitment and unwavering loyalty. (p. 374)

CONCLUDING THOUGHTS

It is a mistake to dismiss political psychology as just another applied subfield in which we can safely assume that the first-order, cognitive and affective principles developed elsewhere in this volume apply with no need for modification. But it is also a mistake to conclude that political psychology is lawless, that its scare-quote-bracketed "truths" are totally contingent upon the quirky cultural–historical contexts in which we find ourselves ideologically embedded. There are principles of political psychology. It is just that those principles are inherently more complex than those in mainstream social psychology. Political psychological principles require taking into simultaneous account the social psychology of the research participants and the political priorities and labeling rules of the researchers. Let us revisit each of the four lines of work featured in this chapter:

1. Prejudice toward outgroup members is a widespread phenomenon for an assortment of basic process, cognitive and affective reasons (Brewer & Brown, 1998). But the decision to label an attitude, implicit or explicit, a manifestation of prejudice is a politically charged one that reflects the prevailing norms and values of the research community. And that decision, in turn, is a lawful phenomenon, driven by a mixture of factual beliefs and value priorities—and subject to systematic modeling that uses conceptual tools such as signal detection and decision theory. How do researchers judge the risk of under- versus overestimating the tenacity of prejudice in shifting political environments? How do they weigh the clashing risks?

2. Cognitive biases—deviations from academic norms of rationality—reach into many spheres of life. But the decision of a researcher to label a point of view biased is a politically charged one that reflects value-laden assumptions about the goals that political actors should be pursuing in particular situations as well as extrapsychological assumptions about the world with which they are struggling to cope. And that decision, in turn, merits political–psychological analysis in its own right. For instance, how do researchers go about judging the relative risks of overconfidence versus underconfidence, of persisting too long with failing policies or prematurely abandoning fundamentally sound policies at the first sign of trouble, and so on?

3. Negotiation blunders are not unusual (Bazerman, 2005) and advocates of the fixed-pie fallacy argument may well be right that the more common error is to be excessively suspicious and competitive and to fail to ferret out integrative agreements (Thompson, 2005). But negotiators can also make the opposite mistake of trusting the untrustworthy. Claims about errors in international negotiation strategies rest on speculative leaps of faith about the researcher's superior ability to read the true intentions of other players. Is it not scientific hubris—even 20 years later—to claim to know whether Kelman (1983) offered a correct political–psychological assessment of Yasser Arafat as a negotiating partner, or whether White (1984) had an accurate read on the mindset of the Soviet Politburo?

4. Conservatives do seem to have different cognitive styles from liberals. But we know, from classic work on

implicit personality theory (Peabody, 1967), that it is easy to reverse the positive or negative connotative spins on the opposing ends of virtually any denotative personality dimension. Risk seekers may be bold or rash; risk avoiders may be cautious or timid. The big question thus remains: Under what conditions should the research community decide that the greater conservative affinity for simpler worldviews, with clear-cut good and bad guys, is a cognitive liability (a sign of rigidity, dogmatism, ignorance, ethnocentrism, etc.) or a cognitive strength (a sign of principled conviction, determination, resolve, decisiveness, etc.)?

The first-order principles of political psychology may well eventually prove quite similar to those of mainstream social psychology, with the important proviso that mainstream work needs to recognize boundary conditions highlighted by independent variables operating in political settings: game-theory pressures to keep up with the competition; role-theory and accountability pressures to take alternative perspectives into consideration; and the impact of periodic dramatic crises (Tetlock, 1998). Political actors are indeed human beings subject to the same basic process laws as ordinary mortals. But the second-order principles of political psychology will prove unsettlingly self-referential to the research community, requiring rigorous scrutiny of how its own ideological preconceptions and values shape its standards of evidence and proof for making pronouncements of irrationality or prejudice or miscalculation. Mapping political psychological principles will require understanding the workings of the minds not just of the research participants but also of the scientific observers. Our ability to rise to that challenge, without getting sucked into a tendentious infinite regress of impugning each other's motives, will ultimately determine the scientific viability of this intellectual enterprise.

REFERENCES

Abelson, R. P. (1959). Modes of resolution of belief dilemmas. *Journal of Conflict Resolution*, *3*, 343–352.

Adorno, T., Frenkel-Brunswick, E., Levinson, D., & Sanford, N. (1950). *The authoritarian personality*. New York: Harper & Row.

Aldag, R., & Fuller, S. R. (1993). Beyond fiasco: A reappraisal of the groupthink phenomenon and a new model of group decision processes. *Psychological Bulletin*, *113*, 533–552.

Allport, G. W. (1968). *The nature of prejudice*. Cambridge, MA: Addison-Wesley. (Original work published 1954)

Altemeyer, R. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.

Art, R. J., & Waltz, K. N. (1983). Technology, strategy, and the uses of force. In R. J. Art & K. N. Waltz (Eds.), *The use of force* (pp. 1–33). Lanham, MD: University Press of America.

Axelrod, R. (1976). *Structure of decision*. Boston: Little, Brown.

Axelrod, R. M. (1984). *The evolution of cooperation*. New York: Basic Books.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association Press.

Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2004). No place for nostalgia in science: A response to Arkes and Tetlock. *Psychological Inquiry*, *15*, 283–285.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, *87*(1), 5–22.

Bazerman, M. H. (2005). *Judgment in managerial decision making*. New York: Wiley.

Bell, D. (1992). *Faces at the bottom of the well*. New York: Basic Books.

Bendor, J. (2001). Aspiration-based reinforcement learning in repeated games: An overview. *International Game Theory Review*, *3*, 159–174.

Berman, L. S. (1982). *Planning a tragedy: The Americanization of the war in Vietnam*. New York: Norton.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, *61*, 27–42.

Blanton, S. L. (1996). Images in conflict: The case of Ronald Reagan and El Salvador. *International Studies Quarterly*, *40*, 23–44.

Blight, J. (1990). *The shattered crystal ball: Fear and learning in the Cuban missile crisis*. New York: Rowman & Littlefield.

Boettcher, W. (1995). Context, methods, numbers, and words: Prospect theory in international relations. *Journal of Conflict Resolutions*, *39*, 561–583.

Boring, E. G. (1950). *A history of experimental psychology*. New York: Appleton.

Bovens, M., & t'Hart, P. (1996). *Understanding policy fiascoes*. New Brunswick, NJ: Transaction.

Brecher, M. (1993). *Crisis in world politics: Theory and reality*. New York: Pergamon Press.

Breslauer, G. W., & Tetlock, P. E. (Eds.). (1991). *Learning in U.S. and Soviet foreign policy*. Boulder, CO: Westview Press.

Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 554–594). New York: McGraw-Hill.

Brewer, M. B., & Kramer, R. M. (1985). The psychology of intergroup attitudes and behavior. *Annual Review of Psychology*, *36*, 219–243.

Brockner, J. & Rubin, J. (1985). *Entrapment in escalating conflicts: A social psychological analysis*. New York: Springer Verlag

Camerer, C. (1995). Individual decision making. In J. H. Hagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 587–704). Princeton, NJ: Princeton University Press.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*, 1–3, 7–42.

Chee, M. W. L., Sriram, N., Soon, C. S., & Lee, K. M. (2000). Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *NeuroReport*, *11*, 135–140.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *12*, 163–170.

Danilovic, D. (2004). Conceptual and selection bias issues in deterence. *Journal of Conflict Resolution*, *45*, 97–125.

Dawes, R. M. (1998). Behavioral decision making and judgment. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 497–548). New York: McGraw-Hill.

Deutsch, M. (1983). The prevention of World War III: A psychological perspective. *Political Psychology*, *4*, 3–31.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*, 62–68.

Downs, G. W. (1991). Arms races and war. In P. E. Tetlock, R. Jervis, P. Stern, J. L. Husbands, & C. Tilly (Eds.), *Behavior, society, nuclear war* (Vol. 2, pp. 73–109). New York: Oxford University Press.

Druckman, J. (2004). Political preference formation: Competition, deliberation, and the (ir)relevance of framing effects. *American Political Science Review*, *98*, 671– 685.

Durkheim, E. (1925). *The elementary forms of the religious life*. London: Allen and Unwin.

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 269–322). New York: McGraw-Hill.

Einhorn, H., & Hogarth, R. (1981). Behavioral decision theory. *Annual Review of Psychology*, *31*, 53–88.

Etheredge, L. S. (1980). *A world of men*. Cambridge, MA: MIT Press.

Farnham, B. (1992). Roosevelt and the Munich crisis: Insights from prospect theory. *Political Psychology*, *13*, 242–272.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Festinger, L. (1964). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I," the "A," and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*(3), 74–147.

Fischhoff, B. (1991). Nuclear decision-making. In P. E. Tetlock, R. Jervis, P. Stern, J. L. Husbands, & C. Tilly (Eds.), *Behavior, society, nuclear war* (Vol. 2, pp. 110–192). New York: Oxford University Press.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin*, *30*, 1611–1624.

Friedman, T. (1999). *The lexus and the olive tree*. New York: Farrar, Straus, Giroux.

George, A. L. (1980). *Presidential decision making in foreign policy: The effective use of information and advice*. Boulder, CO: Westview Press.

Gergen, K. (1978). Experimentation in social psychology: A reappraisal. *European Journal of Social Psychology*, *8*, 507–527.

Gibson, J. L. (2006). Enigmas of intolerance: Fifty years after Stouffer's Civil Liberties. *Perspectives on Politics*, *1*, 21–34.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilbert, D., & Malone, P. S. (1995). Correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Gilovich, T. (1981). Seeing the past in the present: The effect of associations to familiar events on judgments and decisions. *Journal of Personality and Social Psychology*, *40*, 797–308.

Gilovich, T., Griffin, D. W., & Kahneman, D. (2000). *Inferences, heuristics, and biases: New directions in judgment under uncertainty*. New York: Cambridge University Press.

Goldgeier, J., & Tetlock, P. E. (2000). Psychology and international relations theory. In N. Polsby (Ed.), *Annual review of political science* (Vol. 4, pp. 67–92). Palo Alto, CA: Annual Reviews Press.

Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation—the left, the right and the rigid: Comment on Jost et al. (2003). *Psychological Bulletin*, *129*, 376–382.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1991). A terror management theory of social behavior: the psychological functions of self-esteem and cultural worldviews. *Advances in Experimental Social Psychology*, *24*, 93–159.

Greenstein, F. I. (1975). *Personality and politics*. New York: Norton.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.

Gruenfeld, D. H. (1995). Status, ideology, and integrative complexity on the U.S. Supreme Court: Rethinking the politics of political decision making. *Journal of Personality and Social Psychology*, *68*, 5–20.

Guttieri, K., Wallace, M. D., & Suedfeld, P. (1995). The integrative complexity of decision makers in the Cuban missile crisis. *Journal of Conflict Resolution*, *39*, 595–621.

Hempel, C. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.

Heradstveit, D., & Bonham, G. M. (1996). Attribution theory and Arab images of the Gulf war. *Political Psychology*, *17*, 271–292.

Herek, G., Janis, I. L., & Huth, P. (1987). Decision making during international crises: Is the quality of process related to outcomes? *Journal of Conflict Resolution*, *31*, 203–226.

Hermann, M. (1986). *Handbook of political psychology*. San Francisco: Jossey-Bass.

Herrmann, R. (1988). The empirical challenge of the cognitive revolution: A strategy for drawing inferences about perceptions. *International Studies Quarterly*, *32*, 175–203.

Herrmann, R., & Fischerkeller, M. (1995). Beyond the enemy image and spiral model: Cognitive-strategic research after the Cold War. *International Organization*, *49*, 415–450.

Herrmann, R., Tetlock, P. E., & Visser, P. (1999). Mass public decisions on going to war: A cognitive-interactionist framework. *American Political Science Review*, *93*, 553–574.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Holsti, O. R. (1967). Cognitive dynamics and images of the enemy. In R. Fagan (Ed.), *Enemies of politics* (pp. 25–96). Chicago: Rand McNally.

Holsti, O. R. (1989). Crisis decision making. In P. E. Tetlock, R. Jervis, P. Stern, J. L. Husbands, & C. Tilly (Eds.), *Behavior, society, nuclear war* (pp. 8–84). New York: Oxford University Press.

Holsti, O. R., & Rosenau, J. N. (1979). Vietnam, consensus, and the belief systems of American leaders. *World Politics*, *32*, 1–56.

Jackman, M. R., & Jackman, R. (1983). *Class awareness in the United States*. Berkeley: University of California Press.

Janis, I. L. (1982). *Groupthink*. New York: Free Press.

Janis, I. L. (1989). *Crucial decisions: Leadership in policy making and crisis management*. New York: Free Press.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice and commitment*. New York: Free Press.

Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.

Jervis, R. (1989). Rational deterrence: Theory and evidence. *World Politics*, *41*, 183–207.

Jervis, R. (2006). Understanding beliefs. *Political Psychology*, *5*, 641–663.

Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107–117.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003a). Exceptions that prove the rule—Using a theory of motivated social cognition to account for ideological incongruities and political anomalies: Reply to Greenberg and Jonas (2003). *Psychological Bulletin*, *129*, 383–393.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003b). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*, 339–375.

Jost, T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, *38*, 586–602.

Judd, C. M., Blair, I., & Chapleau, K. (2004). Automatic stereotypes vs. automatic prejudice: Sorting out the possibilities in the Payne (2001) weapon paradigm. *Journal of Experimental Social Psychology*, *40*, 75–81.

Jussim, L. (2005). Accuracy: Criticisms, controversies, criteria, components, and cognitive processes. *Advances in Experimental Social Psychology*, *37*, 1–93.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kahneman, D., & Tversky, A. (2001). *Choice, values, frames*. New York: Cambridge University Press.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, *81*, 774–788.

Kelley, H. H. (1965). Experimental studies of threats in negotiations. *Journal of Conflict Resolution*, *9*, 77–105.

Kelman, H. C. (1983). Conversations with Arafat: A social-psychological assessment of the prospects for Israeli-Palestinian peace. *American Psychology*, *38*, 203–216.

Kelman, H. C., & Bloom, A. (1973). Assumptive frameworks in international politics. In J. Knutson (Ed.), *Handbook of political psychology*. San Francisco: Jossey-Bass.

Kennedy, P. (1987). *The rise and fall of the great powers*. New York: Random House.

Khong, Y. F. (1991). *Analogies at war*. Princeton, NJ: Princeton University Press.

Kinder, D. R. (1986). The continuing American dilemma: White resistance to racial change 40 years after Myrdal. *Journal of Social Issues*, *42*, 151–171.

Kinder, D. R. (1998). Opinion and action in the realm of politics. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 778–867). Boston: McGraw-Hill.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis-testing. *Psychological Review*, *94*, 211–228.

Kleiboer, M. A. (1996). Understanding mediation success in international conflict: Review and prospects. *Journal of Conflict Resolution*, *40*, 360–389.

Knutson, J. (Ed.). (1973). *Handbook of political psychology*. San Francisco: Jossey-Bass

Koh, J., Han, J., & Kim, J. (2004). Integrative complexity of South-North Korean correspondences: A time-series analysis, 1984–1994. *Journal of Conflict Resolution*, *46*, 286–304.

Kramer, R. M. (1988). Windows of vulnerability or cognitive illusions? Cognitive processes and the nuclear arms race. *Journal of Experimental Social Psychology*, *25*, 79–100.

Kruglanski, A. (2004). *The psychology of closed-mindedness*. New York: Psychology Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lane, R. E. (1973). Patterns of political belief. In J. N. Knutson (Ed.), *Handbook of political psychology*. San Francisco: Jossey-Bass.

Lasswell, H. D. (1948). *The analysis of political behavior: An empirical approach*. London: Routledge & Kegan Paul.

Lebow, R. N. (1981). *Between peace and war*. Baltimore: Johns Hopkins University Press.

Leng, R. J. (1993). Influence techniques. In P. E. Tetlock, R. Jervis, P. Stern, J. L. Husbands, & C. Tilly (Eds.), *Behavior, society, nuclear war* (Vol. 3, pp. 71–125). New York: Oxford University Press.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159.

Levi, A., & Tetlock, P. E. (1980). A cognitive analysis of Japan's 1941 decision to go to war. *Journal of Conflict Resolution*, *24*, 195–211.

Levi, A., & Whyte, G. (1997). A cross-cultural explanation of reference dependence of crucial group decisions under risk: Japan's 1941 decision for war. *Journal of Conflict Resolution*, *41*, 792–813.

Levy, J. S. (1989). The causes of war: A review of theories and evidence. In P. E. Tetlock, R. Jervis, P. Stern, J. L. Husbands, & C. Tilly (Eds.), *Behavior, society, nuclear war* (Vol. 1, pp. 209–333). New York: Oxford University Press.

Levy, J. S. (1992). Prospect theory and international relations: Theoretical applications and analytical problems. *Political Psychology*, *13,* 283–310.

Levy, J. S. (1994). Learning and foreign policy: Sweeping a conceptual minefield. *International Organization*, *48*, 279–312.

Lewin, K., Lippitt, R., & White, R. (1939). Patterns of aggressive behavior in experimentally created "social climates. *Journal of Social Psychology*, *10*, 271–299.

Lindskold, S. (1978). Trust development, the GRIT proposal and the effects of conciliatory acts on conflict and cooperation. *Psychology Bulletin*, *85*, 772–788.

Loewenstein, R. (2000). *When genius failed*. New York: Random House.

MacCoun, R. (1998). Biases in the interpretation and use of research results. *Annual Review of Psychology*, *49*, 259–287.

Mack, J. (1985). Toward a collective psychopathology of the nuclear arms competition. *Political Psychology*, *6*, 291–321.

Maoz, I., Ward, A., Katz, M., & Ross, L. (2004). Reactive devaluation of an "Israeli" vs. "Palestinian" peace proposal. *Journal of Conflict Resolution*, *46*, 515–546.

Maoz, Z. (1981). The decision to raid Entebbe: Decision analysis applied to crisis behavior. *Journal of Conflict Resolution*, *25*, 677–708.

Maoz, Z., & Shayer, A. (1987). The cognitive structure of peace and war argumentation: Israeli prime ministers versus the Knesset. *Political Psychology*, *8*, 575–604.

McClelland, J. S. (1996). *A history of Western political thought*. London: Routledge.

McClintock, C. G., Stech, F., & Beggan, J. (1987). The effects of commitments to threats and promises upon bargaining behavior and outcomes. *European Journal of Social Psychology*, *17*, 447–64.

McClosky, H. (1967). Personality and attitude correlates of foreign policy orientation. In J. N. Rosenau (Ed.), *Domestic sources of foreign policy* (pp. 51–110). New York: Free Press.

McDermott, R. (1998). *Risk-taking in international politics: Prospect theory in American foreign policy*. Ann Arbor: University of Michigan Press.

McGuire, W. J. (1983). A contextualist theory of knowledge: Its implications for innovation and reform in psychological research. In *Advances in experimental and social psychology* (Vol. 16, pp. 1–47). New York: Academic Press.

Mellers, B. A., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*, 269–275.

Merton, R. (1987). Three fragments from a sociologist's notebooks: Establishing the phenomenon, specified ignorance, and strategic research materials. *Annual Review of Sociology*, *13*, 1–28.

Mills, C. W. (1959). *The sociological imagination*. New York: Oxford University Press.

Mintz, A. (1993). The decision to attack Iraq: A noncompensatory theory of decision-making. *Journal of Conflict Resolution*, *37*, 595–618.

Mitchell, P. G., & Tetlock, P. E. (2006). Anti-discrimination law and the perils of mind reading. *Ohio State University Law Review*.

Mitchell, P. G., Tetlock, P. E., Mellers, B. A., & Ordonez, L. (1993). Judgments of social justice: Compromises between equality and efficiency. *Journal of Personality and Social Psychology, 65*, 629–639.

Mitchell, P. G., Tetlock, P. E., Newman, D., & Lerner, J. (2003). Experiments behind the veil: A hypothetical societies approach to the study of social justice. *Political Psychology*, *24*, 519–547.

Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition*, *19*, 395–417.

Murray, S., & Meyers, J. (2004). Do people need foreign enemies? American leaders' beliefs after the Soviet demise. *Journal of Conflict Resolution*, *43*, 555–569.

Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.

Neustadt, R. E., & May, E. R. (1986). *Thinking in time: The uses of history for decision-makers*. New York: Free Press.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Patchen, M. (1987). Strategies for eliciting cooperation from an adversary: Laboratory and international findings. *Journal of Conflict Resolution*, *31*, 164–185.

Payne, J., Bettman, J. R., & Johnson, E. J. (1992). Behavioral deci-

sion theory: A constructive processing perspective. *Annual Review of Psychology*, *43*, 87–131.

Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology, 7*, 1–18.

Peng, K., & Nisbett, R. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *51*, 1–14.

Peterson, B. E., Winter, D. G., & Doty, R. M. (1994). Laboratory tests of a motivational-perceptual model of conflict escalation. *Journal of Conflict Resolution*, *38*, 719–748.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*, 729–738.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Poehlman, A., Uhlmann, E., Greenwald, A. G., & Banaji, M. (2005). *Understanding and using the Implicit Association Test: Meta-analysis of predictive validity*. Unpublished manuscript, Yale University.

Posner, R. (2005). *Uncertain shield: The U.S. intelligence system in the throes of reform*. Baltimore: Rowman & Littlefield.

Post, J. (1991). Saddam Hussein of Iraq: A political psychology profile. *Political Psychology*, *212*, 279–290.

Postmes, T., Spears, T., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology*, *80*, 918–930.

Proctor, R. (1991). *Value-free science?: Purity and power in modern knowledge*. Cambridge, MA: Harvard University Press.

Pruitt, D. (1998). Social conflict. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 470–503). New York: McGraw-Hill.

Pruitt, D., & Rubin, J. (1986). *Social conflict: Escalation, stalemate, and settlement*. New York: Random House.

Putnam, H. (2002). *The collapse of the fact/value dichotomy and other essays*. Cambridge, MA: Harvard University Press.

Putnam, R. D. (1973). *The beliefs of politicians: Ideology, conflict, and democracy in Britain and Italy*. New Haven, CT: Yale University Press.

Putnam, R. D. (1988). Diplomacy and domestic politics: The logic of two-level games. *International Organization, 42*, 427–460.

Raphael, T. D. (1982). Integrative complexity theory and forecasting international crisis. *Journal of Conflict Resolution*, *26*, 423–50.

Redding, R. (2004). Bias on prejudice?: The politics of research on racial prejudice? *Psychological Inquiry*, *15*, 289–293.

Reiter, D. (1996). *Crucible of beliefs*. Ithaca, NY: Cornell University Press.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

Ross, L., & Griffin, D. (1991). Subjective construal, social inference, and human misunderstanding. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319–359). New York: Academic Press.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test: Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*(2), 139–165.

Sagan, S., & Waltz, K. (1995). *The spread of nuclear weapons: A debate*. New York: Norton.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518.

Schelling, T. C. (1966). *Arms and influence*. New Haven, CT: Yale University Press.

Schuman, H., Steeh, C., Bobo, L., & Krysan, M. (1997). *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press.

Searle, J. (1984). *Minds, brains, and science*. Cambridge, MA: Harvard University Press.

Sears, D. (2004). One survey researcher's perspective on implicit prejudice. *Psychological Inquiry*, *15*(4), 293–297.

Sears, D. Huddy, L., & Jervis, R. (2003). *Oxford handbook of political psychology*. New York: Oxford University Press.

Seligman, M. (1990). *Learned optimism*. New York: Basic Books.

Shephard, G. (1988). Personality effects on American foreign policy, 1969–1984. *International Studies Quarterly*, *32*, 91–123.

Sidanius, J., & Pratto F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Smith, M. B., Bruner, J., & White, R. (1956). *Opinions and personality*. New York: Wiley.

Smith, V. (1991). Rational choice: The contrast between economics and psychology. *Journal of Political Economy*, *99*, 877–897.

Sniderman, P. M., & Carmines, E. G. (1997). *Reaching beyond race*. Cambridge, MA: Harvard University Press.

Sniderman, P. M., & Tetlock, P. E. (1986). Symbolic racism: Problems of motive attribution in political debate. *Journal of Social Issues*, *42*, 129–150.

Spiegel, S. (1991). Learning in U.S. foreign policy: The case of the Middle East. In G. Breslauer & P. E. Tetlock (Eds.), *Learning in U.S. and Soviet foreign policy* (pp. 264–301). Boulder, CO: Westview.

Spillman, K. R., & Spillman, K. (1991). On enemy images and conflict resolution. *International Social Science Journal*, *43*, 57–76.

Staw, B. (1980). Rationality and justification in organizational life. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 2, pp. 45–80). Greenwich, CT: JAI Press.

Staw, B., Sandelands, L., & Dutton, J. (1981) Threat-rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly*, *26*, 501–524.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.

Steele, S. (2006). *White guilt: How blacks and whites together destroyed the promise of the Civil Rights era*. New York: HarperCollins.

Stein, H. (2004, January 30). Dumb and dumber. *Wall Street Journal*, p. A15.

Stein, J., & Tanter, R. (1980). *Rational decision making: Israel's security choices*. Columbus: Ohio State University Press.

Sternberg, R. J., & Soriano, L. (1984). Styles of conflict resolution. *Journal of Personality and Social Psychology*, *47*, 115–126.

Stone, W. P. (1980). The myth of left-wing authoritarianism. *Political Psychology*, *2*, 3–20.

Streufert, S., & Streufert, S. (1978). *Behavior in the complex environment*. Washington, DC: Winston and Sons.

Suedfeld, P. (1992). Cognitive managers and their critics. *Political Psychology*, *13*.

Suedfeld, P., & Tetlock, P. E. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolution*, *21*, 169–84.

Suedfeld, P., & Tetlock, P. E. (1991). Psychologists as policy advocates: The roots of controversy. In P. Suedfeld & P. E. Tetlock (Eds.), *Psychology and social policy* (pp. 1–37). Washington, DC: Hemisphere.

Suedfeld, P., & Tetlock, P. E. (2001). Cognitive styles. In A. Tesser & N. Schwartz (Eds.), *Blackwell international handbook of social psychology: Intra-individual processes* (Vol. 1, pp. 284–304). London: Blackwell.

Suppe, F. (Ed.). (1977). *The structure of scientific theories*. Chicago: University of Illinois Press.

Svenson, O., & Maule, A. J. (1994). *Time pressure and stress in human judgment and decision making*. New York: Plenum Press.

Swets, J. A. (1982). *Evaluation of diagnostic systems: Methods from signal detection theory*. New York: Academic Press.

t'Hart, P. (1994). *Groupthink in government*. Baltimore: Johns Hopkins University Press.

t'Hart, P., Stern, E., & Sundelius, B. (1995). *Beyond groupthink*. Stockholm, Sweden: Stockholm Center for Organizational Research.

Taylor, S. (1998). The social being in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 1, pp. 35–68). New York: McGraw-Hill.

Taylor, S., & Brown, J. (1988). Illusion and well-being: A social psy-

chological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Tetlock, P. E. (1979). Identifying victims of groupthink from public statements of decision makers. *Journal of Personality and Social Psychology*, *37*, 1314–1324.

Tetlock, P. E. (1981). Pre- to post-election shifts in presidential rhetoric: Impression management or cognitive adjustment? *Journal of Personality and Social Psychology*, *41*, 207–212.

Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social Psychology*, *45*, 118–126.

Tetlock, P. E. (1984). Cognitive style and political belief systems in the British House of Commons. *Journal of Personality and Social Psychology*, *46*, 365–375.

Tetlock, P. E. (1985). Integrative complexity of American and Soviet foreign policy rhetoric: A time-series analysis. *Journal of Personality and Social Psychology*, *49*, 1565–1585.

Tetlock, P. E. (1986). A value pluralism model of ideological reasoning. *Journal of Personality and Social Psychology*, *50*, 819–827.

Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331–376). New York: Academic Press.

Tetlock, P. E. (1994). Political psychology or politicized psychology: Is the road to scientific hell paved with good moral intentions? *Political Psychology*, *15*, 509–530.

Tetlock, P. E. (1998). Social psychology and world politics. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 868–912). New York: McGraw-Hill.

Tetlock, P. E. (1999a). Prospecting for evidence for prospect theory. Review of R. McDermott, Risk-taking in international politics: Prospect theory in American foreign policy. *Contemporary Psychology*, *44*, 399–401.

Tetlock, P. E. (1999b). Theory-driven reasoning about possible pasts and probable futures: Are we prisoners of our preconceptions? *American Journal of Political Science*, *43*, 335–366.

Tetlock, P. E. (2000). Cognitive biases and organizational correctives: Do both disease and cure depend on the ideological beholder? *Administrative Science Quarterly*, *45*, 293–326.

Tetlock, P. E. (2002). Social-functionalist frameworks for judgment and choice: The intuitive politician, theologian, and prosecutor. *Psychological Review, 109*, 451–472.

Tetlock, P. E. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton University Press.

Tetlock, P. E., & Arkes, H. (2004). The implicit-prejudice exchange: Islands of consensus in a sea of controversy. *Psychological Inquiry*, *15*, 311–321.

Tetlock, P. E., Armor, D., & Peterson, R. (1994). The slavery debate in antebellum America: Cognitive style, value conflict, and the limits of compromise. *Journal of Personality and Social Psychology*, *66*, 115–126.

Tetlock, P. E., & Belkin, A. (Eds.). (1996). *Counterfactual thought experiments in world politics*. Princeton, NJ: Princeton University Press.

Tetlock, P. E., & Boettger, R. (1989). Cognitive and rhetorical styles of traditional and reformist Soviet politicians: A content analysis study. *Political Psychology*, *10*, 209–232.

Tetlock, P. E., & Boettger, R. (1994). Accountability amplifies the status quo effect when change creates victims. *Journal of Behavioral Decision Making*, *7*, 1–23.

Tetlock, P. E., Hannum, K., & Micheletti, P. (1984). Stability and change in senatorial debate: Testing the cognitive versus rhetorical style hypotheses. *Journal of Personality and Social Psychology*, *46*, 979–990.

Tetlock, P. E., & Lerner, J. S. (1999). The social contingency model: Identifying empirical and normative boundary conditions on the error-and-bias portrait of human nature. In S. Chaiken & Y.

Trope (Eds.), *Dual-process models in social psychology* (pp. 571–585). New York: Guilford Press.

Tetlock, P. E., McGuire, C., & Mitchell, P. G. (1991). Psychological perspectives on nuclear deterrence. *Annual Review of Psychology*, *42*, 239–276.

Tetlock, P. E., & Mitchell, P. G. (1993). Liberal and conservative approaches to justice: Conflicting psychological portraits. In B. Mellers & J. Baron (Eds.), *Psychological perspectives on justice* (pp. 234–255). Cambridge, UK: Cambridge University Press.

Tetlock, P. E., Peterson, R., & Lerner, J. (1996). Revising the value pluralism model: Incorporating social content and context postulates. In C. Seligman, J. Olson, & M. Zanna (Eds.), *Ontario Symposium on Social and Personality Psychology* (Vol. 64, pp. 500–511). Toronto: University of Toronto.

Tetlock, P. E., Peterson, R., McGuire, C., Feld, P., & Chang, S. (1992). Assessing political group dynamics: A test of the groupthink model. *Journal of Personality and Social Psychology*, *63*, 402–423.

Tetlock, P. E., & Tyler, A. (1996). Winston Churchill's cognitive and rhetorical style: The debates over Nazi intentions and self-government for India. *Political Psychology*, *17*, 149–170.

Tetlock, P. E., Visser, P., Singh, R., Polifroni, M., Elson, B., Scott, A., et al. (in press). People as intuitive prosecutors: The impact of social control motives on attributions of responsibility. *Journal of Experimental Social Psychology*.

Thernstrom, S., & Thernstrom, A. (1997). *America in black and white*. New York: Simon & Schuster.

Thompson, L. (1990). Negotiation behavior and outcomes: Empirical evidence and theoretical issues. *Psychological Bulletin*, *108*, 515–532.

Turner, M., Pratkanis, A. R., Probasco, P., & Love, C. (1992). Threat, cohesion, and group effectiveness: Testing a social identity maintenance perspective on groupthink. *Journal of Personality and Social Psychology*, *63*, 781–796.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.

Uhlmann, E., Brescoll, V. L., & Paluck, E. L. (2006). Are members of low status groups perceived as bad, or badly off? Egalitarian negative associations and automatic prejudice. *Journal of Experimental Social Psychology*, *42*, 491–505.

Vertzberger, Y. (1990). *The world in their minds*. Stanford, CA: Stanford University Press.

Vertzberger, Y. (1995). Rethinking and reconceptionalizing risk in foreign policy decision-making: A socio cognitive approach. *Political Psychology*, *16*, 347–380.

Walker, S. (1983). The motivational foundations of political belief systems: A reanalysis of the operational code construct. *International Studies Quarterly*, *27*, 177–201.

Walker, S. G., & Watson, G. L. (1994). Integrative complexity and British decisions during the Munich and Polish crises. *Journal of Conflict Resolution*, *38*, 3–23.

Waltz, K. N. (1979). *Theory of international politics*. Reading, MA: Addison-Wesley.

White, R. K. (1984). *Fearful warriors: A psychological profile of U.S.–Soviet relations*. New York: Free Press.

Wilson, J. Q. (1989). *Bureaucracy*. New York: Free Press

Winter, D. G. (1993). Power, affiliation, and war: Three tests of motivational model. *Journal of Personality and Social Psychology*, *65*, 532–545.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.

Zimmerman, W., & Axelrod, R. M. (1981). The "lessons" of Vietnam and Soviet foreign policy. *World Politics*, *34*, 1–24.

CHAPTER 40

Organizational Behavior

LEIGH THOMPSON
JO-ELLEN POZNER

Organizational behavior (OB) is the study of how people think, feel, and act in organizations and similarly, how they are affected by the activities within organizations. OB is interested in the behavior of people embedded in specific contexts of organizational systems. Within this definition, there are three important levels of analysis that guide the field: the individual, the team, and the organization. Not surprisingly, OB has much in common with social psychology; however, the context of application—the organizational environment—is much more defined than it is in social psychology.

In this chapter, we highlight the key themes and principles that guide the study of OB, paying particular attention to the ways in which OB has been influenced by social psychology. We organize our chapter around the three key levels of analysis that undergird the study of OB. The first level of analysis within OB is that of the organizational actor as an independent decision maker or leader. Within this section, we focus on decision making, psychological contracts, justice and fairness, and, of course, leadership. We appreciate that all the concepts just listed do not occur in a social vacuum and therefore might very well occur in the context of teams and decision making in organizations; what is common among these phenomena, however, is that they are considered properties or characteristics of the organizational actor.

A second level of analysis considers the organizational actor as a team member. Within this section, we delve into perhaps what is considered to be the fundamental building block of most organizations and companies: the team. We are careful to draw a clear distinction between groups as they are commonly studied in the social psychology literature and teams as they are studied in the OB literature. In this section, we analyze the key processes of negotiation and teamwork. A final level of analysis is that of the organization in which actors are embedded. More than any other section, this is where social psychology and organizational behavior often part company.

Threaded throughout our discussion of these key themes in organizational behavior, we articulate the two key considerations facing the OB researcher: methodology and application. In terms of methodology, we argue that in contrast to social psychology, wherein there is near universal agreement that the laboratory method is necessary and (perhaps) sufficient, that considerably more methodological variety exists in OB. Accordingly, we are careful to point domains in which field researchers have placed a "flag" and domains in which lab researchers have claimed territory. In some cases, the subject matter necessitates a field study (e.g., organizational downsizing, social networks, and, often, leadership). In terms of application, we argue that in contrast to social psychology, wherein application is often an afterthought, it is a crucial consideration in organizational behavior research. Thoughtful application of principles of OB requires careful consideration of the realities of organizational life, including cost–benefit analysis, reorganization, globalization, and industry developments.

We conclude by considering the future of OB. To that end, we review some of the topics and themes that do not fit neatly into the aforementioned basic processes. We

suggest that in the future, OB will grow more divergent from social psychology, and we explain why we make this prediction using both theory and paradigmatic logic.

THE ORGANIZATIONAL ACTOR AS INDEPENDENT THINKER AND DECISION MAKER

The fundamental unit of analysis in OB, like social psychology, is the individual. Whereas OB privileges individual cognition, emotion, and behavior, it pays particular attention to the social and organizational contexts in which they are situated, as well as the consequences of those contexts on social life and social interaction.

One well-developed and highly influential area of research is individual decision making, which is inspired by the seminal work of Kahneman and Tversky (1979). Another important area of OB research is procedural justice—people's reaction to events based on their perceptions of the fairness of relevant processes and outcomes—as well as the behavioral consequences that follow. Finally, OB is interested in how individuals as leaders can understand and manage the cognitive and emotional reactions of their subordinates; thus we also review several theories on leadership.

Decision Making

Within the topic of decision making, there are three identifiable streams of research, including (1) flawed and biased decision making; (2) justification processes and pressures; and (3) organizational boundary effects.

Flawed and Biased Decision Making

This area of research in OB is a direct descendant of the seminal research of March and Simon (1958, 1993) and Kahneman and Tversky (1979). The fundamental theme is that organizational decision makers, despite being intelligent and presumably motivated by noble and benevolent reasons, are hopelessly victimized by their own nonrational thought processes (for an overview, see Bazerman, 2006). Consequently, although decisions often deviate from what is strictly rational, according to economists, they are somewhat predictable in light of OB decision-making research.

BEHAVIORAL DECISION RESEARCH TRADITION

With their seminal publication, *Judgment under Uncertainty: Heuristics and Biases*, Kahneman, Slovic, and Tversky (1982) changed the future of management and psychology by ushering in an era of experimental research on descriptive biases. Their book neatly outlined seven types of biases, including representativeness, causality and attribution, availability, covariation and control, overconfidence, multistage evaluation, and risk perception that served as the stage for the next 20 years of research on bias. Whereas bias approach primarily focused on cognitive biases for several years (availability

heuristic, representativeness heuristic, anchoring heuristic, framing bias, endowment effects, etc.), more recent treatments of bias implicate human motivation as a primary source of bias. Examples of biases that have a motivational core include the positive illusion bias, self-serving bias, and judgments that are affected by mood (for a review, see Bazerman, 2006).

The metaphor of the hopelessly "biased" decision maker within micro-organizational behavior has almost exclusively been applied to the individual. Moreover, the great body of research in micro-organizational behavior has relied on the individual as the fundamental unit of analysis. For example, Bazerman's (2006) sixth edition of his book, *Judgment in Managerial Decision Making*, nearly exclusively relies on the individual manager as the source of bias in organizational decision making. However, among macro-organizational theorists, bias occurs at the level of the organization.

BOUNDED RATIONALITY

The concept of bounded rationality was proposed by Herbert Simon (1947) and developed by the Carnegie School (Cohen, March, & Olsen, 1972; Cyert & March, 1963; March, 1962; March & Simon, 1958). Simon's fundamental insight is that the extent to which decision making can be rational is limited, or "bounded," by the fact that decision makers never have access to all the information they need about their alternatives or the consequences of any given decision. Actors therefore satisfice, rather than optimize, and focus their limited attention on a small, rather than expansive, range of issues, which are generally more salient to themselves. Because the situationally and structurally determined attention structures within an organization (Ocasio, 1997) lead the actor to focus on a unique set of issues, which often conflict with those of other actors, and because of the interdependence inherent within organizations, actors are forced to seek coalitions to achieve their goals (March, 1962). Arguing that information-processing capabilities prevent individuals from behaving perfectly rationally, the Carnegie School built a behavioral theory of the firm, dubbed the "garbage can model" of organizational choice.

GARBAGE CAN MODEL

An important implication of bounded rationality at the individual level is that organizational decision making—the outcome of many interdependent individual decisions—cannot itself be rational. Cyert and March (1963) sought to analyze ambiguous behaviors, or behavior that appears to contradict classical theory, at the organizational level. Underlying the garbage can model is the understanding that, when aggregated, bounded rationality leads to extremely uncertain decision environments, where it is impossible to employ classical decision theory.

In addition to bounded rationality, the presence of political coalitions of organizational subunits, conflicting goals, problematic preferences (i.e., the organization as a whole does not have a clear preference ordering), un-

clear technology (i.e., organizational processes are not fully understood by all participants), and fluid participation (i.e., sequential and selective attention to issues leads to different actors being involved in decisions over time) characterize the organizational decision-making environment, which can be termed "organized anarchy" (Cohen et al., 1972). Organizational choice is therefore the result of unique and somewhat unpredictable combinations of problems and solutions facing unique combinations of decision makers at a given point in time. Problems and choices are therefore partially decoupled from one another, and decisions are made only when the combination of problems, solutions, and decision makers facilitates action (Cohen et al., 1972; Cohen, March, & Olsen, 1976; March & Olsen, 1976). Grounded in theory as well as in field studies and computer simulations, the garbage can model and the Carnegie School represent a rich tradition in organizational behavior, which has influenced a wide array of literatures, from more microlevel views of decision making to the study of groups, learning, power, and leadership.

Justification Processes and Pressures

Accountability is generally assumed to have a positive effect on the quality of decision making. Tetlock's (1983) theory of accountability suggests that the decision maker who was effectively responsible for her decision would be more vigilant than the decision maker who was under no particular pressure to justify a decision. In a series of investigations, Tetlock and his colleagues (Tetlock, 1985; Tetlock & Kim, 1987) manipulated accountability, typically by informing decision makers at the outset of a judgment task that they would have to present their findings to a panel of experts or be otherwise reviewed and evaluated. Under such accountability conditions, decision makers were more accurate (Tetlock, 1985; Tetlock & Kim, 1987) and took longer to make decisions than when not under accountability pressure.

Tetlock distinguished different types of accountability, most notably process accountability (the extent to which a person or group must answer questions about how they came to a decision) and outcome accountability (the extent to which a person or group must account for the outcome or conclusion of their decision). Tetlock reasoned that process accountability was more effective than was outcome accountability in producing better decisions. Outcome accountability might lead a decision maker to arrive at a decision that would meet the approval of the organization; process accountability, on the other hand, presumably motivates decision makers to review the evidence without regard for the outcome that it might suggest.

Organizational Boundary Issues That Shape Decision Making

In recent years, the study of decision making has become decidedly more informed by an understanding of the organizational actor's relationship to the company.

INSIDER VERSUS OUTSIDER VIEWS

Kahneman and Lovallo (1993) suggest that organizational decision makers have a strong tendency to consider the organizational problems they face as unique, when in fact problems in organizations reoccur with some regularity. Managers tend to view current problems in isolation, while discounting or neglecting past experience and statistics, which might in fact be helpful in evaluating current decisions. Kahneman and Lovallo refer to this tendency to forget the past as an "insider view" of problems, which anchors predictions about the future on plans and scenarios. Kahneman and Lovallo argue that decision makers should adopt an "outsider view" of problems, which would bring relevant data and information to bear, and allow the decision maker to disassociate him- or herself from a problem.

KNOWLEDGE VALUATION

Menon and her colleagues (Menon & Blount, 2003; Menon, Choi, & Thompson, 2005; Menon & Pfeffer, 2003) find that decision makers' valuations of knowledge affecting their decision making are highly influenced by whether the information is perceived to emanate from organizational insiders or outsiders. Whereas traditional group research in social psychology posits a near law-like ingroup favoritism effect, Menon argues that within the organization, actors battle for status and distinctiveness (see also Sutton & Hargadon, 1996), making the use of knowledge contributed by an insider tantamount to career suicide. In controlled, empirical investigations of insider versus outsider knowledge, Menon finds that people tend to ignore, discount, and devalue knowledge that is presumably sourced from insiders, whereas the same knowledge sourced from an outsider is more carefully attended to and more highly valued. Menon and colleagues (2005) reveal that managers who have an opportunity to affirm themselves are less likely to experience status threat and therefore more likely to make use of valuable insider knowledge.

Prescriptive Approaches in Judgment and Decision Making

The question of how to mitigate or eliminate human bias in organizational decision making has received relatively short shrift in comparison to descriptive accounts of how biases emerge and operate. One explanation for this dearth of prescriptive treatments is that it is very difficult to reverse or mitigate fundamental human bias. Indeed, investigations of expert decision makers reveal that experts often fall prey to the same biases that befall novices (for a review, see Arkes & Hammond, 1986). Bazerman (2006) outlines six strategies for combating bias in individual decision making: (1) acquiring experience and expertise; (2) debiasing techniques, that often involve consciousness raising about biases; (3) analogical reasoning that involves highlighting biases that occur across a range of situations, allowing decision makers to induce a common theme (cf. Loewenstein, Thompson, & Gentner,

1999); (4) taking the outsider view; (5) using linear models; and (6) understanding biases in others.

Justice and Fairness

Pay and access to resources are critical concerns for most organizational actors. Not surprisingly, rewards and resource allocation, as well as actors' reactions to allocation decisions, are therefore central topics in OB. Research that sheds light on the processes by which organizational actors assess fairness, as well as the consequences of those assessments, is important for two reasons. First, it presents a limitation to the traditional economic perspective of individuals as atomistic, self-interested actors from a microperspective. In addition, it contrasts sharply with research on interpersonal decision making (e.g., Messick & Sentis, 1985; Ross & Sicoly, 1979), which focuses on individuals' egocentric preferences for maximizing their own utility relative to others. By placing the study of individual judgments into the contexts in which people interact, we get deeper insight into socially situated decision-making processes.

Although OB is interested in the factors that affect our assessments of what is just, and the consequences of those judgments on organizational outcomes, it also goes beyond the individual level of analysis to consider the effects of social context on justice assessments, as well as the implications of those assessments for organizational systems. This literature therefore considers the effect of the perceiver's relative position within the social system on his judgments (e.g., Kray & Lind, 2002; Mossholder, Bennett, & Martin, 1998). It also considers the implications of individual perceptions of fairness on the social system itself (e.g., Ambrose, Seabright, & Schminke, 2002; Greenberg, 1993; Johnson, Korsgaard, & Sapienza, 2002; Sapienza & Korsgaard, 1996). We deal with both of these topics, paying special attention to the negative consequences of procedural justice assessments in the workplace: sabotage and revenge.

Assessing Procedural Justice

Because most OB research is situated within organizational contexts, it is not surprising that the implications of the organization as a social context are primary to OB studies of procedural justice assessments (Cropanzano & Greenberg, 1997). Inasmuch as organizational structure shapes patterns of interaction and relative participation in decision making, it has clear implications for procedural justice assessments. Actors in highly centralized, vertically complex organizations, where individuals participate less in decision making and face more levels of hierarchy, generally feel a lower sense of distributive and procedural fairness (Schminke, Ambrose, & Cropanzano, 2000; Schminke, Cropanzano, & Rupp, 2002). Similarly, actors within larger and more formalized organizations—those marked by lower levels of social integration and fewer relationships based on intimate personal contact, and where decision making may be more political—also feel less distributive and procedural fairness (Randall, Cropanzano, Bormann, & Birjulin, 1999; Schminke et al., 2000, 2002). These effects are not felt across all levels of an organization, however. Because more control and higher returns accrue to those in higher echelons, higher-ranking actors feel a greater sense of justice than do those in the lower ranks. Likewise, those at higher levels are more forgiving of structural impediments to fairness, so that rank moderates the effects of structure on assessments of justice (Schminke et al., 2002).

Within a given organizational setting, actors' assessments of justice and fairness are likely to be often influenced by those with whom they interact on a daily basis. Thus organizational actors' assessments of justice are moderated by both the context in which those assessments are made as well as their relationships with the other actors involved in the event or process being scrutinized (Cropanzano & Greenberg, 1997). Coworkers' assessments of procedural and distributive justice often influence individuals' own assessments (Folger, Rosenfield, Grove, & Corkran, 1979; Steil, 1983), just as the treatment of other employees during the process of being laid off affects survivors' assessments of fairness (Brockner & Greenberg, 1990; Brockner, Grover, Reed, DeWitt, & O'Malley, 1987; Brockner et al., 1994), perhaps because actors focus on the potential implications for themselves of processes they see affecting others (James & Cropanzano, 1990).

In contrast, assessments are often based, at least in part, on prior personal experience (Lind, Kray, & Thompson, 1998; Lind & Tyler, 1988). Consequently, organizational actors are more likely to feel empathy when they have experienced injustice themselves (Kray & Lind, 2002). Similarly, the degree to which an actor has interacted with those involved in the focal event, particularly the authority figure perceived to be the source of the injustice, moderates perceptions of fairness (Folger & Cropanzano, 2001; Kray & Lind, 2002; Lind, Kray, & Thompson, 2001; Van den Bos, Lind, & Wilke, 2001).

Because much work within organizations focuses on the work group, the dynamics of interaction within such small groups have significant consequences for justice assessments. When organizational actors identify strongly with their work unit, their individual reactions are often filtered through group identification. This effect can be so strong that when authority figures violate neutrality or trust relative to individual team members, other members of the team perceive those actions as violations against the entire work group (Tyler & Lind, 1992). Similarly, perceptions of decision fairness are more associated with the degree to which the decision represents the interests of all group members than of individual group members (Miller, Jackson, Mueller, & Schersching, 1987). Consequently, members of work units are likely to share assessments of justice, affecting other types of employee attitudes, as well: work-unit-level procedural justice assessments explain more variance in individual job satisfaction than do individual perceptions of justice (Mossholder et al., 1998).

Consequences of Procedural Justice Assessments

Assessments of procedural justice at the group level affect the level of individual, self-reported job satisfaction.

Job satisfaction is not the only motivational outcome affected by perceptions of fairness, however; procedural and distributive justice assessments are relevant to an actor's evaluations of many and varied aspects of his organizational environment. Distributive justice, for example, accounts for more variance in satisfaction with compensation than does procedural justice (Folger & Konovsky, 1989). In contrast, perceptions of fairness regarding the performance appraisal process have a direct impact on an actor's trust in his supervisor, whereas perceptions of fairness regarding pay determination, performance appraisal and communication regarding compensation directly impact an actor's commitment to his or her employing organization (Folger & Konovsky, 1989; Scarpello & Jones, 1996). Similarly, the interaction of distributive and procedural justice moderate the effects of each, such that high procedural justice assessments reduce the effects of distributive justice on employee organizational commitment but heighten the effect of distributive justice on employee self-evaluation or self-esteem, implying that actors use information on procedural fairness to make self-attributions for their own outcomes (Brockner et al., 2003).

At the intraorganizational level, Kim and Mauborgne (1991, 1993a, 1993b, 1995, 1996, 1998) find that managers of foreign subsidiaries judge the strategic decision-making process to be fair when bilateral communication exists between the head office and subsidiary involved; subsidiaries are able to challenge and refute the strategic views of managers in the head office; managers in the head office are well informed regarding local conditions of the subsidiaries; subsidiaries are given a full account of the strategic decisions of the head office; and the head office applies decision-making procedures consistently across subsidiaries (Kim & Mauborgne, 1991, 1993b). Not only do these factors influence subsidiary managers' perceptions of procedural justice, they also increase subsidiary managers' willingness to comply (Kim & Mauborgne, 1993a), as well as to engage in extra-role behavior (Kim & Mauborgne, 1996).

Sapienza and Korsgaard (1996) find that the same cognitive and motivational processes that operate in individual perceptions of justice are also at work in interorganizational relationships. Because entrepreneurs cannot build high-potential businesses without the financing and participation of outside investors, harmonious entrepreneur–investor relations are paramount (Bygrave & Timmons, 1992; Larson, 1992). When entrepreneurs provide timely feedback, reducing the need for investor vigilance, investors' perceptions of procedural justice increase, particularly when investors' influence over organizational decisions is relatively low (Sapienza & Korsgaard, 1996).

*The Dark Side: Negative Assessments
of Procedural Justice*

Although organizational justice research has long identified revenge or retaliation as a means of restoring justice (e.g., Adams, 1965), only recently have OB researchers sought to test this assertion empirically. These studies demonstrate the moderating mechanisms that determine whether and how retaliation will be delivered, as well as the importance of interpersonal interaction in assessments of justice. Perhaps the most important question regarding the dark side of justice is *when* actors will take revenge. Retaliation appears to depend on the assignment of blame, and the type and degree of perceived injustice. When they assign blame, actors are more likely to take revenge against the perceived offender and less likely to reconcile with the perceived offender (Aquino, Tripp, & Bies, 2001). Sabotage is a specific type of retaliation in which a person actively tries to destroy another person's character, property, personal effects, and so on. The type of injustice actors feel also determines the type of sabotage they engage in and their goals in doing so: Those who feel distributive injustice are more likely to engage in sabotage to restore equity, whereas sabotage resulting from interpersonal injustice aims to retaliate against the transgressors. Although interpersonal injustices engender the most severe forms of sabotage or revenge, severity increases with perceived distributive, procedural, and interactional injustice (Ambrose et al., 2002; Greenberg, 1993).

In addition to the perceived source of injustice, the target of revenge affects the type and degree of retaliatory behavior in which an actor will engage. In general, wronged actors seem to use revenge to restore their feelings of justice, not to perpetuate the cycle of injustice. Thus the target of sabotage is most often the source of the perceived injustice, such that retaliatory action is seldom taken against individuals when the roots of the injustice are perceived to be structural (Ambrose et al., 2002; Greenberg, 2002). This reluctance is partially in the interests of self-preservation; low-status actors are less likely to take revenge on higher-status actors than are high-status actors on lower-status actors, although the tendency to take revenge decreases with the victim's absolute status (Aquino et al., 2001). Similarly, revenge is judged less harshly when its consequences are symmetric with the instigating injustice, whereas judgments are harsher when the method of revenge is similar to that by which the initial injustice was enacted (Tripp, Bies, & Aquino, 2002). Because revenge is often visible and subject to third-party evaluation, victims focus on aesthetic principles in deciding when and how to retaliate.

Both assessments of justice and fairness and the consequences of those reactions are highly salient to organizational functioning and outcomes. By supplementing the findings of laboratory research with participant observation, interviews, and surveys, organizational behavior is able to understand the interaction of these phenomena with other organizational processes and relationships much more deeply than it could if study were limited to the laboratory. The same is true of organizational behavior's treatment of the related topic of leadership; by moving the study of leadership, and particularly its consequences, out of the laboratory and into the field, we get a much clearer picture of how and why leadership matters.

Leadership

Perspectives on leadership in organizational behavior draw heavily on social psychology and industrial and or-

ganizational psychology. In OB, the focus of leadership studies is the individual actor within the organizational context, rather than individual differences. Therefore, we ultimately look to a discussion of social capital and networks and organizational power and influence—topics we address systematically in subsequent sections—to understand leaders in their organizations fully. In this section, however, we focus primarily on individual leadership styles and their consequences. We do not take up the age-old question of whether leaders are born or made (see Bryman, 1996, for a summary), but we outline three or streams of research on leadership within organizational behavior: leadership style, the consequences or impact of leadership, and leadership biases and blind spots.

Leadership Style

The study of styles of leadership in organizational behavior can be traced to Lewin, Lippitt, and White's (1939) studies of children's reactions to different leaders' approaches. Since then, research has gone in many different directions, from social-exchange theories to individual trait theories (see Judge, Bono, Ilies, & Gerhardt, 2002, for a review). OB researchers generally leave these issues aside, however, to focus primarily on how leadership shapes subordinate performance and satisfaction. Pearce and Sims (2002), building on Manz and Sims's (1991) and Bass and Avolio's (1993) models of leadership behavior, provide a model of five primary leadership types: directive leadership, aversive leadership, transactional leadership, transformational leadership, and empowering leadership. Here we focus on the latter three.

Transactional leadership can be understood as developing and fulfilling a contract with subordinates. Grounded in expectancy theory (Vroom, 1964), equity and exchange theory (Adams, 1963; Homans, 1958, 1961), and reinforcement theory (Luthans & Kreitner, 1985; Scott & Podsakoff, 1982), transactional leadership is associated with personal and material rewards conditional on performance, providing clarification and feedback relating to goals and managing by exception. Because expectancy theory suggests that individuals attempt to maximize expected returns to performance, transactional leadership focuses on clarifying the effort–reward relationship to maximize motivation, giving rewards equitable to input, and motivating higher subordinate input through the provision of higher rewards. Finally, transactional leadership emphasizes reinforcement through rewarding desired behavior.

Transformational leadership, in contrast, highlights the benefits of inspiration, charisma and overarching values in maximizing subordinate performance and satisfaction (Bass, 1985; Bass & Avolio, 1990). Transformational leaders elevate their subordinates' goals and give them confidence to exceed leaders' expectations. This tradition is rooted in charismatic leadership (House, 1977) and transforming leadership (Burns, 1978) and suggests that managers engage in impression management, articulate ideological goals, define subordinate roles in terms of ideological values, engage in role modeling, communicate high expectations and confidence in subordinates,

and arouse appropriate follower motives. Transformational leaders provide vision, express idealism, communicate inspiration and high performance expectations, challenge the status quo, and set high performance expectations (Bass, 1998). Although it aims to help subordinates realize their potential, subordinates often see transformational leaders as exceptional and thus become dependent on the leader for inspiration and guidance (Yukl, 1998).

Building on transformational leadership and the "new leadership approach" (Bryman, 1996), Manz and Sims (1980, 1987, 1989, 1991, 2001) describe empowering leadership. This leadership style involves developing follower self-leadership and self-management skills, and can be called "SuperLeadership" or "leading others to lead themselves" (Sims & Lorenzi, 1992, p. 295). The roots of empowering leadership lie in behavioral self-management (Thorenson & Mahoney, 1974), social cognitive theory (Bandura, 1986), participative goal setting (Erez & Arad, 1986; Locke & Latham, 1990), and cognitive-behavioral modification research (Meichenbaum, 1977). Fundamentally, this school holds that by modeling self-leadership behavior, leaders induce their subordinates to become similar self-leaders. Thus empowering leaders often reconceptualize obstacles as opportunities for learning (Manz & Sims, 1989, 2001) and set specific and difficult goals to stimulate performance (Locke & Latham, 1990). They emphasize independent action, teamwork, self-development, and self-reward and encourage participative goal setting, transforming followers into leaders (Bryman, 1996).

Consequences of Leadership

Although some OB scholars argue that organizational outcomes are marginally, if at all, affected by leadership (House & Aditya, 1997; Yukl, 1999), there is evidence of a causal link. Using archival statistics for professional baseball, Kahn (1993) finds that there is a strong correlation between the quality of managers' transactional leadership, team success, and player performance. Not only do higher-quality managers—measured by predicted pay relative to salary based on overall market trends—achieve better winning percentages, controlling for player skills, but also, players tend to perform better relative to prior performance the higher the manager's quality (Kahn, 1993). This study also suggests that transactional leadership, although often treated as problematic by OB scholars, can be quite consequential in certain contexts; further research regarding the limits of transactional leadership may therefore be merited.

Similar findings emerge not only for professional sports but also in more typical organizational settings. Transformational leadership, in particular, has a positive impact on subordinate performance. In a field study in which army officers were given training in either exclusively transformational or a variety of leadership techniques, transformational leaders had a more positive impact on direct follower development and indirect follower performance than did the control group (Dvir, Eden, Avolio, & Shamir, 2002). Although transforma-

tional leadership leads subordinates to identify with and become dependent on their leaders, it also leads to subordinate empowerment in terms of individual and collective self-efficacy and organization-based self-esteem (Kark, Shamir, & Chen, 2003); this relationship is partially mediated by subordinates' social identification with the group.

The relationship between transformational leadership and organizational outcomes seems to be linked to subordinate self-concordance, or the degree to which subordinates' own goals and values are aligned with those articulated by the leader (Bono & Judge, 2003; Sheldon & Elliot, 1999). By presenting work in terms of values and ideology that transcend individual interests, transformational leadership leads subordinates to internalize their leader's and group's goals, giving them internalized motivation, which is associated with more extra-role behavior and lower intention to leave (Bono & Judge, 2003; O'Reilly & Chatman, 1986). Employees see their work as more self-expressive, thus consequently perform better, and are more motivated, committed to their organizations, and satisfied with their leaders (Bono & Judge, 2003).

Even transformational leadership does not improve organizational outcomes under all conditions. For example, CEOs with transformational leadership styles improve organizational performance only when actors across the organization perceive environmental uncertainty (Waldman, Ramirez, House, & Puranam, 2001). Uncertainty appears to make actors more receptive to charismatic leadership and gives CEOs more discretion in their actions and behavior toward subordinates, with consequences for firm profitability, controlling for external influences. CEOs with transactional leadership styles, in contrast, do not improve firm performance, even under uncertainty (Waldman et al., 2001). Similarly, although transformational leadership is associated with improved future performance, it does not have a relationship to prior firm performance, suggesting a strong causal relationship between leadership and future performance (Waldman, Javidan, & Varella, 2004).

Leadership is clearly critical to the success of self-leading or self-managing teams (Barry, 1991). Often without a formal leadership position, self-managing teams may flounder if team members have previously experienced only aversive, directive, or transactional leaders in the past, as may those who have become dependent on transformational leaders (Bryman, 1996). Self-leadership is positively related to performance when teams are engaged in conceptual tasks and when team member interdependence is either very high or very low; in contrast, self-leading teams engaged in productive or behavioral tasks increase performance at moderate levels of team member interdependence (Stewart & Barrick, 2000).

Because empowered teams are almost forced to share leadership, distributing leadership roles and influence across team members (Pearce, Perry, & Sims, 2001; Yukl, 1998), they are better equipped to self-manage. In such shared leadership teams, the agents and targets of influence are peers. Pearce and Sims (2002) find that shared leadership is more closely associated with team effective-

ness than vertical leadership, particularly in teams with transformational and empowerment leadership styles. Shared leadership is a significant predictor of self-managing teams, particularly when engaged in complex tasks (Pearce & Sims, 2002). Self-managing teams are most successful when they have external leaders who span boundaries to build relationships, interact with outside constituencies, build internal and external support, and empower their teams (Druskat & Wheeler, 2003). External leaders of self-managing teams are less effective, however, when engaging in traditional leadership behaviors (Manz & Sims, 1987).

Despite evidence to the contrary, some OB researchers argue that the impact of leaders on organizational outcomes, particularly unexpected success, is overemphasized or "romanticized" (Meindl & Ehrlich, 1987; Meindl, Ehrlich, & Dukerich, 1985). Because individuals need to comprehend causally ambiguous, yet significant, organizational events, to attribute outcomes to human causes, and to validate their belief that organizations are effective, they give credit for unexpectedly positive outcomes to leadership (Meindl et al., 1985). This suggests that the value of leadership lies in the ability to control the meanings and interpretations that important stakeholders assign to events. This line of research points to an important consideration: that leadership may comprise both real effects, as indicated previously, as well as the interpretational effects suggested here.

Leadership and Biases

Because leaders are in both formal and informal positions of power and control, they may rely excessively on heuristics and mental models and may thus be subject to distorted perceptions, particularly in regard to their subordinates. The approach/inhibition theory of power (Keltner, Gruenfeld, & Anderson, 2003) argues that the experience of control increases the tendency to approach and decreases the tendency to inhibit. Because they experience subjective feelings of power, leaders are likely to feel more positive and less negative emotion, and are thus more attuned to rewards than threats (Anderson & Berdahl, 2002), suggesting that leaders may neglect subordinates' grievances. Similarly, leaders are likely to pay less attention to subordinates and thus more susceptible to stereotyping (Fiske, 1993). Moreover, those experiencing power are more likely to act, regardless of consequences (Galinsky, Gruenfeld, & Magee, 2003).

Those evaluating leaders may also show systematic biases, which are investigated in field studies and surveys by industrial and organizational psychologists and OB researchers. Implicit leadership theory holds that, over time and through personal perceptions and expectations, employees develop implicit leadership theories (ILTs), or personal assumptions about the skills, abilities, and personal characteristics of an ideal business leader (Kenney, Schwartz-Kenney, & Blascovich, 1996) that are stored and activated when they interact with people in a leadership position. Unless actors are perceived as conforming to those characteristics, they will not be per-

ceived as leaders (House & Aditya, 1997). Because ILTs are relatively constant across individuals, work contexts, and personal characteristics, as well as over time, not conforming to ILTs (Epitropaki & Martin, 2004) may hinder a leader's effectiveness.

Women leaders are particularly affected by biased attributions. Although evaluations of women leaders overall are only slightly less favorable than those of male leaders, meta-analysis shows that women tend to suffer more significant devaluation when their leadership styles are perceived as masculine (Eagly, Makhijani, & Klonsky, 1992). Similarly, women experience more devaluation when they occupy male-dominated roles and when their evaluators are men (Eagly et al., 1992). Women are evaluated more negatively than men even when their behavior meets the needs of a leadership role because of perceived incongruity of leadership and female gender roles, making it more difficult for women to achieve success as leaders than for men (Eagly & Karau, 2002). Evaluations of women are likely misguided, however, as meta-analysis shows that women display more transformational leadership behaviors positively related to effectiveness and fewer transactional leadership behaviors than do men (Eagly, Johannesen-Schmidt, & van Engen, 2003).

The study of leadership—similar to the study of decision making and justice and fairness—is clearly grounded in social psychology. Nevertheless, it often seeks to understand processes and consequences particular to organizational contexts, which cannot be replicated exactly in the laboratory. By measuring and controlling for outside influences, these studies are able to isolate and analyze the unique contributions and effects of leadership on organizational outcomes.

ORGANIZATIONAL ACTOR AS TEAM PLAYER

Within OB, teams are analyzed at several levels, including the nature of the intrateam relationships (also referred to as group dynamics) as well as the relationship between the team and the rest of the organization (commonly referred to as social networks).

Groups and Teams

It is nearly impossible to talk about organizations without mentioning or understanding teams. Not surprisingly, OB has borrowed and built on social psychology's groundbreaking studies of groups. Most of the research that OB has leveraged has been prior to 1990.

Many scholars (and certainly people in the business world) use groups and teams interchangeably when they should not. A *group* is an aggregate or collection, whose participants may or may not have a (shared) goal and may or may not be interdependent; a *team* is a group of people who share a goal and are interdependent. Thus, a team is a specific type of group—ideally composed of fewer than 10 people—working toward a shared goal that has organizational relevance (Hackman, 1990). The social psychology of groups is largely descriptive; the orga-

nizational study of teams is both descriptive and prescriptive. Key OB research on groups that derives from social psychology but has been claimed by OB includes transactive memory, the common information effect, minority influence, and conflict.

Transactive Memory

A transactive memory system is a group-level information-processing system that is an extension of the individual information-processing system (Wegner, 1986; Wegner, Giuliano, & Hertel, 1985). Whereas Wegner (1986) first measured and identified transactive memory systems in couples, Moreland, Argote, and Krishnan (1998) extended the concept to teamwork. In their investigation, teams were asked to assemble AM radios as part of a training experience (Moreland, Argote, & Krishnan, 1996). Teams were randomly assigned to either individual training sessions or group training sessions. In the training phase, everyone—both individuals and groups—received the same information; groups were not given any explicit instructions on how they should organize themselves. One week later, the participants were asked to assemble the radios again without the benefit of written instructions. The groups that had trained as a team performed better, in terms of successfully completing the assembly and with fewer errors. Most notably, the transactive memory effect has been examined in actual work teams, including hip and knee replacement surgery teams (doctors who have performed surgery as a team have patients with faster recoveries and fewer complications; Moreland & Argote, 2003) and coal-mining injuries and deaths (Goodman & Garber, 1988).

Common Information Effect

The common information effect, or the tendency for groups to discuss information that is common to everyone, as opposed to information that is unique to individuals in the team, has been demonstrated with a variety of paradigms and group tasks (cf. Christensen et al., 1998; Gigone & Hastie, 1996; Stasser & Titus, 1985). The common information effect is regarded to be dysfunctional for groups and, in most investigations, the optimal group answer can only emerge if group members are fully sharing all information. OB researchers quickly focused on two critical implications of the common information effect: practical solutions and the interface between the type of information and the type of team members (Phillips, Mannix, Neale, & Gruenfeld, 2004). Practical solutions to avoid excessive focus on common information have included instructing groups to rank rather than choose (Gigone & Hastie, 1996; Hollingshead, 1996), spend an equal amount of time discussing each alternative (Larson, Foster-Fishman, & Keys, 1994), define the goal as a "problem to be solved" rather than a "judgment to be made" (Laughlin, 1980; Stasser & Stewart, 1992), suspend initial judgment (Schulz-Hardt, Frey, Lüthgens, & Moscovici, 2000), and record facts during the discussion that justify their decision (Larson, Christensen, Franz, & Abbott, 1998).

Phillips and her colleagues (Gruenfeld, Mannix, Williams, & Neale, 1996; Phillips, 2003; Phillips & Loyd, 2004; Phillips, Mannix, et al., 2004) have examined not only how the distribution of information affects the likelihood that members will consider information but also the way that friendships among group members impact the common information effect. In this way, Phillips has linked the cognitive-information-processing aspects of teamwork with the sociorelational aspects. For example, team members who are familiar with one another are less likely to make poor decisions resulting from the common knowledge problem than are teams whose members are unacquainted (Gruenfeld et al., 1996). According to Gruenfeld and colleagues (1996), this result stems from the fact that teams that are more familiar with one another tend to pool their unique information to find the best choice, while unfamiliar teams tend to aggregate their preferences and adopt the choice of the majority; this suggests that familiarity prevents teams from adopting dysfunctional responses to normative influence. Moreover, Phillips and colleagues have shown that when the distribution of unique versus shared information is congruent with the relationship status of group members, such that the unfamiliar team member possesses the unique information and the previously acquainted team members share information, teams are more likely to discuss unique and shared information and exhibit superior performance (Phillips, 2003; Phillips, Mannix, et al., 2004). Phillips, Mannix, and colleagues (2004) found that when a familiar member who was in the numerical majority held unique information, groups performed worse than when that information belonged to a stranger. Likewise, Phillips (2003) argued that it is members' assumptions of similarity (e.g., Allen & Wilder, 1975) that lead them to alter their behavior when they find that they do not agree with whom they expect to agree. For instance, individuals express greater surprise and irritation with a differing perspective when it is expressed by a socially similar other than when it is expressed by someone who is socially different (Phillips, 2003; Phillips & Loyd, 2004). Thus, the sharing of unique information and performance of groups that face hidden profile situations is also affected by who agrees with whom in the group, which is frequently a function of the social relationships among group members and the decision-making strategies they employ as a consequence.

Minority Influence

Minority influence focuses on the extent to which people that hold nondominant opinions in a group can influence a group's ultimate decision. The mere presence of a minority stimulates more thorough information and more thoughtful group decision making (Gruenfeld, 1995). Minority opinions can arise from one of two sources—from a member of one's own work team (i.e., an ingroup member) or a member of another group (an outgroup member). Both can be effective; however, a minority opinion offered by an ingroup member is often more influential (David & Turner, 1996). People in work teams are often reluctant to express divergent views and

disagreement with outgroup members is more tolerable than disagreement with ingroup members (Phillips, 2003).

Intensive examinations of actual organizations reveal that members in the majority exhibit greater integrative complexity than do members in the minority (Gruenfeld, 1995). For example, the authors of majority opinions in the U.S. Supreme Court who are exposed to members who hold a minority point of view show increases in their own levels of integrative thinking (Gruenfeld, 1995). In contrast, people exposed to majority opinions or unanimous groups actually experience a decrease in integrative thinking (Gruenfeld, Thomas-Hunt, & Kim, 1998).

One type of minority influence hinges on the length of time a person has been a member of a group. Because reorganizations are so prevalent in businesses, OB researchers have been particularly interested in how the changing composition of a group affects its productivity. Gruenfeld, Martorana, and Fan (2000) investigated the consequences of temporary membership changes for itinerant members (i.e., members who leave their core group to visit a new work group, and then subsequently return) and indigenous members of those foreign and native groups. Although it would seem that itinerant members would learn new ideas that would transfer back to their native group once they returned, this was not observed. Members of all groups produced more unique ideas after itinerant members returned to their native group than before they left or during the temporary change period. However, indigenous group members were less likely to want to include them in a group project designed to draw on knowledge of the work team. Moreover, indigenous group members regarded the itinerant members to be more argumentative than they were before their new assignment, and their contributions were perceived as less valuable. Thus, itinerant group members had less direct influence after their reorganization than before. In this sense, groups benefit from their "worldly" members, but they do not appreciate them.

Conflict

The study of conflict in teams has been almost exclusively the domain of management theorists, such as Jehn (1995; Jehn & Mannix, 2001; Shah & Jehn, 1993; see also Amason, 1996; Guetzkow & Gyr, 1954). Jehn (1995) proposed that three types of conflict occur in teams: relationship conflict (also known as emotional conflict, A-type conflict), task conflict (conflict about the substance of an issue or project), and process conflict (conflict about how to do a project). Many research investigations measured the types of conflict within laboratory teams as well as actual work teams; several investigations reported that relationship conflict was negatively related to task performance, whereas task conflict was positively related to task performance (cf. Jehn, 1997). Recently, De Dreu and Weingart (2003) performed a meta-analysis of the three types of conflict; in all cases, both relational as well as task conflict was negatively related to group perfor-

mance. Behfar, Peterson, Mannix, and Trochim (2004) examined 51 work teams and found that three factors were associated with enhanced performance: members focus on content, rather than style; members assign work to others in the team based on expertise, not convenience; and members communicate about factors that affect their ability to carry out assignments, something that Behfar and colleagues call communicating about work and quirks.

Brainstorming and Creativity

Research on team brainstorming and creativity has a strong hold in psychology as well as management science. Whereas the social psychological approach to brainstorming focuses nearly exclusively on the question of why groups are inferior to individuals (or nominal groups; for a review, see Paulus, Brown, & Ortega, 1999; see also Diehl & Stroebe, 1987), the OB approach assumes that teams must be creative and consequently concerns itself with improving the effectiveness of intact brainstorming teams (cf. Sutton & Hargadon, 1996). Three subareas of research have been remarkably influential in this area: intrinsic versus extrinsic motivation, psychological safety, and boundary spanning.

PSYCHOLOGICAL SAFETY

Edmondson and Mogelof (2006; see also Edmondson, 1999) explain how the presence of psychological safety in a team provides a critical foundation for creativity. Psychological safety encompasses the taken-for-granted beliefs that others will respond positively when a person takes a risk. Because the creative process involves divergent thinking, it is a risky behavior. Edmondson and Mogelof investigated several project teams and measured psychological safety at key points in time during a project. Importantly, individual differences in terms of two of the "big five" personality factors—neuroticism and openness to experience—were positively associated with greater psychological safety in teams.

CREATIVITY AND BOUNDARY SPANNING

According to Ancona (1990; Ancona & Caldwell, 1988, 1990, 1992a), for teams to be creative, they need to navigate organizational boundaries and divides. Ancona and Caldwell's (1992a) groundbreaking investigation of 45 teams reveals four prototypical patterns: ambassadorialism, task coordination, scouting, and isolationism. Ancona and Bresman (2006) suggest that creative teams, called, X-teams, "beg, borrow, and steal" best practices across the company. They distinguish the process of creative idea generation from the organizational tightrope of winning support for ideas and ultimately receiving organizational buyin. They argue that the most innovative teams have a marked disrespect for traditional teams and even organizational boundaries, and that this fluidity drives the innovative process. They identify five key components of X-teams with respect to the innovative process: external activity, extensive (network) ties, expand-

able tiers (managing up and out), flexible membership, and coordination among ties.

Hargadon (2006) argues that creativity involves two complementary but seemingly opposing processes: bridging and building. Bridging, according to Hargadon, requires that two previously distinct worlds or domains be brought together via pattern recognition or making a new connection. To this end, Hargadon provides compelling case examples from physics to rock and roll. The building process, according to Hargadon, requires that new patterns be built involving both understanding and action within those social groups that serve as the arbiters of the creative output. Hargadon bases his theory in the microsociology of creativity.

Negotiations

Negotiation research is the child of at least three literatures: Morton Deutsch's (1973) seminal work in social psychology; Raiffa's (1982) work in applied game theory; and, oddly enough, Kahneman and Tversky's (1979) work in decision making. Negotiation takes front and center stage in OB, in part because organizational actors routinely find themselves involved in intra- and interorganizational conflict and exchange. Modern negotiation research in OB has dominated the micro-OB journals for over two decades, beginning in the 1980s when Bazerman and Neale redefined the study of negotiation as a joint decision-making process that contained the same flaws as individual decision making. We outline three central themes in negotiation research: judgment and decision making; relationships and reputation; and social perception and social identity. These themes are highly interrelated, but we distinguish them here to focus on how cognitive and motivational processes affect decision making in negotiation; how social relationships affect negotiation; and, finally, how subtle changes in context can affect negotiated outcomes.

Judgment and Decision Making

The great majority of research in negotiation focuses on the shortcomings that befall negotiators. The key finding is that negotiators are rather poor at negotiating. In one investigation, more than 40% of negotiators failed to realize when their interests were completely compatible (i.e., aligned) with others, and nearly 25% of these negotiators reached lose–lose outcomes (Thompson & Hastie, 1990; Thompson & Hrebec, 1996). Negotiation biases fall into biases that are general in nature and negotiation-specific biases that include the framing effect, anchoring and adjustment, and overconfidence (for reviews, see Neale & Bazerman, 1991). For example, several investigations have examined the impact of framing (gain vs. loss frame) on negotiation outcomes. The general finding is that negotiators who are led to adopt a gain frame are more likely to reach settlement than are negotiators who are led to adopt a loss frame (cf. Bottom & Studt, 1993; Neale & Bazerman, 1991).

The second set of biases is negotiation specific, such as the fixed-pie perception (Bazerman & Neale, 1983;

Thompson & Hastie, 1990), coercion bias (Rothbart & Hallmark, 1988); attribution bias (Morris & Larrick, 1995; Morris, Larrick, & Su, 1999), and concession aversion (or the grass-is-greener bias; Ross & Stillinger, 1991). Perhaps of all the negotiation-specific biases, it is the fixed-pie perception that has garnered the most attention. The fixed-pie perception, coined by Fisher and Ury (1983), refers to the often faulty belief that people's interests are fundamentally and directly opposed, such that a gain for one party represents a loss for the other party. Numerous research programs have established the presence and persistence of the fixed-pie perception in producing suboptimal negotiation agreements, preventing negotiators from reaching mutually beneficial agreements (cf. Bazerman & Neale, 1983). The implications of the fixed-pie perception for the quality of negotiated outcomes are serious; when negotiators believe that their interests are completely opposed, they cannot dovetail their interests.

Another well-documented general bias that has been studied in context of negotiations is overconfidence. Overconfidence refers to unwarranted levels of confidence in people's judgments of their abilities and the occurrence of positive events and underestimates of the likelihood of negative events. For example, in negotiations involving third-party dispute resolution, negotiators on both sides of the table believe that a "neutral" third party will adjudicate in their favor (Farber, 1981; Farber & Bazerman, 1986, 1989). Obviously, the third party cannot rule in favor of both parties. Similarly, in final-offer arbitration, wherein parties each submit their final offer to a third party, who then makes a binding decision between the two proposals, negotiators consistently overestimate the probability that a neutral arbitrator will choose their own offer (Bazerman & Neale, 1982; Neale & Bazerman, 1983). Whereas overconfidence about potential outcomes is a detriment to effective negotiation, evidence suggests that overconfidence about the value of the other party's walkaway point might serve the negotiator well. Specifically, negotiators who are optimistically biased (i.e., they think that their counterpart will concede more than he or she really can) have a distinct bargaining advantage (Bottom & Paese, 1999). Whereas negotiation research had initially been exclusively cognitive, the research tide turned and investigations of mood and emotion in negotiation grew in popularity (for a review, see Barry, 1999; Barry, Fulmer, & Goates, 2005). For example, several investigations have explored the implications of Forgas's (1995) affect infusion model on negotiated outcomes.

Relationships and Reputations

Somewhere in the late 1980s and early 1990s, a small rebellion took place against the overly cognized view of the negotiator. Rather than picturing him or her as a biased information processor, people began to focus on the context of the negotiation and the relationship among organizational actors. Pruitt and Rubin's (1986) dual-concern model, derived from Thomas and Kilman's (1974) concept of people and task focus, suggests that integrative agreements result when negotiators are concerned about themselves and the other party (Pruitt & Carnevale, 1993; Pruitt, Carnevale, Ben-Yoav, Nochajski, & Van Slyck, 1983). Several investigations have examined the impact of friendship and relationships on negotiator effectiveness (for a review, see Valley, Neale, & Mannix, 1995). Another aspect of relationships is the social-organizational network. Several investigations of networks in negotiations were conducted and, recently, the study of reputations in negotiation communities has become an important topic (Tinsley, O'Connor, & Sullivan, 2002). For example, Tinsley and colleagues (2002) find that negotiators who perform well in an organizational community but who then gain a reputation for being a tough negotiator encounter opponents who behave more aggressively toward them.

Glick and Croson (2001) argue that negotiators' reputations are built fairly quickly in negotiation communities and affect how others deal with them. They found that investigations of reputations in a community of MBA students yielded four profiles: liar-manipulators (who will do anything to gain advantage), tough but honest (negotiators who don't lie, but make very few concessions), nice and reasonable people (who make concessions), and cream puffs (who always make concessions). People act much tougher when dealing with someone who has a reputation for being a liar but are much more reasonable with other profiles. And, most notably, people are more likely to use tough strategies with liars for defensive reasons but might use tough strategies with cream puffs for opportunistic reasons.

Somewhat paradoxically, while people may expect that they will act tougher around those they expect to be tough, these same individuals also may be likely to cave in. Diekmann, Tenbrunsel, and Galinsky (2003) recently demonstrated that during negotiations, there can be a disconnect between individuals' intended reaction to a competitive opponent and how they actually behave. For example, MBA students reported that they would negotiate more aggressively with a competitive opponent relative to a noncompetitive opponent. However, when these same individuals conducted a negotiation 5 weeks later, the same participants who claimed they would be more aggressive when facing a competitive opponent set less aggressive expectations and ultimately agreed to worse outcomes than participants who believed they were facing a noncompetitive opponent. Why? When negotiators forecast how they believe they might behave, they fail to consider their motivation to reach an agreement.

The rise of the Internet era ushered in a new area of research on the impact of communication modality (face-to-face, telephone, email, etc.) on negotiation outcomes (for a review, see Nadler & Shestowsky, 2006). In general, it is believed that the richer the communication medium, the better the negotiated outcome (cf. Moore, Kurtzberg, Thompson, & Morris, 1999); however, negotiators who are not prepared or are in a relatively weak bargaining position are best advised to negotiate via email (Loewenstein, Morris, Chakravarti, Thompson, & Kopelman, 2005).

Social Perception and Social Identity

The classic study of social perception in negotiation is that of Kelley and Stahelsski (1970) who found that the outcomes of a multiround Prisoner's Dilemma Game depended on the parties' perceptions of one another as well as their social stereotypes in general. Kramer and Brewer (1984) used social identity theory to predict how negotiators would behave in a social dilemma task. They hypothesized that negotiators might be more cooperative if they shared a common identity. They examined the incidence of competitive and cooperative behavior in a social dilemma task. Although the stakes and payoffs were exactly the same across conditions, when group members were told to focus on a superordinate identity, they behaved more cooperatively (donated more to a common pool of resources) than when they were instructed to focus on their individual identities. In a similar sense, Ross and Ward (1996) found that people playing a Prisoner's Dilemma Game, titled "the Wall Street game," were much more competitive than people playing the same game when it was titled "the community game."

Morris and colleagues (1999) examined how people develop impressions of others in negotiations based on their behaviors. Morris and colleagues reasoned that negotiators would fall prey to the fundamental attribution error when attempting to explain the behavior of a demanding counterpart. The situational explanation for a demanding counterpart would attribute excessive demands to the quality of that person's alternatives (best alternative to negotiated agreement, or BATNA), whereas a dispositional explanation would invoke a personality attribution. Morris and colleagues found that negotiators with attractive options outside their negotiation were indeed labeled as more demanding (dispositional attribution) than negotiators who did not have as attractive outside options.

The tidal wave of unconscious processes in social psychology was not lost on OB research. Several investigations of below-conscious-awareness information processing and negotiator behavior have been conducted. For example, Kray, Thompson, and Galinsky (2001) applied stereotype threat to the study of negotiation. They found that when gender is "primed," women perform less well than do men (Kray, Thompson, & Galinsky, 2001). However, they also found evidence for a "stereotype reactance" effect, wherein when gender is explicitly primed, women are less likely to identify with traditional gender roles and perform even better than do males (Kray, Galinsky, & Thompson, 2002).

Research on cross-cultural negotiation differences has burgeoned in recent years (cf. Brett, 2001). One subarea within this research stream focuses on cultural differences as a personality difference among members of different cultures; another subarea in this research stream focuses on cultural differences that can be temporarily activated in people of different cultures. As an example of the first type of research, Brett and her colleagues (1998) find that direct information sharing about interests is not normative in collectivistic cultures; rather, collectivistic cultures share information indirectly, which

can be just as effective in reaching optimal outcomes. As an example of the second type of research, Seeley, Thompson, and Gardner (2003) temporarily primed negotiators to be either interdependent (similar to collectivistic) or independent (similar to individualistic) and found that interdependently negotiators are more cooperative in one-on-one negotiations, but more competitive in team-on-team negotiations.

Demography and Diversity

One of the contextual factors that influences work-team processes is the demography of the team itself, particularly the diversity of its composition. Not only does composition affect team processes and outcomes, it also affects the way team members feel about each other and their work. These effects are generally found to result from social categorization (Kramer, 1991), similarity/attraction (Byrne, 1971), and information sharing in decision making (Gruenfeld et al., 1996). The effects of work-team diversity are not limited to those created by gender or race but, rather, can stem from any characteristics used to identify others as different, which are often context specific (Williams & O'Reilly, 1998).

Organizations implement teams based on the assumption that decisions made by groups with diverse expertise and experience will be better than those made by individuals. Some, following the "value-in-diversity" school (Cox, Lobel, & McLeod, 1991), believe that diversity can improve organizational performance. Others, in contrast, show that diversity impedes group functioning (see Brewer, 1979; Messick & Massie, 1989) and argue that homogeneity might improve group performance, because similar people tend to like each other more than dissimilar people (Byrne, 1971) and undergo relatively little conflict in the course of completing a task. We review the effects on group process and performance of both task-oriented diversity (group tenure and background) and relations-oriented diversity (gender and race and ethnicity).

Task-Oriented Diversity

GROUP TENURE DIVERSITY

As argued by Pfeffer (1983, 1985), homogeneity in group tenure is generally found to lead to increased communication, social integration, and cohesion (Good & Nelson, 1971; Lott & Lott, 1965; O'Reilly, Caldwell, & Barnett, 1989). Social integration may be affected by heterogeneity in tenure indirectly, through decreased, less open communication and more distortion of messages (O'Reilly, Snyder, & Boothe, 1993; Smith et al., 1994). Similarly, tenure diversity is associated with increased conflict and political activity (O'Reilly, Williams, & Barsade, 1997; O'Reilly et al., 1993; Pelled, 1993).

These findings do not necessarily indicate that tenure diversity is detrimental to team performance. Heterogeneous groups are more likely to have access to nonredundant information because their networks do not overlap (Burt, 1997). Generally, diverse groups are

better able to identify and understand multiple external cues and thus may recognize subtle changes in task demands sooner than homogeneous group members (Thompson, 1967). Similarly, diverse teams may enjoy the benefit of diverse experience in attacking complex tasks (Ancona & Caldwell, 1992b; O'Reilly et al., 1997). In addition, the conflict engendered by diversity may lead to disagreements—such as those regarding minority opinions mentioned earlier—that improve decision making (Amason, 1996; Eisenhardt, Kahwajy, & Bourgeois, 1997; Pelled, 1996a). Nevertheless, heterogeneity in tenure increases turnover by creating conflict and poor social integration (O'Reilly et al., 1989, 1993; Pfeffer & O'Reilly, 1987). Together, these results suggest that tenure diversity is advantageous based on experience and perspective, but that groups may fail to realize these benefits.

FUNCTIONAL DIVERSITY

Diversity in background can refer to education, functional specialty, and job experience, indicating training in different expertise, abilities, and perspectives. Because of their access to diverse skill sets and unique information, often accessed through increased communication with those outside the team, cross-functional teams are assumed to be more effective than less diverse teams (Ancona & Caldwell, 1992b; Cummings, 2004; Jackson, 1992). As with tenure diversity, however, the advantages of cross-functional teams may be squandered through the increased conflict that often accompanies diversity, or by team members' unwillingness to share information held unevenly across the team (Wittenbaum & Stasser, 1996). As Gruenfeld and colleagues (1996) show, heterogeneous groups perform best when information is fully shared, whereas groups that are more familiar with each other perform better when information is unevenly distributed. In addition, because it leads to the surfacing of multiple points of view and task orientations, functional diversity also increases task conflict, which improves cognitive task performance (Jehn, Northcraft, & Neale, 1999; Pelled, 1993).

In fact, much OB research shows that teams composed of actors with diverse backgrounds are able to achieve superior performance (e.g., Barsade, Ward, Turner, & Sonnenfeld, 2000). Although functional diversity drives task conflict, because those with different specialized backgrounds are naturally more likely to disagree on process issues, this conflict leads them to think more deeply about task requirements, improving performance (Pelled, Eisenhardt, & Zin, 1999). The mere fact of recognizing that diverse actors have specialized knowledge allows these groups to explore divergent and innovative ideas (Janssens & Brett, 2003). Similarly, because they look to different external cues, functionally diverse teams respond better to environmental change (Keck & Tushman, 1993; Murmann & Tushman, 1997). Moreover, intrapersonal functional diversity improves information sharing and decision quality, as team members are more willing to accept others' functional expertise (Bunderson & Sutcliffe, 2002). Similarly, structurally

diverse work groups—whose members fill different roles and affiliations—achieve superior performance by exposing the group to unique knowledge and skills (Cummings, 2004). Thus, functional diversity improves performance by increasing the availability of task-relevant knowledge, leading to greater task conflict and information sharing.

Demographic Diversity

GENDER DIVERSITY

Gender is perhaps the most studied, yet least conclusive, aspect of work-group diversity. Gender is a particularly salient dimension of identification and categorization, with immediate consequences for workplace interaction. For example, Eagly's gender role theory (see Eagly et al., 1992) holds that actors develop expectations about their own and others' behavior based on their beliefs about the behavior that is appropriate for men and women, and they may evaluate women negatively when their behavior differs significantly from expectations. This is thought to be more salient and have a stronger effect for women than for men, because women are generally in the minority in the workplace—particularly in leadership roles—which reinforces negative stereotypes (Eagly et al., 1992).

There is evidence of increased conflict and process loss in mixed-gender groups (e.g., Alagna, Reddy, & Collins, 1982; Pelled, 1996b; Sackett, DuBois, & Noe, 1991), although these findings are not conclusive (Williams & O'Reilly, 1998). Kanter (1977) notes that the effects of gender diversity on group process are moderated by the proportion of men and women in a given group, an issue not taken into account by most gender studies. Although gender diversity was thought to increase conflict, there is no evidence to support this hypothesis (O'Reilly et al., 1997; Pelled, 1996a; Pelled et al., 1999). Moreover, the effect of being different varies between men and women (Bacharach & Bamberger, 2004; South, Bonjean, Markham, & Corder, 1982; Tsui, Egan, & O'Reilly, 1992), partially because groups dominated by females tend to isolate male minority members less than male-dominated groups do female minority members (Fairhurst & Snavely, 1983; Schreiber, 1979). Similarly, the effect of exclusion of women from male-dominated networks is also inconclusive (Brass, 1985; Ibarra, 1992, 1997; Mehra, Kilduff, & Brass, 1996).

The findings regarding gender diversity and organizational performance are also mixed, with some studies finding positive effects on at least some measures of performance (Fenwick & Neal, 2001; Jackson, Joshi, & Erhardt, 2003; Pazy & Oron, 2001; Rentsch & Klimoski, 2001), and others finding no significant effects (Richard, 2000; Watson, Johnson, & Merritt, 1998). Gender diversity may decrease perceived productivity in work groups, partially as a result of increased emotional conflict (Pelled, 1996b). Similarly, male-dominated work groups demonstrate less prosocial organizational behavior, which is also related to increased emotional conflict (Kizilos, Pelled, & Cummings, 1996). Proportionality and

gender of the focal actor are also consequential to the effects of diversity on individual outcomes such as satisfaction, commitment, and turnover (Tsui et al., 1992). Men in mixed work teams are less satisfied and have lower self-esteem than do men in male- or female-dominated teams, whereas women are most satisfied in male-dominated teams and least satisfied in female-dominated teams (Wharton & Baron, 1987, 1991). Women are more likely to leave homogeneous groups than are men, despite women's expression of greater commitment, positive affect, and perceived cooperation when working in all-female teams (Chatman & O'Reilly, 2004). Nevertheless, women are more supportive of their peers and view female superiors more positively when management is more diverse, perhaps due to higher perceived chances for advancement (Cohen, Broshak, & Haveman, 1996; Ely, 1994).

RACIAL AND ETHNIC DIVERSITY

Given that racial and ethnic diversity is significantly less studied than gender diversity, it is not surprising that its effects are somewhat murky. There is evidence that ethnic minorities behave more cooperatively than do Anglo-Americans in Prisoner's Dilemma Games, indicating collectivism (Cox et al., 1991; Earley, 1989). Contradictory evidence also exists, indicating that ethnic minorities behave more competitively when in the minority on a given task than do Anglo-Americans, whereas both groups are equally cooperative when in the majority (Espinoza & Garza, 1985). There is no more conclusive evidence, unfortunately, regarding racial diversity and conflict; racial diversity is found to cause emotional conflict (Pelled, 1993; Pelled et al., 1999), although not when the sample studied includes a large proportion of minorities (O'Reilly et al., 1997; Pelled, 1996b). This may be due to expectations of congruence for ingroup members, as disagreement is less disruptive to group processes when expressed by minorities or perceived outgroups (Phillips, 2003; Phillips, Mannix, et al., 2004). As with gender, proportionality seems to influence the effect of racial diversity.

The effects of racial and ethnic diversity on group performance are also mixed. Evidence from class-related group activity indicates that racially diverse groups perform similarly to racially homogeneous groups, but that they may consider more perspectives and generate more alternatives (McLeod & Lobel, 1992; Watson, Kumar, & Michaelsen, 1993). The results of field studies indicate a positive relationship between racial diversity and group creativity and implementation (O'Reilly et al., 1997). One possible explanation for these findings is "aversive racism" (Gaertner & Dovido, 1986), whereby individuals attempt to overtly demonstrate their lack of prejudice to compensate in situations in which perceived prejudices may be salient. Many studies, however, find no effect or weakly negative effects of racial diversity on performance (Kizilos et al., 1996; Pelled, 1996b; Pelled et al., 1999). Regarding individual-level outcomes, those in the minority are generally less satisfied and less committed to the organization, although these effects are stronger for whites

than for ethnic minorities (Riordan & Shore, 1997; Tsui et al., 1992), and such asymmetries have not been found consistently (see Bacharach & Bamberger, 2004). In summary, it seems that finer-grained studies are needed to fully understand the effects of racial and gender diversity on group process and performance.

Our review of groups and teams, negotiations, and demography and diversity is uniquely influenced by the organizational setting and the application of fundamental social psychological principles to organizational contexts. In the next section, we move more explicitly to the study of individuals embedded within the structure of organizations to examine how the reciprocal interaction of actors and their environments is affected by social psychological processes and how this in turn affects the environments in which actors operate.

THE ACTOR EMBEDDED WITHIN THE ORGANIZATION

The area of OB that is most different from traditional social psychology deals with the organizational actor as fully embedded within the organization. Here the focus is still on the individual, but the context within which he or she operates, that constrains and enables the individual's actions, is brought into much sharper focus. The organization as an identifiable and salient entity is highly developed. The focus is still on interaction of individuals, as in the previous section. However, the structure of the environment and social system plays a greater role in shaping the interaction itself. And often, the interaction has a feedback effect, influencing the social structure and environment, in turn. Because of the complexity of relationships and interactions represented by these phenomena, they are studied primarily in the field, distinguishing them from traditional social psychological research. In this section, we focus on the critical topics of social capital and networks, and power.

Social Capital and Networks

Of all programs of OB research, the study of social capital and networks, or the structure of interorganizational relationships, most explicitly views the actor's behavior as embedded in and, more important, constrained by complex social relationships and systems, as it sets all individual experiences in the context of the larger social structure, which is beyond the control of individuals. This research looks at different patterns of behavior that result from differences in structure among social systems, as well as the individual actor's relative position within those structures. The implications of network structure on macro-organizational outcomes have received a great deal of attention in sociology and strategy, whereas research in OB has focused primarily on the consequences of network structure on individual action and behavior. In this section, we first review the foundational work in this area, then review more recent research on the consequences of network composition for individual actors, paying particular attention to minority

group members, as well as consequences for organizational outcomes.

Social Networks and Embeddedness

The concept of social networks and the methodologies used to analyze them have developed through macro-organizational and sociological research over the past 30 years, moving from highly mathematical models in the natural sciences and information science to the social sciences. Granovetter (1973, 1985) provides the primary point of departure, arguing that because all interaction takes place in or is "embedded" within a social context (i.e., does not occur between atomistic actors without histories or futures), understanding that context itself is critical to understanding social interaction. Inasmuch as social networks structure small-scale, highly local interactions, they also impact the social structure in which all interaction occurs.

Granovetter (1973) also differentiates weak from strong social ties, which can be thought of as interpersonal relationships or connections within a network. Drawing on Homans (1950), Granovetter argues that the strength of a social tie is a function of the amount of time, intimacy, emotional intensity, and reciprocity that characterize the relationship it represents; that is, the more time actors spend together, the stronger their feelings of friendship and connectedness to each other. Based on homophily, assumed similarity, and cognitive balance theory (Heider, 1958; Newcomb, 1961), the stronger the tie between two actors, the larger the proportion of ties with third parties they will have in common. Strong ties are valuable because they breed trust, openness and exchange, even in the absence of formal contracts (Granovetter, 1985). Weak ties are also valuable, however, because they can form bridges that link otherwise unconnected actors, thereby facilitating communication and diffusion across the network (Granovetter, 1973).

Strong ties may be more interesting, however, inasmuch as they engender social capital, a property of social relationships and structures that facilitates action by the actors within those structures that could not be achieved without the relationship (Coleman, 1988). That is, strong ties give individual actors access to resources that they can use to in pursuit of their own interests, and so help explain differences among individual outcomes, as well as differences among macrosocial outcomes. Similar to cohesion, social capital inheres within the structure of relations between and among actors, rather than the actors themselves. Benefits accrue to actors with social capital because strong ties give them access to valuable information channels, engender obligations and expectations of reciprocity and trustworthiness, and promote the development of norms and sanctions that make trust less risky (Coleman, 1988). The advantages of social capital inherent in strong ties grant actors access to the unique resources of others within the network, while insulating them from opportunism and exploitation. The more social capital one has, the better connected one is, and the more advantages and rewards one captures.

In contrast to Coleman's focus on the cohesive nature of social networks, Burt (1992, 1997) focuses on the opportunistic side of social networks: weak ties. Burt (1992) argues that social capital can be measured not only in terms of access to people with specific resources but also in terms of the social structure in its own right. Structural hole theory poses social capital as a function of brokerage opportunities in a network, or the ability of one actor to bridge others on opposite sides of a hole in the social network, with positive economic effects for the network. Brokerage gives the focal actor several advantages: timing benefits resulting from unique information regarding new opportunities created by needs in one group that might be served by the skills of another; the ability to bring actors from opposite sides of the structural hole together when it would be profitable, or referral benefits; and control over whose interests are served and what information flows across the structural hole (Burt, 1997). Similarly, the presence of the broker gives the network access to more, nonredundant information.

Individual Consequences of Network Position

Building on Granovetter (1973, 1985), Coleman (1988), and Burt (1992, 1997), OB researchers argue that who we know is as consequential to individual outcomes as who we are. The process of elaborating contingencies and moderators associated with individual outcomes began with Burt (1997), who argued that bridging structural holes was more valuable to actors in senior positions and to managers working across significant boundaries within or across the organization. To do their jobs effectively, actors occupying such positions must capitalize on their social ties, whereas their access to and control of information and opportunities gives them more entrepreneurial prospects, making them seem more capable and more creative than others. Thus, structural hole brokers tend to be promoted earlier than their peers, although because they extract value from the uniqueness of their opportunities, the value of their social capital decreases as the number of actors occupying similar positions increases (Burt, 1997).

Not only does relative position within the network structure impact material outcomes, it also has consequences for effectiveness, performance, and satisfaction. Network centrality—the number of direct contacts and access of a given actor within a network—should, for example, increase individual performance, inasmuch as it bestows information and referral benefits to the focal actor. This relationship is contingent upon the type of network, such that centrality in friendship and communication networks is associated with satisfaction and performance, whereas centrality in adversarial networks lowers satisfaction and performance (Baldwin, Bedell, & Johnson 1997).

Although networks and social capital are largely structural theories, their impact on individual outcomes can be based on perception alone. For example, being perceived of as having high-reputation and high-status friends within an organization boosts the focal actor's reputation, although actually having such a friend has no

incremental effect on reputation (Kilduff & Krackhardt, 1994). Similarly, actors with more accurate mental representations of intra-organizational networks are seen as more powerful, independent of their informal and formal structural positions (Krackhardt, 1990). Being perceived as prominent within a given network is often a function of being visible, affiliated with central organizations, and connected to a social network through both strong and weak ties (Burt, 1982; Cole, 1979; Cole & Cole, 1973). These same characteristics also make actors poor self-estimators, making more internal attributions based on downward social comparison (Cerulo, 1990; Festinger, 1954). Thus, prominent scientists, for example, overestimate their own productivity when their prominence is attributed to internal effort or personal characteristics, whereas they underestimate their relative productivity when their prominence is attributed to external circumstances (Cerulo, 1990).

Social Capital and Minority Groups

The impact of who you know is particularly salient for minority group members. A great deal of research therefore focuses on the impact of the structure of network ties on outcomes consequently achieved by minorities. In general, minority group members benefit from having heterogeneous, nonoverlapping networks, whereas majority group members benefit from having homogeneous networks characterized by multiplex ties. Men generally have more multiplex, homophilous, and strong ties than do women and are more likely to hold central positions in organizational interaction networks; in contrast, women are more likely to differentiate their personal networks, gaining social support and friendship from other women, but instrumental access through network ties to men (Ibarra, 1992).

Like women, minority managers tend to have more racially heterogeneous and fewer intimate network relationships than do nonminorities (Ibarra, 1995). Nevertheless, minorities achieve the best outcomes when they use different groups as social resources. Minority group members with the highest potential for advancement balance the number of cross-race and same-race ties and high-status and low-status ties within their networks, ensuring themselves access to both information and psychosocial support (Ibarra, 1995). Because they maintain potential status distinctions, differentiated networks of this sort may also be beneficial to the interaction of racially heterogeneous work groups. Although actors in racially homogenous groups that segment work-related and nonwork relationships experience lower levels of behavioral and psychological integration with their work teams, similar segmentation in the context of racially diverse work teams actually increases in-team integration (Phillips, Rothbard, & Dumas, 2004).

These findings highlight the need for minority group members to proactively manage their social networks to achieve superior organizational outcomes, as heterogeneous and differentiated networks do not readily emerge for most minority group members. Distinctiveness theory (McGuire, 1984) suggests that actors identify with others with whom they share a relatively rare characteristic within a given social context, because relative rarity heightens the salience of such characteristics. Relative minority group members, therefore, tend to make identity and friendship choices within their own minority groups, rather than with the majority group (Mehra, Kilduff, & Brass, 1998), although Ibarra's (1992, 1995) research points to superior outcomes for those with more balanced networks. Integration into heterogeneous networks may be more difficult for women than for racial and other minorities, however, as racial minorities' relative network marginalization tends to result from individual preferences for same-race friends, whereas women's marginalization is attributable more to exclusionary pressures than to preferences for female friends (Mehra et al., 1998).

Organizational Consequences of Social Networks

At the intraorganizational level, just as network structure influences the process of microinteraction, microinteraction influences network structure, and hence the organization itself. Networks have a significant impact on employee turnover, which tends to occur in clusters, based on employee positions in advice networks. When actors see others in similar positions leaving an organization, they tend to reconsider their options and alternatives, and ultimately leave, as well (Krackhardt & Porter, 1985, 1986). Similarly, attitudes toward work are often subject to the social influence of those with whom an actor interacts regularly and are similar to those of structural equivalents, or actors in similar positions (Erickson, 1988; Krackhardt & Brass, 1994; Krackhardt & Kilduff, 1990; Rice & Aydin, 1991). Evaluations of job design are also affected by network position; actors' evaluations are influenced by those with whom they interact most frequently (Krackhardt & Brass, 1994; Salancik & Pfeffer, 1978). Power, leadership, and motivation are also partially determined through network membership and position (Krackhardt & Brass, 1994).

When OB researchers go beyond the individual level, the material consequences of network structure become more apparent. Field studies at the business-unit level, for example, demonstrate that network structure can help explain outcomes related to information transfer. Strong ties between business units within the same organization are more capable of transferring tacit, complex, uncodified, and interdependent knowledge from one to the other than are weak ties. Nevertheless, weak ties are capable of enabling the transfer of simple, codified, and independent knowledge (Hansen, 1999). Thus the ability of an organization to exploit its own knowledge internally, and hence to innovate and to compete, is contingent upon the structure of its own internal networks. Uzzi (1997, 1999) finds similar results at the interorganizational level in his field studies. The trust, information, communication, and coordination ability inherent in strong ties that are embedded in multiplex relationships—that is, relationships that are both social as well as transactional—lower the cost of transacting among members of such relationships, as they do among

members of networks marked by strong ties (Uzzi, 1997, 1999).

Although social capital and social networks can be built and used instrumentally, the study of social networks does not focus explicitly on their strategic use. Rather, such research generally falls into the domain of power and influence, to which we now turn our attention.

Power and Influence

Organizations are influenced and characterized by the search for and exercise of power. Whereas social psychology deals extensively with the behavioral antecedents and consequences of power and its phenomenological experience (e.g., Anderson & Berdahl, 2002; Dovidio & Ellyson, 1982; Keltner et al., 2003; Overbeck & Park, 2001; Reynolds, Oakes, Haslam, Nolan, & Dolnik, 2000; Tiedens & Fragale, 2003; etc.), and interpersonal influence and persuasion (e.g., Petty & Cacioppo, 1986a, 1986b, 1979a, 1979b; Petty & Krosnick, 1995), OB is more concerned with power as a social phenomenon, situated within the organizational context.

Most OB research on power posits completing views of dependence relationships within organizations or systems. Ocasio (2002) identifies three main schools of thought: the functional school, which sees power emerging in response to organizations' efforts to interact with and manage their environments; the structural school, which emphasizes the criticality of structural position in determining power relationships; and the institutional school, which highlights the role of culture and institutions in distributing power (Ocasio, 2002). Such explanations are not fully satisfactory, however, given the often mixed empirical support for resource dependence theories of power (Astley & Zajac, 1990; Lachman, 1989). A broader view of power has therefore been posited by Ocasio and Pozner (2004), who also take the social psychological mechanisms associated with status into account.

Functional Perspectives

Functional views of power take organizations as social systems adapting to changing environments. Power therefore accrues to the individuals and organizational subunits that contribute to adaptation, survival, and success given the problems facing an organization at a given time. March (1962) and Cyert and March (1963) develop the idea of the organization as resolving conflict with political coalitions, whose composition is contingent upon the contribution made by members given the organizational context. Perrow (1961) builds on Cyert and March by explaining that power accrues to subunits that resolve the tasks most important to the organization, which he defines as securing the resources needed to function and grow; legitimizing operations; assembling necessary skills; and coordinating member activities and relations with other organizations and consumers.

Hickson, Hinings, Less, Schneck, and Pennings (1971) find that subunits gain power to the extent that they cope with the critical uncertainties of other subunits in a way that cannot be achieved by third parties. Thus it is not the performance of Perrow's four basic tasks that generate power but, rather, the extent to which this performance makes others dependent on the focal subunit. Hickson and colleagues also point out that power decreases as a subunit's tasks become routinized and as the information and uncertainty absorption provided by that subunit become less critical. Salancik and Pfeffer (1974) argue that the most power accrues to the organizational subunits that contribute the most critical and scarce resources, and that this power is used to influence resource allocation, in turn.

Whereas these arguments are intuitively compelling, they are not always supported empirically. Functional arguments are supported when research focuses on the contemporaneous coincidence of power with control over contingencies and resources, yet in a longitudinal study, Lachman (1989) finds no correlation between control over contingencies and power over time. Similarly, Astley and Zajac (1990) find that nondirectional interdependence among organizational units is a better predictor of power than dyadic dependencies. This suggests that, in addition to resources and constraints that accrue to actors based on long-term commitments and relationships, power is a function of properties of the broader organizational system.

Structural Perspectives

Structural perspectives focus on the embeddedness of groups and individuals within the structure of social relationships embodied by the organization. They hold that it is not only the ability to perform certain functions or provide resources that generates power but also one's relative position within formal and informal organizational structures. This view is perhaps best expressed by Emerson's (1962) assertion that power is a function of the social relation, not the individual actor. Cook and Emerson (1978) demonstrate further not only that social exchange establishes power but also that the constraints and resources imposed on actors by long-term social commitments play a role. This concept has been successfully extended to multilateral relationships, such as exchange networks (Cook & Emerson, 1978; Cook, Emerson, Gilmore, & Yamaguchi, 1983; Willer, 2003), and different types of exchange (e.g., reciprocal or negotiated) (Molm, Peterson, & Takahashi, 1999). Brass (1984) reinforces the importance of structural position to power distribution in his study of individual power; he finds that certain actors, such as boundary spanners, are likely to accrue more power based not simply on their own abilities but because of their central and critical network positions.

Similarly, much theory and many studies show that power can accrue to actors through others' perceptions of the power and prestige associated with their positions and affiliations, rather than actual capacities and skills (Adler & Kwon, 2002; Bourdieu, 1977, 1985, 1989; Finkelstein, 1992; Krackhardt, 1990). Network position is also central to Burt's (1997) finding that social capital,

or the ability to broker network holes, is more valuable to certain organizational actors: those with relatively few peers, in senior positions, working across significant intra- or interorganizational boundaries, or who are more visible. Individual actors' personal affiliations also have implications for power attributions, as individuals in groups may use the perceived power of a new member to attribute power to the group, and outsiders may base their perceptions of individual group members on the group's reputation (Fiol, O'Connor, & Aguinis, 2001; Fombrun, 1983). Kilduff and Krackhardt (1994) show that friendship with prominent actors increases attributions of the focal actor's own power and skills.

Institutional Perspectives

Institutional perspectives on power agree that structural and functional arguments do not adequately capture the sources of organizational power, because they are embedded in institutions. Rather, they suggest that prevailing rules, norms, and culture shape both structure and function as well as power distribution. The neo-Weberian argument articulated by Jackall (1988) posits that it is both hierarchical position and the control over culture—which represents organizational values, relevant categories, and acceptable alternatives—that form the bases for power. Selznick (1957) argues that power relationships determined by the institutionalization of value-infused organizational commitments and self-definitions. This is echoed by Stinchcombe (1965) and Kimberly (1975), who pose that the structures created at an organization's founding delineate the power of individuals and subgroups, which is resistant to change. Finally, Brown (1978), Oakes, Townley, and Cooper (1998), and Thornton and Ocasio (1999) present a political–cultural approach to organizational power. Brown, supported by the empirical findings of Oakes and colleagues, describes paradigm control and control over agendas as sources of power, whereas Thornton and Ocasio find that certain determinants of power are more salient than others based on cultural framing of organizational attention.

Status and Power

In contrast to perspectives of power and organizational dependence, a more interactionist approach is taken by the sociological social psychologists who study status. This intellectual tradition is rooted in early social psychological studies of status determination in small groups (Sherif, White, & Harvey, 1955). Status can be defined as one's position of honor within a social field (Weber, 1946), and it is often accompanied by beliefs about social worth, such that an actor in a given status position is generally thought to be "better than" or "worse than" an actor in a different position (Sell, Lovaglia, Mannix, Samuelson, & Wilson, 2004; Sewell, 1992). Social fields are thus marked by status structures, or rank-ordered relationships among social actors. Such hierarchies are inherently relational and contingent upon shared values within the social system; that is, two systems with mem-

bers identical on relevant dimensions but different rules of the game would have differently ranked status structures. Although they often conflate the two concepts (e.g., Berger, Conner, & Fisek's [1974] power and prestige order), these theorists explain the complex social psychological links between status and power.

One way to think about the relationship between power and status is through the mechanism of status conferral. Social actors bestow status on each other by enacting voluntary compliance, deference, acceptance, and liking without threat or coercion (Kemper & Collins, 1990). When status-conferring microinteractions are repeated through interaction ritual chains, they contribute to the definition of the macrosocial status structure (Collins, 1981). Status conferral increases the likelihood that the high-status actor will be deferred to again, implying that high-status actors can exert power through deference rather than coercion.

Another interactionist argument is presented by expectation states theory, an outgrowth of social exchange theory. This theory holds that based on their relative status positions, actors form expectations about their own and each other's abilities, skills, and performance; these expectations ultimately determine the distribution of participation, influence, and further prestige (Berger, Cohen, & Zelditch, 1972; Humphreys & Berger, 1981). Based on intersubjective expectations, higher-status actors are given more opportunities to contribute, are more likely to offer contributions, to receive positive evaluations of their contributions, and to be influential in group decisions (Berger et al., 1974). Behavior interchange patterns develop over time into stable interactions that are consistent with the power and prestige order and are generalized to situations outside those in which the original status ordering was created (Fisek, Berger, & Norman, 1991). Status thus becomes an organizing structure for group interaction, in that it defines the terms under which power may be legitimately and acceptably used (Sell, Lovaglia, Mannix, Samuelson, & Wilson, 2004). Snyder and Stukas (1999) find that such status expectations create self-fulfilling prophecies, whereby the perceiver adopts behavior that causes the relatively high-status actor to appear to confirm the perceiver's expectations.

Social psychology has produced a number of findings that help OB scholars understand the microlevel mechanisms working on both the relatively high-status actor and the lower-status actor. High-status actors receive abundant resources, both physical and social, such as attention, flattery, and respect, giving them the phenomenological sense of power (Keltner et al., 2003). Actors experiencing a sense of power have been shown to be more likely to take action, regardless of its potential risks (Galinsky et al., 2003), to attend more to rewards and less to threat and punishment, and to be less inhibited in their behavior (Keltner et al., 2003). This engenders cognitive and behavioral effects on those with whom the relatively high-status actor interacts, as he is perceived by others as more powerful, engendering a submissive reaction on the part of the perceiver (Dovidio & Ellyson, 1982; Tiedens & Fragale, 2003). The perceiver then inter-

prets the actor's dominant behavior as an indicator of status, reinforcing the status-to-power link (Ellyson & Dovidio, 1985).

Status also has an effect on the course of that interaction. Status may increase access to resources that can be used to create dependence (Thye, 2000). Similarly, perceptions of status affect the most microprocesses of social interaction to such an extent that it can create and perpetuate feelings of power, which in turn increase perceptions of power. At the same time, one can have status without power, as well as power without status, and power use may lead to status-reducing conflict and resistance (Willer, Lovaglia, & Markovsky, 1997). Thus, although the link between power and status is sufficient to demonstrate that power is not simply a structural–functional phenomenon, it is still insufficient to explain power alone.

Organizational behavior takes a rather broad view of power, incorporating interactionist, structural, functional, cultural, and experiential perspectives. Because it approaches power from many angles, it has yet to produce a unified theory of power. Nevertheless, this thorough investigation of power, from the psychological to the structural, will ultimately provide a comprehensive view that can be applied not only to organizational contexts but also to any arena of social interaction.

CONCLUSION

As most reviews eventually admit, we have been openly selective in our review of OB research. We focused on research that has a timeless quality and that has already spawned new streams of research. In selecting this research, we hope to show that organizational behavior is not merely an application of social psychological research to organizational contexts. Moreover, we mean to demonstrate the benefits of field research in elaborating the complex mechanisms at work in many social phenomena, particularly those involving small groups and larger social units. Finally, we intend to show how situating social psychological research within a macrosocial context such as the organization provides insight into the dynamics of social interaction, that is, the way that actors' individual responses to social stimuli influence not only their own behavior but also the reactions of other actors as well as the context in which they interact.

We have relied on a levels-of-analysis approach to guide our review of organizational behavior. In doing so, we have revealed how OB scholars define themselves. In this regard, the tensions that define social psychology (individual vs. group research) are similar to those that occur within OB. If anything, the divisions run deeper in OB. A given problem, such as an underperforming team may be accounted for via an individual level of analysis, an interpersonal level of analysis, a team approach, and in many cases a more macroapproach. OB has challenged young researchers to heal the rift and engage in what is euphemistically called mesoresearch. The idea is that a mezzanine is the floor that is halfway between the lobby (micro) and the upper (macro) levels of the organization. Some have tried to navigate this crevasse, but many have failed.

As a field, organizational behavior could be likened to a teenager, desperately trying to gain independence from its mother and father disciplines: social psychology and sociology. Thus, far, mom and dad are not overly impressed by OB. Moreover, the ungrateful teenager enjoys all the benefits of having an influential and successful mom and father but is not reciprocating—at least judging by the frequency with which social psychology and sociology journals cite OB papers (as opposed to vice versa). And, like most teenagers, OB does not really desire to follow in mom and dad's footsteps. And, like many teenagers, OB is not bashful about wanting to take whatever new car (theory or paradigm) that social psychology might have developed out for its own spin. We suspect that as OB continues to mature, there will be a greater independence and perhaps a deepened appreciation between the two fields.

REFERENCES

Adams, S. (1963). Toward an understanding of inequity. *Journal of Abnormal Social Psychology*, *67*, 422–436.

Adams, S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York: Academic Press.

Adler, P. S., & Kwon, S.-W. (2002). Social capital: Prospects or a new concept. *Academy of Management Review*, *27*(1), 17–40.

Alagna, S., Reddy, D., & Collins, D. (1982). Perceptions of functioning in mixed-sex and male medical training groups. *Journal of Medical Education*, *57*, 801–803.

Allen, V. L., & Wilder, D. A. (1975). Categorization, beliefs similarity, and intergroup discrimination. *Journal of Personality and Social Psychology*, *32*, 971–977.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview Press.

Amason, A. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal*, *39*(1), 123–148.

Ambrose, M. L., Seabright, M. A., & Schminke, M. (2002). Sabotage in the workplace: The role of organizational injustice. *Organizational Behavior and Human Decision Processes*, *89*, 947–965.

Ancona, D. G. (1990). Outward bound: Strategies for team survival in an organization. *Academy of Management Journal*, *33*, 334–365.

Ancona, D., & Bresman, H. (2006). Begging, borrowing, and building on ideas from the outside to create pulsed innovation inside teams. In L. Thompson & H.-S. Choi (Eds.), *Creativity in organizational teams* (pp. 183–198). Mahwah, NJ: Erlbaum.

Ancona, D. G., & Caldwell, D. F. (1988). Beyond task maintenance: Defining external functions in groups. *Group and Organizational Studies*, *13*, 468–494.

Ancona, D. G., & Caldwell, D. F. (1990). Beyond boundary spanning: Managing external dependence in product development teams. *Journal of High Technology Management Research*, *1*, 119–135.

Ancona, D. G., & Caldwell, D. F. (1992a). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly*, *37*, 634––665.

Ancona, D., & Caldwell, D. (1992b). Demography and design: Predictors of new product team performance. *Organization Science*, *3*, 321–341.

Anderson, C., & Berdahl, J. L. (2002). The experience of power: Examining the effects of power on approach and inhibition ten-

dencies. *Journal of Personality and Social Psychology*, *83*(6), 1362–1377.

Aquino, K., Tripp, T. M., & Bies, R. J. (2001). How employees respond to personal offense: The effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace. *Journal of Applied Psychology*, *86*(1), 52–59.

Arkes, H. R., & Hammond, K. R. (1986). *Judgment and decision making: An interdisciplinary reader*. Cambridge, UK: Cambridge University Press.

Astley, W. G., & Zajac, E. J. (1990). Beyond dyadic exchange: Functional interdependence and subunit power. *Organization Studies*, *11*, 481–501.

Bacharach, S. B., & Bamberger, P. A. (2004). Diversity and the union: The effect of demographic dissimilarity on members' union attachment. *Group and Organization Management*, *29*, 385–418.

Baldwin, T. T., Bedell, M., & Johnson, J. J. (1997). Social networks in a team-based MBA program: Effects on student satisfaction and performance. *Academy of Management Journal*, *40*, 1369–1397.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barry, B. (1999). The tactical use of emotion in negotiation. In R. J. Bies, R. J. Lewicki, & B. H. Sheppard (Eds.), *Research on negotiation in organizations* (Vol. 7, pp. 93–121). Stamford, CT: JAI Press.

Barry, B., Fulmer, I. S., & Goates, N. (2006). Bargaining with feeling: The role of emotion in negotiation. In L. Thompson (Ed.), *Frontiers of social psychology: Negotiations.* New York: Psychology Press.

Barry, D. (1991). Managing the bossless team: Lessons in distributed leadership. *Organizational Dynamics*, *20*(1), 31–47.

Barsade, S., Ward, A., Turner, J., & Sonnenfeld, J. (2000). To your heart's content: A model of affective diversity in top management teams. *Administrative Science Quarterly*, *45*, 802–836.

Bass, B. M. (1985). *Leadership and performance beyond expectations.* New York: Free Press.

Bass, B. M. (1998). *Transformational leadership: Industrial, military, and educational impact.* Hillsdale, NJ: Erlbaum.

Bass, B. M., & Avolio, B. J. (1990). Developing transformational leadership: 1992 and beyond. *Journal of European Industrial Training*, *14*, 21–37.

Bass, B. M., & Avolio, B. J. (1993). Transformational leadership: A response to critiques. In J. G. Hunt, B. R. Baliga, H. P. Dachler, & C. A. Schriesheim (Eds.), *Emerging leadership vistas* (pp. 29–40). Lexington, MA: Lexington Books.

Bazerman, M. H. (2006). *Judgment in managerial decision making* (6th ed.). Hoboken, NJ: Wiley.

Bazerman, M. H., & Neale, M. A. (1982). Improving negotiation effectiveness under final offer arbitration: The role of selection and training. *Journal of Applied Psychology*, *67*(5), 543–548.

Bazerman, M. H., & Neale, M. A. (1983). Heuristics in negotiation: Limitations to effective dispute resolution. In M. Bazerman & R. Lewicki (Eds.), *Negotiating in organizations* (pp. 51–67). Beverly Hills, CA: Sage.

Behfar, K. J., Peterson, R. S., Mannix, E. A., & Trochim, W. M. K. (2004). *Exploring conflict resolution strategies in autonomous work groups: An example of adaptive structuration in teams*. Paper under review.

Berger, J., Cohen, B. P., & Zelditch, M., Jr. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*, 241–255.

Berger, J., Conner, T. L., & Fisek, M. H. (Eds.). (1974). *Expectations states theory: A theoretical research program*. Cambridge, MA: Winthrop.

Bono, J. E., & Judge, T. A. (2003). Self-concordance at work: Toward understanding the motivational effects of transformational leaders. *Academy of Management Journal*, *46*(5), 554–571.

Bottom, W. P., & Paese, P. W. (1999). Judgment accuracy and the asymmetric cost of errors in distributive bargaining. *Group Decision and Negotiation*, *8*, 349–364.

Bottom, W. P., & Studt, A. (1993). Framing effects and the distributive aspect of integrative bargaining. *Organizational Behavior and Human Decision Processes*, *56*, 459–474.

Bourdieu, P. (1977). Symbolic power. In D. Gleeson (Ed.), *Identity and structure: Issues in the sociology of education* (pp. 112–119). Driffield, UK: Nafferton Books.

Bourdieu, P. (1985). The social space and the genesis of groups. *Social Science Information*, *24*(2), 195–220.

Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, *7*(1), 14–25.

Brass, D. J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science Quarterly*, *29*(4), 518–539.

Brass, D. J. (1985). Men and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management Journal*, *28*, 327–343.

Brett, J. M. (2001). *Negotiating globally: How to negotiate deals, resolve disputes, and make decisions across cultural boundaries*. San Francisco: Jossey-Bass.

Brett, J. M., Adair, W., Lempereur, A., Okumura, T., Shikhirev, P., Tinsley, C., et al. (1998). Culture and joint gains in negotiation. *Negotiation Journal*, pp. 61–86.

Brewer, M. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*, 307–324.

Brockner, J., & Greenberg, J. (1990). The impact of layoffs on survivors: Insights from procedural and distributive justice. In J. S. Carroll (Ed.), *Applied social psychology and organizational settings* (pp. 45–75). Hillsdale, NJ: Erlbaum.

Brockner, J., Grover, S., Reed, T., DeWitt, R., & O'Malley, M. (1987). Survivors' reactions to layoffs: We get by with a little help for our friends. *Administrative Science Quarterly*, *32*(4), 526–541.

Brockner, J., Heuer, L., Magner, N., Folger, R., Umphress, E., van den Bos, K., et al. (2003). High procedural fairness heightens the effect of outcome favorability on self-evaluations: An attributional analysis. *Organizational Behavior and Human Decision Processes*, *91*, 51–68.

Brockner, J., Konovsky, M. A., Cooper-Schneider, R., Folger, R., Martin, C., & Bies, R. J. (1994). The interactive effects of procedural justice and outcome negativity on victims and survivors of job loss. *Academy of Management Journal*, *37*, 397–409.

Brown, R. H. (1978). Bureaucracy as praxis: Toward a political phenomenology of formal organizations. *Administrative Sciences Quarterly*, *23*(3), 365–382.

Bryman, A. (1996). Leadership in organizations. In S. R. Clegg, C. Hardy, & W. R. Nord (Eds.), *Handbook of organization studies* (pp. 276–292). Thousand Oaks, CA: Sage.

Bunderson, J. S., & Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of Management Journal*, *45*, 875–893.

Burns, J. M. (1978). *Leadership.* New York: Harper & Row.

Burt, R. S. (1982). *Toward a structural theory of action: Network models of social structure, perception, and action.* New York: Academic Press.

Burt, R. S. (1992). *The social structure of competition*. Cambridge, MA: Harvard University Press.

Burt, R. S. (1997). The contingent value of social capital. *Administrative Science Quarterly*, *42*, 339–365.

Bygrave, W. D., & Timmons, J. A. (1992). *Venture capital at the crossroads.* Boston: Harvard Business School Press.

Byrne, D. (1971). *The attraction paradigm.* New York: Academic Press.

Cerulo, K. (1990). To err is social: Network prominence and its effects on self-estimation. *Sociological Forum*, *5*(4), 619–634.

Chatman, J. A., & O'Reilly, C. (2004). Asymmetric reactions to work group sex diversity among men and women. *Academy of Management Journal*, *47*(2), 193–208.

Christensen, C., Larson, J. R., Abbott, A., Ardolino, A., Franz, T., &

Pfeiffer, C. (1998). *Decision making of clinical teams: Communication patterns and diagnostic error.* Manuscript under review.

Cohen, L., Broschak, J., & Haveman, H. (1996). *And then there were more?: The effect of organizational sex composition on hiring and promotion.* Unpublished manuscript, Haas School of Business, University of California, Berkeley.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, *17*(1), 1–25.

Cohen, M. D., March, J. G., & Olsen, J. P. (1976). People, problems, solutions, and the ambiguity of relevance. In J. G. March & J. P. Olsen (Eds.), *Ambiguity and choice in organizations* (pp. 24–37). Bergen, Norway: Universitetsforlaget.

Cole, J. R. (1979). *Fair science: Women in the scientific community.* New York: Free Press.

Cole, J. R., & Cole, S. (1973). *Social stratification in science.* Chicago: University of Chicago Press.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95–S120.

Collins, R. (1981). On the micro-foundations of macro-sociology. *American Journal of Sociology*, *86*, 984–1014.

Cook, K. S., & Emerson, R. M. (1978). Power, equity and commitment in exchange networks. *American Sociological Review*, *43*, 721–739.

Cook, K. S., Emerson, R. M., Gillmore, M. R., & Yamagishi, T. (1983). The distribution of power in exchange networks: Theory and experimental results. *American Journal of Sociology*, *89*, 275–305.

Cox, T. Lobel, S., & McLeod, P. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Academy of Management Journal, 34*, 827–847.

Cropanzano, R., & Greenberg, J. (1997). Progress in organizational justice: Tunneling through the maze. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 317–372). New York: Wiley.

Cummings, J. N. (2004). Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, *50*, 352–364.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm.* Englewood Cliffs, NJ: Prentice Hall.

David, B., & Turner, J. C. (1996). Studies in self-categorization and minority conversion: Is being a member of the outgroup an advantage? *British Journal of Social Psychology*, *35*, 179–200.

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, *88*(4), 741–749.

Deutsch, M. (1973). *The resolution of conflict.* New Haven, CT: Yale University Press.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, *53*, 497–509.

Diekmann, K. A., Tenbrunsel, A. E., & Galinsky, A. D. (2003). From self-prediction to self-defeat: Behavioral forecasting, self-fulfilling prophecies, and the effect of competitive expectations. *Journal of Personality and Social Psychology*, *85*(4), 672–683.

Douglas, M. (1986). *How institutions think.* Syracuse, NY: Syracuse University Press.

Dovidio, J. F., & Ellyson, S. L. (1982). Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly*, *45*(2), 106–113.

Druskat, V. U., & Wheeler, J. V. (2003). Managing from the boundary: The effective leadership of self-managing work teams. *Academy of Management Journal*, *46*(4), 435–457.

Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, *45*(4), 735–744.

Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, *129*(4), 569–591.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory or prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598.

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*(1), 3–22.

Earley, P. C. (1989). Social loafing and collectivism. *Administrative Science Quarterly*, *34*, 565–581.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383.

Edmondson, A. C., & Mogelof, J. P. (2006). Explaining psychological safety in innovation teams: Organizational culture, team dynamics, or personality? In L. Thompson & H.-S. Choi (Eds.), *Creativity in organizational teams* (pp. 109–136). Mahwah, NJ: Erlbaum.

Eisenhardt, K., Kahwajy, J., & Bourgeois, L. (1997). Conflict and strategic choice: How top management teams disagree. *California Management Review*, *39*, 42–62.

Ellyson, S. L., & Dovidio, J. F. (Eds.). (1985). *Power, dominance, and nonverbal behavior.* New York: Springer-Verlag.

Ely, R. (1994). The effects of organizational demographics and social identity on relationships among professional women. *Administrative Science Quarterly*, *39*, 203–238.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, *27*(1), 31–41.

Epitropaki, O., & Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology*, *89*(2), 293–310.

Erez, M., & Arad, R. (1986). Participative goal-setting: Social, motivational, and cognitive factors. *Journal of Applied Psychology*, *71*, 591–597.

Erickson, B. H. (1988). The relational basis of attitudes. In B. Wellman & S. D. Berkowitz (Eds.), *Social structures: A network approach* (pp. 99–121). Cambridge, UK: Cambridge University Press.

Espinoza, J., & Garza, R. (1985). Social group salience and inter-ethnic cooperation. *Journal of Experimental Social Psychology*, *21*, 380–392.

Fairhurst, G., & Snavely, B. (1983). A test of the social isolation of male tokens. *Academy of Management Journal*, *26*, 353–361.

Farber, H. S. (1981). Splitting the difference in interest arbitration. *Industrial and Labor Relations Review*, *35*, 70–77.

Farber, H. S., & Bazerman, M. H. (1986). The general basis of arbitrator behavior: An empirical analysis of conventional and final offer arbitration. *Econometrica*, *54*, 1503–1528.

Farber, H. S., & Bazerman, M. H. (1989). Divergent expectations as a cause of disagreement in bargaining: Evidence from a comparison of arbitration schemes. *Quarterly Journal of Economics*, *104*, 99–120.

Fenwick, G. D., & Neal, D. J. (2001). Effect of gender composition on group performance. *Work and Organization*, *8*, 205–225.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.

Finkelstein, S. (1992). Power in top management teams: Dimensions, measurement, and validation. *Academy of Management Journal*, *35*, 505–538.

Fiol, C. M., O'Connor, E. J., & Aguinis, H. (2001). All for one and one for all? The development and transfer of power across organizational levels. *Academy of Management Review*, *26*(2), 224–242.

Fisek, M. H., Berger, J., & Norman, R. Z. (1991). Participation in homogeneous and heterogeneous groups: A theoretical integration. *American Journal of Sociology*, *97*, 114–142.

Fisher, R., & Ury, W. (1983) *Getting to yes: Negotiating agreement without giving in.* New York: Penguin Books.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628.

Folger, R., & Cropanzano, R. (2001). Fairness theory: Justice as accountability. In J. Greenberg & R. Cropanzano (Eds.), *Advances in organizational justice* (pp. 1–55). Stanford, CA: Stanford University Press.

Folger, R., & Konovsky, M. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of Management Journal*, *32*(1), 115–130.

Folger, R., Rosenfield, D., Grove, J., & Corkran, L. (1979). Effects of "voice" and peer opinions on responses to inequity. *Journal of Personality and Social Psychology*, *37*(12), 2253–2261.

Fombrun, C. J. (1983). Attributions of power across a network. *Human Relations*, *36*, 493–508.

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychological Bulletin*, *117*(1), 39–66.

Gaertner, S., & Dovidio, J. (1986). The aversive form of racism. In J. Dovidio & S. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 61–89). New York: Academic Press.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*, 453–466.

Gigone, D., & Hastie, R. (1996). The impact of information on group judgment: A model and computer simulation. In E. H. Witte & J. H. Davis (Eds.), *Understanding group behavior: Consensual action by small groups* (Vol. 1, pp. 221–251). Mahwah, NJ: Erlbaum.

Glick, S., & Croson, R. (2001). Reputations in negotiation. In S. Hoch & H. Kunreuther (Eds.), *Wharton on decision making* (pp. 177–186). New York: Wiley.

Good, L. & Nelson, D. (1971). Effects of person-group and intragroup similarity on perceived group attractiveness and cohesiveness. *Psychonomic Science*, *25*, 215–217.

Goodman, P. S., & Garber, S. (1988). Absenteeism and accidents in a dangerous environment: Empirical analysis of underground coal mines. *Journal of Applied Psychology*, *73*(1), 81–86.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360–1379.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, *91*(3), 481–510.

Greenberg, J. (1993). Stealing in the name of justice: Informational and interpersonal moderators of theft reactions to underpayment inequity. *Organizational Behavior and Human Decision Processes*, *54*, 81–103.

Greenberg, J. (2002). Who stole the money, and when? Individual and situational determinants of employee theft. *Organizational Behavior and Human Decision Processes*, *89*, 985–1003.

Gruenfeld, D. H. (1995). Status, ideology, and integrative complexity on the U.S. Supreme Court: Rethinking the politics of political decision making. *Journal of Personality and Social Psychology*, *68*(1), 5–20.

Gruenfeld, D. H., Mannix, E. A., Williams, K., & Neale, M. A. (1996). Group composition and decision making: How member familiarity and information distribution affect process and performance. *Organizational Behavior and Human Decision Processes*, *67*(1), 1–15.

Gruenfeld, D. H, Martorana, P. V., & Fan, E. (2000). Integrative complexity and changing group composition, *Organizational Behavior and Human Decision Processes*, *82*, 45–59.

Gruenfeld, D. H., Thomas-Hunt, M., & Kim, P. H. (1998). Divergent thinking, accountability, and integrative complexity: Public versus private reactions to majority and minority status. *Journal of Experimental Social Psychology*, *34*, 202–226.

Guetzkow, H., & Gyr, J. (1954). An analysis of conflict in decision-making groups. *Human Relations*, *7*, 367–381.

Hackman, J. R. (1990). Work teams in organizations, an orienting framework. In J. Hackman (Ed.), *Groups that work (and those that don't)* (pp. 1–14). San Francisco: Jossey-Bass.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, *44*(1), 82–111.

Hargadon, A. B. (2006). Bridging old worlds and building new ones: Towards a microsociology of creativity. In L. Thompson & H.-S. Choi (Eds.), *Creativity in organizational teams* (pp. 199–216) Mahwah, NJ: Erlbaum.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hickson, D. J., Hinings, C. R., Less, C. A., Schneck, R. E., & Pennings, J. M. (1971). A strategic contingencies theory of intraorganizational power. *Administrative Science Quarterly*, *16*(2), 216–229.

Hollingshead, A. B. (1996). The rank-order effect in group decision making. *Organizational Behavior and Human Decision Processes*, *68*(3), 181–193.

Homans, G. (1950). *The human group.* New York: Harcourt, Brace & World.

Homans, G. (1958). Social behavior as exchange. *American Journal of Sociology*, *63*(6), 597–606.

Homans, G. (1961). *Social behavior: Its elementary forms.* New York: Harcourt, Brace & World.

House, R. J. (1977). A 1976 theory of charismatic leadership. In J. G. Hunt & L. L. Larson (Eds.), *Leadership: The cutting edge* (pp. 189–207). Carbondale: Southern Illinois University Press.

House, R. J., & Aditya, R. M. (1997). The social scientific study of leadership: Quo vadis? *Journal of Management*, *23*, 409–473.

Humphreys, P., & Berger, J. (1981). Theoretical consequences of the status characteristics formulation. *American Journal of Sociology*, *86*(5), 953–983.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, *37*(3), 422–447.

Ibarra, H. (1995). Race, opportunity, and diversity of social circles in managerial networks. *Academy of Management Journal*, *38*(3), 673–703.

Ibarra, H. (1997). Paving an alternate route: Gender differences in network strategies for career development. *Social Psychology Quarterly*, *60*(1), 91–102.

Jackall, R. (1988). *Moral mazes: The world of corporate managers.* New York: Oxford University Press.

Jackson, S. E. (1992). Team composition in organizational settings: Issues in managing a diverse work force. In S. Worchel, W. Wood, & J. Simpson (Eds.), *Group process and productivity* (pp. 138–173). Beverly Hills, CA: Sage.

Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent research on team and organizational diversity: SWOT analysis and implications. *Journal of Management*, *29*, 801–830.

James, K., & Cropanzano, R. (1990). Perceived equity of a colleague's outcome: Effects on performance. *Social Justice Research*, *4*, 169–185.

Janssens, M., & Brett, J. M. (2003). *Fusion collaboration in global teams.* Research report, Department of Applied Economics, Katholieke Universiteit Leuven, Belgium.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, *40*, 256–282.

Jehn, K. A. (1997). A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly*, *42*, 530–557.

Jehn, K. A., & Mannix, E. A. (2001). The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of Management Journal*, *44*(2), 238–251.

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why some differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly*, *44*, 741–763.

Johnson, J. P., Korsgaard, M. A., & Sapienza, H. J. (2002). Perceived fairness, decision control, and commitment in international joint venture management teams. *Strategic Management Journal*, *23*, 1141–1160.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Per-

sonality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*(4), 765–780.

Judge, T. A., & Ferris, G. (1993). Social context of performance evaluation decisions. *Academy of Management Journal*, *36*, 80–105.

Kahn, L. M. (1993). Managerial quality, team success, and individual player performance in major league baseball. *Industrial and Labor Relations Review*, *46*(3), 531–547.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*(1), 17–31.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under certainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kanter, R. M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, *82*, 965–990.

Kark, R., Shamir, B., & Chen, G. (2003). The two faces of transformational leadership: Empowerment and dependency. *Journal of Applied Psychology*, *88*(2), 246–255.

Keck, S., & Tushman, M. (1993). Environmental and organizational context and executive team structure. *Academy of Management Journal*, *36*, 1314–1344.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*(1), 66–91.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*(2), 265–284.

Kemper, T., & Collins, R. (1990). Dimensions of microinteraction. *American Journal of Sociology*, *96*(1), 32–68.

Kenney, R. A., Schwartz-Kenney, B. M., & Blascovich, J. (1996). Implicit leadership theories: Defining leaders described as worthy of influence. *Personality and Social Psychology Bulletin*, *22*(11), 1128–1143.

Kilduff, M., & Krackhardt, D. (1994). Bringing the individual back in: A structural analysis of the internal market for reputation in organizations. *Academy of Management Journal*, *37*, 87–108.

Kim, W. C., & Mauborgne, R. A. (1991). Implementing global strategies: The role of procedural justice. *Strategic Management Journal*, *12*, 125–143.

Kim, W. C., & Mauborgne, R. A. (1993a). Procedural justice, attitudes, and subsidiary top management compliance with multinationals' corporate strategic decisions. *Academy of Management Journal*, *36*(3), 502–526.

Kim, W. C., & Mauborgne, R. A. (1993b). Procedural justice theory and the multinational corporation. In S. Ghoshal & D. E. Westney (Eds.), *Organization theory and the multinational corporation* (pp. 237–255). London: Macmillan.

Kim, W. C., & Mauborgne, R. A. (1995). A procedural justice model of strategic decision making: Strategy content implications in the multinational. *Organization Science*, *6*(1), 44–61.

Kim, W. C., & Mauborgne, R. A. (1996). Procedural justice and managers' in-role and extra-role behavior: The case of the multinational. *Management Science*, *42*(4), 499–515.

Kim, W. C., & Mauborgne, R. A. (1998). Procedural justice, strategic decision making, and the knowledge economy. *Strategic Management Journal*, *19*(4), 323–338.

Kimberly, J. (1975). Environmental constraints and organizational structure: A comparative analysis of rehabilitation organizations. *Administrative Science Quarterly*, *20*, 1–9.

Kizilos, M., Pelled, L., & Cummings, T. (1996). *Organizational demography and prosocial organizational behavior*. Unpublished manuscript.

Krackhardt, D. (1990). Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly*, *35*(2), 342–369.

Krackhardt, D., & Brass, D. J. (1994). Intra-organizational networks: The micro side. In S. Wasserman & J. Galaskiewicz (Eds.), *Advances in the social and behavioral sciences from social network analysis* (pp. 209–230). Beverly Hills, CA: Sage

Krackhardt, D., & Kilduff, M. (1990). Friendship patterns and culture: The control of organizational identity. *American Anthropologist*, *92*, 142–154.

Krackhardt, D., & Porter, L. W. (1985). When friends leave: A structural analysis of the relationship between turnover and stayers' attitudes. *Administrative Science Quarterly*, *30*, 242–261.

Krackhardt, D., & Porter, L. W. (1986). The snowball effect: Turnover embedded in communication networks. *Journal of Applied Psychology*, *71*(1), 50–55.

Kramer, R. M. (1991). Intergroup relations and organizational dilemmas: The role of categorization processes. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 13, pp. 191–228). Greenwich, CT: JAI Press.

Kramer, R. M., & Brewer, M. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*, 1044–1057.

Kray, L. J., Galinsky, A., & Thompson, L. (2002). Reversing the gender gap in negotiations: An exploration of stereotype regeneration. *Organizational Behavior and Human Decision Process*, *87*(2), 386–409.

Kray, L. J., & Lind, E. A. (2002). The injustices of others: Social reports and the integration of others' experiences in organizational justice judgments. *Organizational Behavior and Human Decision Process*, *89*, 906–924.

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, *80*(6), 942–958.

Lachman, R. (1989). Power from what? A reexamination of its relationship with structural conditions. *Administrative Science Quarterly*, *34*, 231–352.

Larson, A. (1992). Network dyads in entrepreneurial settings: A study of the governance of exchange relationships. *Administrative Science Quarterly*, *37*, 76–104.

Larson, C. E., Foster-Fishman, P. G., & Keys, C. B. (1994). Discussion of shared and unshared information in decision-making groups. *Journal of Personality and Social Psychology*, *67*(3), 446–461.

Larson, J. R., Christensen, C., Franz, T. M., & Abbott, A. S. (1998) Diagnosing groups: The pooling, management, and impact of shared and unshared case information in team-based medical decision-making, *Journal of Personality and Social Psychology*, *75*(1), 93–108.

Laughlin, P. R. (1980). Social combination processes of cooperative problem-solving groups on verbal interactive tasks. In M. Fishbein (Ed.), *Progress in social psychology* (Vol. 1, pp. 127–155). Hillsdale, NJ: Erlbaum.

Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created social climates. *Journal of Social Psychology*, *10*, 271–279.

Lind, E. A., Kray, L., & Thompson, L. (1998). The social construction of injustice: Fairness judgments in response to own and others' unfair treatment by authorities. *Organizational Behavior and Human Decision Processes*, *75*(1), 1–22.

Lind, E. A., Kray, L., & Thompson, L. (2001). Primacy effects in justice judgments: Testing predictions from fairness heuristic theory. *Organizational Behavior and Human Decision Processes*, *85*(2), 189–201.

Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum Press.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Loewenstein, J., Morris, M., Chakravarti, A., Thompson, L., & Kopelman, S. (2005). At a loss for words: Dominating the conservation and the outcome in negotiation as a function of intricate arguments and communication media. *Organizational Behavior and Human Decision Processes*, *98*(1), 28–38.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical

encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin and Review*, 6(4), 586–597.

Lott, A., & Lott, B. (1965). Group cohesiveness as interpersonal attraction: A review of relationships with antecedent and consequent variables. *Psychological Bulletin*, 64, 259–309.

Luthans, F., & Kreitner, R. (1985). *Organizational behavior modification and beyond*. Glenview, IL: Scott-Foresman.

Manz, C. C., & Sims, H. P., Jr. (1980). Self-management as a substitute for leadership: A social learning theory perspective. *Academy of Management Review*, 5, 361–367.

Manz, C. C., & Sims, H. P., Jr. (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly*, 32, 106–1293.

Manz, C. C., & Sims, H. P., Jr. (1989). *SuperLeadership: Leading others to lead themselves*. New York: Prentice Hall.

Manz, C. C., & Sims, H. P., Jr. (1991). SuperLeadership: Beyond the myth of heroic leadership. *Organizational Dynamics*, 19(4), 18–35.

Manz, C. C., & Sims, H. P., Jr. (2001). *The new SuperLeadership: Leading others to lead themselves*. San Francisco: Berrett-Koehler.

March, J. G. (1962). The business firm as a political coalition. *Journal of Politics*, 24(4), 662–678.

March, J. G., & Olsen, J. P. (Eds.). (1976). *Ambiguity and choice in organizations*. Bergen, Norway: Universitetsforlaget.

March, J. G., & Simon, H. A. (1958) *Organizations*. Cambridge, MA: Blackwell.

March, J. G., & Simon, H. A. (with Guetzkow, H.). (1993). *Organizations* (2nd ed.). Cambridge, MA: Blackwell.

McGuire, W. J. (1984). Search for the self: Going beyond self-esteem and the reactive self. In R. A. Zucker, J. Aronoff, & A. I. Rabin (Eds.), *Personality and the prediction of behavior* (pp. 73–120). New York: Academic Press.

McLeod, P., & Lobel, S. (1992, August). *The effects of ethnic diversity on idea generation in small groups*. Paper presented at the Annual Academy of Management Meeting, Las Vegas, NV.

Mehra, A., Kilduff, M., & Brass, D. J. (1996). *Relegation to the margins: Race and gender differences in network position, friendship strategies, and social identity*. Unpublished working paper.

Mehra, A., Kilduff, M., & Brass, D. J. (1998). At the margins: A distinctiveness approach to the social identity and social networks of underrepresented groups. *Academy of Management Journal*, 41, 441–52.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York: Plenum Press.

Meindl, J. R., & Ehrlich, S. B. (1987). The romance of leadership and the evaluation of organizational performance. *Academy of Management Journal*, 30, 91–109.

Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. (1985). The romance of leadership. *Administrative Science Quarterly*, 30(1), 78–102.

Menon, T., & Blount, S. (2003). The messenger bias: A relational model of knowledge valuation. In B. M. Staw & R. M. Kramer (Eds.), *Research in organizational behavior* (Vol. 25, pp. 137–186). Westport, CT: JAI Press.

Menon, T., & Pfeffer, J. (2003). Valuing internal vs. external knowledge: Explaining the preference for outsiders. *Management Science*, 49(4), 497–513.

Menon, T., Thompson, L., & Choi, H.-S. (2006). Tainted knowledge versus tempting knowledge: Why people avoid knowledge from internal rivals and seek knowledge from external rivals. *Management Science*, 52(8), 1129–1144.

Messick, D., & Massie, D. (1989). Intergroup relations. *Annual Review of Psychology*, 40, 45–81.

Messick, D. M., & Sentis, K. P. (l985). Estimating social and nonsocial utility functions from ordinal data. *European Journal of Social Psychology*, 15(4), 389–399.

Miller, C. E., Jackson, P., Mueller, J., & Schersching, C. (1987). Some social psychological effects of group decision rules. *Journal of Personality and Social Psychology*, 52, 325–332.

Molm, L. D., Peterson, G., & Takahashi, N. (1999). Power in negotiated and reciprocal exchange. *American Sociological Review*, 64(6), 876–890.

Moore, D. A., Kurtzberg, T. R., Thompson, L. L., & Morris, M. W. (1999). The long and short routes to success in electronically-mediated negotiations: Group affiliations and good vibrations. *Organizational Behavior and Human Decision Processes*, 77, 22–43.

Moreland, R. L., & Argote, L. (2003). Transactive memory in dynamic organizations. In R. Peterson & E. Mannix (Eds.), *Leading and managing people in the dynamic organization* (pp. 135–162). Mahwah, NJ: Erlbaum.

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition?* (pp. 57–84). Thousand Oaks, CA: Sage.

Moreland, R. L., Argote, L., & Krishnan, R. (1998). Training people to work in groups. In R. S. Tindale, L. Heath, J. Edwards, E. Posavac, P. B. Bryant, Y. Suarez-Balcazar, et al. (Eds.), *Theory and research on small groups* (pp. 36–60). New York: Plenum Press.

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, 102, 331–355.

Morris, M. W., Larrick, R. P., & Su, S. K. (1999). Misperceiving negotiation counterparties: When situationally determined bargaining behaviors are attributed to personality traits. *Journal of Personality and Social Psychology*, 77, 52–67.

Mossholder, K. W., Bennett, N., & Martin, C. L. (1998). A multilevel analysis of procedural justice context. *Journal of Organizational Behavior*, 19(2), 131–141.

Murmann, P., & Tushman, M. (1997). *The effects of executive team characteristics and organizational context on organizational responsiveness to environmental shock*. Unpublished manuscript, Graduate School of Business, Columbia University.

Nadler, J., & Shestowsky, D. (2006). Negotiation, information technology, and the problem of the faceless other. In L. Thompson (Ed.), *Negotiation theory and research* (pp. 145–172). New York: Psychology Press.

Neale, M. A., & Bazerman, M. H. (1983). The effects of framing and negotiator overconfidence on bargaining behavior and outcomes. *Academy of Management Journal*, 28(1), 34–49.

Neale, M. A., & Bazerman, M. H. (1985). The effects of framing and negotiator overconfidence on bargainer behavior. *Academy of Management Journal*, 28, 34–49.

Neale, M. A., & Bazerman, M. H. (1991). *Cognition and rationality in negotiation*. New York: Free Press.

Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart & Winston.

Oakes, L. S., Townley, B., & Cooper, D. J. (1998). Business planning as pedagogy: Language and control in a changing institutional field. *Administrative Science Quarterly*, 43(2), 257–292.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic Management Journal*, 18(S1), 187–206.

Ocasio, W. (2002). Organizational power and dependence. In J. Baum (Ed.), *The Blackwell companion to organizations* (pp. 363–385). Oxford, UK: Blackwell.

Ocasio, W., & Pozner, J. E. (2004). *Beyond dependence: A political capital perspective on power in organizations*. Unpublished manuscript, Northwestern University, Evanston, IL.

O'Reilly, C., Caldwell, D., & Barnett, W. (1989). Work group demography, social integration, and turnover. *Administrative Science Quarterly*, 34, 21–37.

O'Reilly, C., & Chatman, J. (1986). Organizational commitment and psychological attachment: The effects of compliance, identification, and internalization of prosocial behavior. *Journal of Applied Psychology*, 71, 492–499.

O'Reilly, C., Snyder, R., & Boothe, J. (1993). Effects of executive team demography on organizational change. In G. Huber & W. Glick (Eds.), *Organizational change and redesign* (pp. 147–175). New York: Oxford University Press.

O'Reilly, C., Williams, K., & Barsade, S. (1997). Group demography and innovation: Does diversity help? In E. Mannix & M. Neale (Eds.), *Research in the management of groups and teams* (Vol. 1, pp. 183–207). Greenwich, CT: JAI Press.

Overbeck, J. R., & Park, B. (2001). When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology, 81,* 549–565.

Paulus, P. B., Brown, V., & Ortega, A. H. (1999). Group creativity. In R. E. Purser & A. Montuori (Eds.), *Social creativity* (Vol. 2, pp. 151–176). Cresskill, NJ: Hampton Press.

Pazy, A., & Oron, I. (2001). Sex proportion and performance evaluation among high-ranking military officers. *Journal of Organizational Behavior, 22,* 689–702.

Pearce, C. L., Perry, M. L., & Sims, H. P., Jr. (2001). Shared leadership: Relationship management to improve NPO effectiveness. In T. D. Connors (Ed.), *The nonprofit handbook: Management* (pp. 624–641). New York: Wiley.

Pearce, C. L., & Sims, H. P., Jr. (2002). Vertical versus shared leadership as predictors of the effectiveness of change management teams: An examination of aversive, directive, transactional, transformational, and empowering leader behaviors. *Group Dynamics: Theory, Research, and Practice, 6*(2), 172–197.

Pelled, L. H. (1993). *Work group diversity and its consequences: The role of substantive and affective conflict.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Pelled, L. H. (1996a). Demographic diversity, conflict, and work group outcomes: An intervening process theory. *Organization Science, 7,* 615–631.

Pelled, L. H. (1996b). Relational demography and perceptions of group conflict and performance: A field investigation. *International Journal of Conflict Resolution, 7,* 230–246.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly, 44,* 1–28.

Perrow, C. (1961). The analysis of goals in complex organizations. *American Sociological Review, 26*(6), 854–866.

Petty, R. E., & Cacioppo, J. T. (1979a). Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin, 5,* 173–176.

Petty, R. E., & Cacioppo, J. T. (1979b). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology, 37,* 1915–1926.

Petty, R. E., & Cacioppo, J. T. (1986a). *Communication and persuasion: Central and peripheral routes to attitude change.* New York: Springer-Verlag.

Petty, R. E., & Cacioppo, J. T. (1986b). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York: Academic Press.

Petty, R. E., & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences.* Mahwah, NJ: Erlbaum.

Pfeffer, J. (1983). Organizational demography. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 5, pp. 299–357). Greenwich, CT: JAI Press.

Pfeffer, J. (1985). Organizational demography: Implications for management. *California Management Review, 28,* 67–81.

Pfeffer, J., & O'Reilly, C. (1987). Hospital demography and turnover among nurses. *Industrial Relations, 26,* 158–173.

Phillips, K. W. (2003). The effects of categorically based expectations on minority influence: The importance of congruence. *Personality and Social Psychology Bulletin, 29*(1), 3–13.

Phillips, K. W., & Loyd, D. L. (2004). *When surface and deep-level diversity meet: The effects on dissenting group members.* Unpublished manuscript, Northwestern University, Evanston, IL.

Phillips, K. W., Mannix E. A., Neale, M. A., & Gruenfeld, D. H. (2004). Diverse groups and information sharing: The effects of congruent ties. *Journal of Experimental Social Psychology, 40,* 497–510.

Phillips, K. W., Rothbard, N. P., & Dumas, T. L. (2004). *It's not that I don't like you: How status drives preferences for segmentation and social integration in diverse environments.* Working paper, Northwestern University.

Pruitt, D. G., & Carnevale, P. J. (1993). *Negotiation in social conflict.* Pacific Groves, CA: Brooks/Cole.

Pruitt, D. G., Carnevale, P. J., Ben-Yoav, O., Nochajski, T. H., & Van Slyck, M. (1983). Incentives for cooperation in integrative bargaining. In R. Tietz (Ed.), *Aspiration levels in bargaining and economic decision making* (pp. 22–34). Berlin: Springer.

Pruitt, D. G., & Rubin, J. Z. (1986). *Social conflict: Escalation, stalemate, and settlement.* New York: Random House.

Raiffa, H. (1982). *The art and science of negotiation.* Cambridge, MA: Belknap.

Randall, M. L., Cropanzano, R., Bormann, C. A., & Birjulin, A. (1999). Organizational politics and organizational support as predictors of work attitudes, job performance, and organizational citizenship behaviors. *Journal of Organizational Behavior, 20,* 159–174.

Rentsch, J. R., & Klimoski, R. J. (2001). Why do "great minds" think alike? Antecedents of team member schema agreement. *Journal of Organizational Behavior, 22,* 107–122.

Reynolds, K., Oakes, P., Haslam, A., Nolan, M., & Dolnik, L. (2000). Responses to powerlessness: Stereotyping as an instrument of social conflict. *Group Dynamics: Theory, Research and Practice, 4,* 275–290.

Rice, R. E., & Aydin, C. (1991). Attitudes towards new organizational technology: Network proximity as a mechanism for social information processing. *Administrative Science Quarterly, 36,* 219–244.

Richard, O. C. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. *Academy of Management Journal, 43,* 164–177.

Riordan, C., & Shore, L. (1997). Demographic diversity and employee attitudes: Examination of relational demography within work units. *Journal of Applied Psychology, 82,* 342–358.

Ross, L., & Stillinger, C. (1991). Barriers to conflict resolution. *Negotiation Journal, 7*(4), 389–404.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In T. Brown, E. Reed, & E. Turiel (Eds.), *Values and knowledge* (pp. 103–135). Hillsdale, NJ: Erlbaum.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability attribution. *Journal of Personality and Social Psychology, 8,* 322–336.

Rothbart, M., & Hallmark, W. (1988). In-group and out-group differences in the perceived efficacy of coercion and conciliation in resolving social conflict. *Journal of Personality and Social Psychology, 55,* 248–257.

Sackett, P., DuBois, C., & Noe, A. (1991). Tokenism in performance evaluation: The effects of work representation on male–female and black–white differences in performance ratings. *Journal of Applied Psychology, 76,* 263–267.

Salancik, G. R., & Pfeffer, J. (1974). The bases and use of power in organizational decision making: The case of a university. *Administrative Science Quarterly, 19*(4), 453–473.

Sapienza, H. J., & Korsgaard, M. A. (1996). Procedural justice in entrepreneur-investor relations. *Academy of Management Journal, 39*(3), 544–574.

Scarpello, V., & Jones, F. F. (1996). Why justice matters in compensation decision making. *Journal of Organizational Behavior, 17*(3), 285–299.

Schminke, M., Ambrose, A. L., & Cropanzano, R. (2000). The effect of organizational structure on perceptions of procedural fairness. *Journal of Applied Psychology, 85,* 294–304.

Schminke, M., Cropanzano, R., & Rupp, D. (2002). Organization structure and fairness perceptions: The moderating effects of organizational level. *Organizational Behavior and Human Decision Processes, 89,* 881–905.

Schreiber, C. (1979). *Changing places: Men and women in transitional occupations.* Cambridge, MA: MIT Press.

Schulz-Hardt, S., Frey, D., Lüthgens, C., & Moscovici, S., (2000). Biased information search in group decision making. *Journal of Personality and Social Psychology, 78*(4), 655–669.

Scott, W. E., & Podsakoff, P. M. (1982). Leadership, supervision, and behavioral control: Perspectives from an experimental analysis. In L. Frederickson (Ed.), *Handbook of organizational behavior management* (pp. 36–69). New York: Wiley.

Seeley, E., Thompson, L., & Gardner, W. (2003, August). *Power and exploitation in groups: Effects of construal and group size*. Paper presented at the annual meeting of the Academy of Management, Seattle, WA.

Sell, J., Lovaglia, M. J., Mannix, E. A., Samuelson, C. D., & Wilson, R. K. (2004). Investigating conflict, power, and status within and among groups. *Small Group Research*, *35*(1), 44–72.

Selznick, P. (1957). *Leadership in administration: A sociological interpretation*. New York: Harper & Row.

Shah, P. P., & Jehn, K. A. (1993). Do friends perform better than acquaintances? The interaction of friendship, conflict, and task. *Group Decision and Negotiation*, *2*(2), 149–165.

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*, 546–557.

Sherif, M., White, B. J., & Harvey, O. J. (1955). Status in experimentally produced groups. *American Journal of Sociology*, *60*, 370–379.

Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. New York: Macmillan.

Sims, H. P., & Lorenzi, P. (1992). *The new leadership paradigm: Social learning and cognition in organizations*. Newbury Park, CA: Sage.

Smith, K., Smith, K., Olian, J., Sims, H., O'Bannon, D., & Scully, J. (1994). Top management team demography and process: The role of social integration and communication. *Administrative Science Quarterly*, *39*, 412–438.

Snyder, M., & Stukas, A. A. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology*, *50*, 273–303.

South, S., Bonjean, C., Markham, W., & Corder, J. (1982). Social structure and intergroup interaction: Men and women of the federal bureaucracy. *American Sociological Review*, *47*, 587–599.

Stasser, G., & Stewart, D. D. (1992). Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology*, *63*, 426–434.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*, 1467–1478.

Steil, J. M. (1983). The responses to injustice: Effects of varying levels of social support and position of advantage or disadvantage. *Journal of Experimental Social Psychology*, *19*, 239–253.

Stewart, G. L., & Barrick, M. R. (2000). Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type. *Academy of Management Journal*, *43*(2), 135–148.

Stinchcombe, A. L. (1965). Social structure and organizations. In J. G. March (Ed.), *Handbook of organizations* (pp. 142–193). Chicago: Rand-McNally.

Sutton, R. I., & Hargadon, A. (1996). Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly*, *41*, 685–718.

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, *45*, 74–83.

Tetlock, P. E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly*, *48*, 227–236.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology*, *52*, 700–709.

Thomas, K. W., & Kilman, R. H. (1974). *The Thomas–Kilman conflict mode instrument*. Palo Alto, CA: Consulting Psychologists Press.

Thompson, J. (1967). *Organizations in action*. New York: McGraw-Hill.

Thompson, L., & Hastie, R. (1990). Social perception in negotiation. *Organizational Behavior and Human Decision Processes*, *47*(1), 98–123.

Thompson, L., & Hrebec, D. 1996. Lose-lose agreements in interdependent decision making. *Psychological Bulletin*, *120*(3), 396–409.

Thoreson, E. E., & Mahoney, M. J. (1974). *Behavioral self-control*. New York: Holt, Rinehart & Winston.

Thornton, P. H., & Ocasio, W. (1999). Institutional logics and the historical contingency of power in organizations: Executive succession in the higher education publishing industry, 1958–1990. *American Journal of Sociology*, *105*(3), 801–843.

Thye, S. R. (2000). A status value theory of power in exchange networks. *American Sociological Review*, *65*(3), 407–432.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementary in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*(3), 558–568.

Tinsley, C. H., O'Connor, K. M., & Sullivan, B. A. (2002). Tough guys finish last: The perils of a distributive reputation. *Organizational Behavior and Human Decision Processes*, *88*(2), 621–642.

Tripp, T. M., Bies, R. J., & Aquino, K. (2002). Poetic justice or petty jealousy? The aesthetics of revenge. *Organizational Behavior and Human Decision Processes*, *89*, 966–984.

Tsui, A., Egan, T., & O'Reilly, C. (1992). Being different: Relational demography and organizational attachment. *Administrative Science Quarterly*, *37*, 549–579.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–192). New York: Academic Press.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, *42*, 35–67.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking capital. *American Sociological Review*, *64*, 481–505.

Valley, K. L., Neale, M. A., & Mannix, E. A. (1995). Friends, lovers, colleagues, strangers: The effects of relationships on the process and outcome of dyadic negotiations. In R. J. Bies, R. J. Lewicki, & B. H. Sheppard (Eds.), *Research on negotiation in organizations* (Vol. 5, pp. 65–93). Greenwich, CT: JAI Press.

Van den Bos, K., Lind, E. A., & Wilke, H. A. M. (2001). The psychology of procedural and distributive justice viewed from the perspective of fairness heuristic theory. In R. Cropanzano (Ed.), *Justice in the workplace: From theory to practice* (pp. 49–66). Mahwah, NJ: Erlbaum.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Waldman, D. A., Javidan, M., & Varella, P. (2004). Charismatic leadership at the strategic level: A new application of upper echelons theory. *Leadership Quarterly*, *15*, 355–380.

Waldman, D. A., Ramirez, G. G., House, R. J., & Puranam, P. (2001). Does leadership matter?: CEO leadership attributes under conditions of perceived environmental uncertainty. *Academy of Management Journal*, *44*, 134–143.

Watson, W. E., Johnson, L., & Merritt, D. (1998). Team orientation, self-orientation, and diversity in task groups: Their connection to team performance over time. *Group and Organization Management*, *23*, 161–189.

Watson, W., Kumar, K., & Michaelsen, L. (1993). Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. *Academy of Management Journal*, *36*, 590–602.

Weber, M. (1946). *From Max Weber: Essays in sociology* (H. H. Girth & C. W. Mills, Eds.). New York: Oxford University Press.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York: Springer-Verlag.

Wegner, D. M., Giuliano, T., & Hertel, P. (1985). Cognitive interdependence in close relationships. In W. J. Ickes (Ed.), *Compatible and incompatible relationships* (pp. 253–276). New York: Springer-Verlag.

Wharton, A., & Baron, J. (1987). So happy together?: The impact of

gender segregation on men at work. *American Sociological Review*, *52*, 574–587.

Wharton, A., & Baron, J. (1991). Satisfaction?: The psychological impact of gender segregation on women at work. *Sociological Quarterly*, *32*, 365–387.

Willer, D. (2003). Power-at-a-distance. *Social Forces*, *81*(4), 1295–1334.

Willer, D., Lovaglia, M., & Markovsky, B. (1997). Power and influence: A theoretical bridge. *Social Forces*, *76*(2), 571–603.

Williams, K. Y., & O'Reilly, C. (1998). Demography and diversity in organizations: A review of 40 years of research. In B. Staw & R. Sutton (Eds.), *Research in organizational behavior* (Vol. 20, pp. 77–140). Greenwich, CT: JAI Press.

Wittenbaum, G., & Stasser, G. (1996). Management of information in small groups. In J. Nye & M. Brower (Eds.), *What's social about social cognition? Social cognition research in small groups* (pp. 3–28). Thousand Oaks, CA: Sage.

Yukl, G. (1998). *Leadership in organizations* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Yukl, G. (1999). An evaluation of conceptual weaknesses in transformational and charismatic leadership theories. *Leadership Quarterly*, *10*, 285–305.

CHAPTER 41

Social Action

MARK SNYDER
ALLEN M. OMOTO

In many ways, working alone and working together, people take action that benefits society. Not all these efforts are necessarily motivated by an explicit desire to benefit society, certainly, but the combined effects of individual action can have a profound effect on society. For example, people may practice the habits of recycling and conserving energy, as well as use mass transit, in order to preserve and conserve natural resources or even to save a few dollars or to avoid the stress of driving at rush hour. They may serve as volunteers and provide services to other people who have difficulty caring for themselves. They may participate in programs in schools and in the workplace that provide opportunities for community service. Or, in an effort to spend more time with their children and be a positive influence in their lives, they may coach little league sports teams or serve as the leader or chaperone of youth groups. They may join neighborhood groups and organizations and, where none exist to meet the needs of their communities, take the initiative to found them and assume leadership roles in them. They may vote and work on political campaigns, and even run for office themselves, in order to elect political leaders who will work on behalf of causes they value. They may engage in lobbying and advocacy efforts to arouse the passions and efforts of other people or to work for the passage of legislation of concern to them. They may join and be active in social movements that are dedicated to causes of concern to them, such as improving the living conditions of disadvantaged groups in society, protecting and expanding human rights, and working for peace at home and abroad.

These activities are all instances of individuals seeking to address problems of society by engaging in what is often referred to as civic engagement, citizenship behaviors, or (more generically) *social action*. It has often been suggested that one way to solve many of the problems confronting society is to promote such forms of social action, that is, to encourage people to act in ways that will benefit not only themselves as individuals but also the larger communities and the society of which they are members (e.g., Omoto, 2005a; Omoto & Snyder, 2002; Snyder & Omoto, 2001; Van Vugt & Snyder, 2002). Social action by individuals and groups can and does take many forms. Some social action is explicitly political, such as involving oneself in the political process by voting or working for an election campaign, but many forms of participation are not necessarily politically motivated or guided, such as volunteering one's time to help others in need, looking out for one's neighbors to deter crime, conserving natural resources, and becoming active in community groups or organizations.

SOCIAL ACTION AND SOCIAL PROBLEMS

More generally, many social commentators have emphasized the importance of social action as a form of "citizenship" and "citizen participation" involving individuals taking action in response to societal problems (e.g., Boyte & Kari, 1996) and in generating "social capital"—bonds of trust among citizens that are built through participation in the affairs of their larger communities

(Coleman, 1990; Portes, 1998; Putnam, 1993, 1995, 2000). In fact, there is a substantial body of research that examines such citizenship behaviors and the ways in which the coordinated activities of individuals can serve the common good (e.g., Kymlicka & Norman, 1994; Putnam, 1993, 1995; Van Vugt, Snyder, Tyler, & Biel, 2000; Verba, Schlozman, & Brady, 1995).

The role of social action and citizen participation in solving societal problems is an enduring fixture of American society and is widely endorsed, regardless of political leaning (e.g., Chambré, 1989). Those on the political right (in the United States, and elsewhere) may promote civic participation as a means to save government money, to enhance local control over important issues, and to avoid what they see as potentially oppressive government-imposed solutions to local problems. And, those on the political left may advocate for grassroots organizing, and encourage individuals and groups to cooperate in order to effect positive community change, including community growth and empowerment driven by citizen participation. Although they may not agree on which problems are pressing or on the best solutions to adopt, there appears to be agreement in valuing and encouraging involvement of individuals in taking action to solve social problems. As one indication of how much a part of the shared ideology of the U.S. social action is, one need only note that every President from John F. Kennedy to George W. Bush, over a period of close to a half century, in accepting the nomination of his party or in his inaugural address to the nation, has stressed the importance of citizen involvement and doing good works. (See also the historical account of de Tocqueville, 1835/1969, who stressed how deep the roots run in North America for interest in and commitment to public and shared service as well as efforts to join together and work with others without regard to political differences.)

Our concern in this chapter is with understanding the nature of social action—to articulate principles that can aid in understanding how and why individuals take action on behalf of others and that benefit society. The phenomenon of social action is intriguing for a variety of reasons. From a practical standpoint, social action involves real people taking real actions in the service of real causes, often doing so over extended periods of time and at considerable personal cost and sacrifice. Moreover, the outcomes of social action have real consequences for the well-being of individuals and the effective functioning of society. From a theoretical standpoint, they represent compelling instances of motivated and goal-directed activities, ones that capture the bridging of individual and collective action in pursuit of ends that benefit not only those taking action but also society at large. In fact, as we discuss in greater depth later in this chapter, considerations of social action simultaneously span and link several levels of analysis, and, accordingly, research on these topics holds the promise of helping to develop broader and more comprehensive understandings of many phenomena of interest to all stripes of psychologists as well as sociologists, economists, historians, anthropologists, and political scientists.

In this chapter, we seek an understanding of the principles that account for and govern social action. Our analysis proceeds sequentially and on two fronts, seeking descriptive principles and explanatory principles. The quest for *descriptive principles* seeks answers to the general question of what social action is, including the tasks of discovering basic organizing principles that reveal the characteristic features of social action, that tell us what can profitably be included in the category of social action and what might be excluded from that category. The search for *explanatory principles* seeks answers to the general questions of what causes and what accounts for social action and involves a search for the causal determinants of social action and the purposes served by social action.

DESCRIPTIVE PRINCIPLES: WHAT SOCIAL ACTION IS

To pursue the goal of explicating descriptive principles of social action, let us articulate, by building on the examples with which we began this chapter, an organizing framework of characteristic features that cross-cut various forms of social action. In so doing, we seek to specify what is to be included in the category of social action and what is to be excluded from that category (as well as some phenomena that test the limits of what is and what is not social action). Moreover, understanding the nature of social action may inform the search for an understanding of the processes that generate and account for social action; that is, knowing what social action is may help explain what causes social action. We hasten to point out that ours is an initial attempt to provide organizing principles for this realm and not an exhaustive list of defining and mutually exclusive characteristics. In other words, we have a more modest goal of trying to offer a set of principles that can help to map the terrain of social action—a taxometric tool, so to speak, that can be used to link and distinguish different forms of social action.

An Individual Phenomenon, but Also a Social Phenomenon

To begin with, we propose as our first descriptive principle that social action is both an individual phenomenon and a social phenomenon. The phenomena of social action are *individual* phenomena in that they involve the actions of individuals, reflecting their individual concerns, their personal values, their own motives, and their particular goals. These phenomena are *social* as well, and they are social in several different ways. First, they are social in that they are engaged in not only by individuals but also by collections of people, and often by people who band together to perform these activities in groups, organizations, and communities (for a related discussion of the "levels" at which helping and prosocial behavior can occur, see Penner, Dovidio, Piliavin, & Schroeder, 2005). Second, they are also social in that they are done on behalf of other people, who are the beneficiaries of social action, and often for the betterment of society at large.

That is, they are actions that are generally intended to serve a social good. Thus, social action involves activities that have social and societal impact, impact that is arguably most evident when considered collectively, as the combined consequences of individual actions. Furthermore, and building on this idea that social action is, at one and the same time, both an individual and a collective phenomenon, social action represents a connection between the individual and the social, a way for individuals to join their interests with the interests of other people, to bond with their communities, and to become engaged with society at large. That is, and as we shall see, considerations of social action provide a template for understanding the linkages between and among individuals, the groups to which they belong, the communities of which they are members, and with the larger society—and for bridging individual, group, and societal perspectives on the human condition.

Socially Valued, but Not Socially Mandated

Next, we note that many forms of social action have in common that, at least in North America, they are socially valued but they are not socially mandated. Indeed, surveys of public opinion in the United States consistently reveal widespread support for the ideals of helping others and getting involved in the affairs of one's community to make society a better place for all. For example, in the case of volunteerism as a form of social action, national surveys reveal that, by margins of over 3 to 1, Americans agree that "people should volunteer some of their time to help other people and thereby make the world a better place" and that "nonprofit organizations generally play a major role in their communities" (Independent Sector, 1988, 1999). The history and value of social action, of course, varies by culture and by country (e.g., Curtis, Grabb, & Baer, 1992; Levine, Norenzayan, & Philbrick, 2001), but our general point here is that many social action activities are culturally and socially valued, and especially when they are seen to contribute to a common good.

However, as much as social action is valued, there are no laws that mandate participation in such activities, no rules or commandments that dictate that one must be a volunteer or that one must join a community organization or that one must participate in a social movement or even (in the United States, at least) that one must vote. Rather, when people become involved in social action—whether it is when they volunteer, when they join community organizations, when they join social movements, or when they vote—they do so as because they choose to do so and because they want to do so. That is, involvement in social action, when it occurs, is a volitional or chosen action, undertaken without the requirement or the obligation to become involved. To the list of defining characteristics of social action, then, we add "activity that is seen as socially valued," but to the list of characteristics that help us to exclude certain behaviors from the category of social action we add "activity in response to requirements or coercion."

Reactive Impetus, Proactive Action

As much as social action occurs without obligation and with volition, the fact of the matter is that the impetus for social action typically is "reactive," because it occurs in response to a perceived need to change something or solve a problem. Thus, individuals who are concerned about threats to the environment may respond to those threats by forming an environmental action group to work to protect and preserve the quality of the environment. Similarly, those who are concerned about the lack of child-care services in their neighborhood may be moved to organize a network of volunteers to provide the needed services. And, in response to low wages and bad working conditions, workers may form unions to change their situations. In each case, the impetus for action, and the formation of vehicles for taking action, is provided by the perception of problems that need to be solved. Furthermore, the veridicality of the perception and, even how widely held the perception is, are much less important than the simple fact that one or more individuals think that something needs to be done.

However, whereas the precipitating conditions for social action may be reactive, when individuals mobilize themselves to take action, the actions they take are "proactive" in the sense that individuals must take the initiative to get involved, make the time and invest the effort to work toward their goals, and overcome obstacles to participation (including, at times, the skeptical reactions of friends and family who doubt whether their actions will really make a difference). Thus, for example, a social movement may come about because of a need for change, but then people who wish to participate in this social movement must figure out when, where, and how to act. As such, their actions take on a proactive character.

Moreover, at times, social action can and does occur in the context of institutional structures or mandates. For example, voting occurs in the context of clearly defined social structures that dictate when, where, and how it will occur. Other times, social action occurs outside formal structural contexts. For instance, many forms of volunteering are informal, as when neighbors help each other out and devise structures for their involvement, including "town hall" meetings, phone trees, and neighborhood block watch associations. Nevertheless, whether it occurs in the context of formal structures or not, social action can be construed as simultaneously reactive and proactive. And, in fact, it can be speculated that there may be important consequences for encouraging social action participation as well as the effects of social action activities to the extent that these activities are construed as relatively reactive versus proactive in nature.

Individual Actions, Aggregated Consequences

Although people may engage in social action as individuals, the impact of social action becomes particularly evident in the aggregation of the actions of individuals and the aggregated consequences of individuals' actions.

Thus, another descriptive principle of social action is that, commonly, it involves individual actions and aggregated consequences. For example, in voting, it is the aggregated actions of individual voters that decide elections. Similarly, in human rights movements, it is the aggregated actions of individual participants that change society. And, in volunteerism, it is the aggregated actions of individual volunteers that help to solve societal problems, such as when tutors who teach children, one by one, to read contribute to solving the problem of illiteracy.

That social action involves individual actions and aggregated consequences serves to underscore, yet again, the principle that social action is, at one and the same time, both an individual phenomenon, involving the actions of individual social actors, and a collective phenomenon, involving the aggregated consequences of the actions of many individuals working alone or in coordinated fashion.

Multiple Pathways, Common Goals

In many forms of social action, the same outcome can come from different means. These multiple pathways to common goals can be illustrated in the case of participation in a social movement (e.g., Klandermans, 1997), say an environmental action movement. Here, someone participating in such a movement who wants to take action to protect the environment and conserve resources could engage in recycling, could turn down the thermostat, could join an environmental action group, and could recruit their friends and neighbors to engage in these activities. All these actions are different means toward the shared end of protecting the environment, conserving and preserving resources, and enhancing general sustainability (e.g., Oskamp, 2000). Stated another way, these actions reflect the descriptive principle of multiple pathways to a common goal.

Similarly, in many forms of social action, there is often a certain, if even implicit, "substitutability," both on the part of those who engage in social action and on the part of those who benefit from it. For example, in the case of volunteerism as a form of social action, a volunteer Big Brother/Big Sister could shift from working with one child to another without his or her status as a volunteer changing. Similarly, as the recipient of the services of a Big Brother/Big Sister program, a child could shift from one volunteer to another but still be a beneficiary of social action. To some extent, this substitutability is related to, and may even derive from, the fact that many forms of social action are focused on helping "categories" of people rather than specific individuals, and on meeting broad needs and serving general causes. That is, even though social action may be carried out in the context of specific individuals delivering services to specific recipients, it is generally conceptualized by those who engage in social action, and by society at large, in terms of helping categories of people (e.g., helping "the needy" rather than helping the specific recipients of help) or furthering broad social causes (e.g., the Green movement).

Consequences at Several Levels, Direct and Indirect Effects

Just as there are numerous routes to the goals of social action, so too are there diverse consequences and outcomes of social action. In keeping with the view of social action as both an individual and a collective phenomenon, these consequences and outcomes can be viewed as spanning a continuum that runs from the individual through the collective. Thus, when it comes to the outcomes of pursuing agendas for social action, many of these outcomes will be at the level of the individual (such as the increases to self-esteem, affirmation of values, and new skills that may accrue to those who engage in social action). Other outcomes will be of a more interpersonal nature (such as the new people one meets, the new friends one makes, and the new patterns of socializing that will develop). And, other outcomes will affect one's relationships (such as changes in patterns of social support, impact of participation on one's existing relationships). Still other outcomes will affect groups, organizations, and society (such as when organizations and movements meet their goals through the service of their members, or nations becoming "kinder and gentler" through the good works of their citizens).

As well, the consequences of social action range in terms of their directness and their concreteness. Some consequences are specific and concrete whereas others are diffuse and abstract. For example, in the case of volunteerism, there are the direct benefits to volunteers of becoming involved (e.g., increased self-esteem, meeting people, and making friends) and the direct benefits to the recipients of volunteer service (e.g., the companionship provided by volunteers to homebound elderly persons and the tutoring provided by volunteers to children who cannot yet read).

At the same time, there are growing indications of indirect benefits, not necessarily foreseen or sought by volunteers, but not necessarily unwelcome either. Thus, among the consequences of volunteerism (and other forms of social action) are better health and psychological functioning, and even a longer life (e.g., Berkman & Glass, 2000; Brown, House, Brown, & Smith, 2004; Brown, Nesse, Vinokur, & Smith, 2003; House, 2001; House, Robbins, & Metzner, 1982; Kawachi & Berkman, 2000; Musik, Herzog, & House, 1999; Piliavin, 2003, 2004a, 2004b; Schwartz, Meisenhelder, Ma, & Reed, 2003; Thoits & Hewitt, 2001). On the other hand, there may also be unforeseen and negative consequences of volunteer work, such as when volunteers feel embarrassed, stigmatized, or otherwise uncomfortable when their acts of volunteerism, and especially the causes and people that their efforts support, become known to their broader social networks (e.g., Ratner & Miller, 2001; Smith, Omoto, & Snyder, 2001b).

Dimensions of Variability across Forms of Social Action

In the context of the shared features of social action, there are some important dimensions of variability—ways

in which not all forms of social action are created alike. Thus, involvement in social action may vary in how extended over time it is. For example, voting takes but a few moments of one's time, but serving in political office extends for the length of one's run for office and, if elected, one's term of office. Nevertheless, even social actions that, individually, take little time, such as recycling, may become repeated actions, occurring with regularity and predictability and acquiring the status of habits that may even become integrated into one's identity (e.g., Callero, Howard, & Piliavin, 1987; Grube & Piliavin, 2000). Giving blood and helping to build houses for homeless or impoverished individuals are other examples of repeated actions that require relatively more time. Moreover, many forms of helping, such as participating in immediate disaster relief, may be limited in time, but other forms, such as volunteering on an ongoing basis, may extend over considerable periods of time, with many volunteers serving for periods of years and even decades. In short, an important distinction here is between *episodic* versus *sustained* activities (e.g., see Wilson, 2000).

Social action also varies in the degree of institutional support and structure surrounding it. For example, social action can occur in the context of *formal* organizations (such as the Red Cross or the Salvation Army) that recruit volunteers and clients, and that engage in selection, training, placement, and monitoring of service. Or, it can occur in relatively more *informal* settings, such as when neighbors take the initiative to help others on their blocks who need assistance (e.g., reciprocal baby-sitting arrangements, or helping elderly neighbors by mowing their lawns or cleaning their gutters). This dimension of "formalness" or degree of "institutionalization" has consequences for estimating the prevalence of social action activities, to be sure, but it also may have implications for why people might be attracted to certain social action activities as well as how, and how easily, they are carried out. Related to our substitutability point discussed earlier (the principle of "multiple pathways, common goals"), we offer the conjecture that social actions that occur within formal organizations may have a higher degree of substitutability associated with them than actions that take place in more informal contexts.

When formal organizational structures for social action exist, furthermore, they may vary in whether they are *local* in reach (e.g., a neighborhood block watch), *national* in focus (e.g., the Sierra Club), or even *international* in scope (e.g., Amnesty International). Thus, there is variability in the extent to which social action efforts are embedded in or organized around structures, with these structures themselves differing in degree of breadth. In many cases, formal structures for social action are linked in umbrella networks across multiple levels. For example, many local animal protection organizations are part of regional affiliations as well as tied to a national organization, the Humane Society of the United States. Moreover, the activities and efforts at the multiple levels are often coordinated (e.g., national campaigns that have local events), a fact that helps to underscore our point about the impact of social action being most acutely observed

with aggregation (the principle of "individual actions, aggregated consequences").

What Is Not Social Action?

Taken together, our set of characteristic features of social action and the dimensions of variability that we have presented also give some indication of what we believe should be *excluded* from the category of social action. First, actions that occur under obligation, whether legal or moral or contractual, such as helping members of one's own family or one's spouse, would not qualify as social action because they occur with obligations stemming from familial and marital bonds. Similarly, the payment of taxes, although of benefit to society and a benefit that is most apparent in the aggregation of the inputs of many taxpayers, would not qualify as social action because it occurs with the legal obligation to pay taxes and the threat of penalties and sanctions for evading payment. More generally, actions performed under institutional pressure and in response to obligations to comply, and those that have sanctions for noncompliance, would not qualify as social action.

In addition, we wish to make clear that social action is not simply synonymous with behavior that is good for everyone. Some activities that can be characterized by our descriptive principles, for example, may benefit some individuals while disadvantaging or harming others. This state of affairs is especially evident in intergroup contexts or in situations of direct competition between social groups. In fact, many social actions, including social movements and services, develop and flourish in settings that are characterized by competition between movements or groups. Individuals may engage in social action activities intended to support one group, but those activities may simultaneously have negative effects on other (often opposing) groups.

Returning to our principle of "socially valued, but not socially mandated," therefore, we add the *caveat* that "socially valued" is determined within a specific group, societal, or cultural context. For example, social action on behalf of gay rights or the right of women to terminate a pregnancy may not be universally accepted and admired. In fact, depending on one's political leanings and one's religious or moral values, such activities may be construed as actions that work to the detriment of another group or against one's sense of the greater good. As another example, consider Ku Klux Klan activities and rallies. These activities may not be generally endorsed by the majority of Americans, yet they can be characterized by many of the descriptive principles that we have discussed. What is good for one social movement or cause may not be good for everyone or for all social actors. Thus, the simple point is that social action cannot be characterized or defined simply by its effects—all social action is not necessarily good for all people.

Similarly, although many examples of social action are also prosocial behaviors (e.g., Batson, 1998; Dovidio, Piliavin, Schroeder, & Penner, 2006), we do not wish to imply that helping and prosocial behaviors are the same

things as social action. While it is true that many forms of helping behavior share characteristics of social action, our descriptive principles can be applied, as noted previously, to many "unhelpful" behaviors as well. In addition, many instances of helping occur in direct response to an appeal for help or out of an obligation to provide aid and cannot logically be aggregated across actions or actors. These helping behaviors may be *ad hoc*, informal, and personally directed, and they include examples such as providing a ride to an elderly neighbor, fixing a meal for a coworker who has just had a baby, or making a one-time donation to a charitable cause in memory of a deceased family member or friend. These are all prosocial actions, to be sure, and they are interesting and important phenomena. However, they do not possess many of the characteristics of social action that we have explicated, rendering a full treatment of them outside the scope of this chapter. Thus, we view prosocial behavior and social action as only partially overlapping sets, and for the sake of conceptual clarity, we use the descriptive principles we have outlined to determine when prosocial behavior can also be considered a form of social action.

Finally, there are cases that test the boundaries of what constitutes social action. For example, service learning and mandatory volunteer programs in schools and corporations (that require people to volunteer or do other forms of community service) and "incentived" volunteerism (such as elder volunteer programs that pay limited stipends to volunteers) build institutional structures around what would otherwise be voluntary actions and create curious mixtures of voluntary and nonvoluntary social action. In these programs, incentives (e.g., course credit and stipends) are generally small and may be just enough to get individuals to engage in social action. In addition, individuals may retain choice about the causes and organizations with which to volunteer and donate their time. Thus, the *act* of volunteering or providing assistance may be dictated by an external agent or barely sufficient justification (e.g., Festinger & Carlsmith, 1959), but the social actor him- or herself has a fair degree of choice about where, when, and how to engage. And, for many programs, social action (e.g., volunteering) is but one option that individuals have for meeting requirements. (For further discussion of service learning and mandatory volunteerism programs, see, for example, Bringle & Duffy, 1998; Stukas, Clary, & Snyder, 1999; Stukas, Snyder, & Clary, 1999; Tschirhart, Mesch, Perry, Miller, & Lee, 2001.)

Thus, in some cases it may be meaningful to talk about activities as forms of social action that are characterized by the principles that we have outlined earlier. In other instances, the "fit" to our descriptive principles of social action may be less good. We do not include further discussion of "hybrid" variants of social action in this chapter (such as service learning). However, we note that they share several of the characteristic features of social action that we have described. In addition, they are topics worthy of research, for both practical reasons revolving around program implementation and improvement, and also to more fully build the body of scientific information on social action and related phenomena. At this time, however, we are not yet ready to fully apply our principles of social action to these intriguing and important behaviors.

Summary

At the outset, we suggested that our first goal for this chapter was to offer a set of descriptive principles that could be used to characterize social action. Having reviewed several principles, we emphasize again that this is neither an exhaustive nor necessarily a prioritized list of principles. That is, we leave open the possibility that our list of principles may need to be modified, most likely, we believe, needing to be expanded rather than reduced, in order to better capture and characterize the many diverse forms of social action. And, in the absence of compelling evidence, we are currently reluctant to claim that any principle or characteristic is more important than any other in defining or understanding social action. We have offered our descriptive set of principles in hopes of providing a conceptual framework for approaching the study of social action, and to bring some (even tentative) order to a dizzying array of social behaviors that sometimes have been studied together and sometimes been considered apart. As we suggested, our goal has been a modest one—simply to provide some guideposts and roadmaps for traversing the topography of social action.

To summarize, then, we have articulated a set of descriptive principles of social action. These principles include characteristic features of social action as well as dimensions of variability across forms of social action. Together, these principles help to organize the wide range of behaviors that can be considered forms of social action, and they also offer guides for determining which phenomena are to be included and which are to be excluded from the category of social action. Moreover, these descriptive principles serve to define what various forms of social action have in common, as well as some of the ways in which they differ from each other. They also aid in determining the degree of "prototypicality" of the different forms of social action, with some forms (such as volunteerism) matching the prototype of social action better than others (such as service learning). And, as we shall see, these descriptive principles point the way toward a consideration of the explanatory principles of social action that help to account for why social action occurs and may determine the forms that it takes.

EXPLANATORY PRINCIPLES: WHY SOCIAL ACTION OCCURS

Clearly, social action occurs. It occurs at the level of individual behaviors such as recycling, energy conservation, blood donation, voting, charitable giving, and volunteering. And, it occurs at the level of collective actions such as participating in neighborhood groups, community organizations, self-help groups, political campaigns, and social movements. But, *why* does it occur? What leads peo-

ple to engage in social action, and what keeps them involved in social action? It is to these questions that we now turn, as we seek to articulate a set of explanatory principles of social action.

Many of the characteristic features of social action suggest that it involves individuals and groups actively setting and pursuing agendas for action. Specifically, that various forms of social action occur without obligation, that individuals seek out opportunities to engage in social action, that they persist in these activities over extended periods of time, often incurring personal costs at the same time as they do social good, and that social action may involve multiple routes to the same ends, all are suggestive that social action is a purposeful phenomenon in which individuals set and pursue agendas that take them along a behavioral course in which they do social good and contribute to society. And, of course, many of these efforts are coordinated, with the full impact of them felt only as a result of aggregation across individuals.

In this sense, then, the descriptive principles of social action provide a springboard for consideration of the causal or explanatory principles of social action, which help to account for why social action occurs and the forms that it takes. That is, the analysis of descriptive principles of social action suggests that attempts to understand the causes of social action should pay careful attention to considerations of motivation as individuals seek to find ways to take action on behalf of the social good of society. Accordingly, it is with the matter of motivation that we begin our consideration of the explanatory principles of social action.

The Role of Motivation

Across diverse literatures on numerous forms of social action, investigators have adopted a *motivational* perspective, focusing theoretically and empirically on the role of motivations in "disposing" people to take action, in channeling them into particular forms of action, in guiding them through the course of their involvement, and in sustaining their efforts over time. To some extent, this emphasis on motivation reflects the fact that many of the characteristic features of social action appear to be describing and defining *motivated* forms of action—chosen activities that are entered into freely and without obligation, continuing over extended periods of time and even in the face of substantial personal costs, and in which there are multiple routes to reaching the same end of contributing to the social good. In this spirit, researchers have searched for personally based motivations that move people to seek out forms of social action, that lead them to initiate involvement in social action, and that guide and sustain their actions over time.

Volunteerism

One form of social action in which the role of motivation is clearly evident is volunteerism. Every year, in countries around the world, millions upon millions of individuals volunteer their time and efforts to directly help others (Curtis et al., 1992). In the United States alone, it is estimated that 44% of adults volunteer (Independent Sector, 2002); they volunteer for such activities as providing one-to-one companionship to the lonely, tutoring to the illiterate, counseling to the troubled, and health care to the sick. Other volunteers invest their time and effort in political campaigns, advocacy efforts, and other causes, the goals of which are not so much the direct delivery of help and assistance to others but, rather, the improvement of the conditions of life of entire groups. Accordingly, volunteerism represents a noteworthy example of social action. Volunteers contribute their time and effort with the purposes of solving problems faced by their communities, of alleviating the suffering of others, and of generally bettering the human condition. And, indeed, volunteers play critical roles in ameliorating such problems as hunger and poverty, illiteracy, and alcohol and drug abuse, as well as in advancing the agendas of various social movements and political causes such as environmental action and human rights movements (see Wilson, 2000, for a review of research on volunteerism).

Research on volunteerism has examined the motivations that dispose and sustain involvement in this form of social action (e.g., Clary & Orenstein, 1991; Clary et al., 1998; Davis, Hall, & Meyer, 2003; Davis et al., 1999; Omoto & Snyder, 1995; Penner & Finkelstein, 1998; Piliavin, 2005; Simon, Stuermer, & Steffens, 2000). Much of this research has been guided by *functionalist* theorizing that emphasizes the purposes served by action and the role of such purposes in initiating, guiding, and sustaining action (e.g., Snyder, 1993; Snyder & Cantor, 1998). In the case of volunteerism, a functional analysis concerns the needs being met, the motives being fulfilled, and the functions being served by engaging in volunteer service (Clary & Snyder, 1991; Omoto & Snyder, 1990; Snyder, Clary, & Stukas, 2000; Snyder & Omoto, 2000).

A central tenet of functionalist theorizing is that different people can and do engage in the very same behaviors to serve quite different psychological functions. According to this logic, acts of volunteerism that are quite similar on the surface may reflect markedly different underlying motivations; that is, they may be serving distinctly different psychological functions. Thus, for example, several individuals may all engage in the same form of volunteerism, say working in a shelter for individuals who are homeless, but do so in the service of quite different motives, with one person doing so to make friends, another to boost a fragile sense of self-worth, and yet another to acquire skills relevant to a career in social services. In accord with this functional principle, research has revealed a diversity of motivations that bring people to volunteerism and that sustain their involvement in volunteerism, including affirming values, enhancing self-esteem, making friends, acquiring skills, and community concern.

Several inventories having been developed to assess motivations for volunteerism, some seeking to measure motivations of generic relevance to volunteerism and

some seeking to measure motivations for specific forms of volunteerism (e.g., Clary et al., 1998; Omoto & Snyder, 1995; Ouellette, Cassel, Maslanka, & Wong, 1995; Reeder, Davison, Gibson, & Hesson-McInnis, 2001; Schondel, Shields, & Orel, 1992). Nevertheless, strong family resemblances exist in the sets of motivations identified in these diverse programs of research, and across diverse ages of the volunteer populations studied (e.g., Okun, Barr, & Herzog, 1998; Omoto, Snyder, & Martino, 2000). As well, the inventories of motivations identified in studies of volunteers have been successfully "translated" to tap the motivations for other forms of social action, such as involvement in community leadership programs (e.g., Bono, Snyder, & Duehr, 2005) and involvement in the political process (e.g., Miller, 2004).

Although there is some variability in the precise numbers of motivations identified for volunteerism, there is a set of recurring motivations. Specifically, it is common for volunteers to express motivations related to personal *values*, including humanitarian concern about others or other personal guiding values, convictions, and beliefs. Another important type of motivation revolves around *community concern*, or the desire to support and assist a specific community of people, whether or not the volunteer considers him- or herself to be a member of that community. Some people volunteer for reasons that are relatively more self-focused. For example, motivations have been identified that include volunteering for *career* reasons, either to bolster career and networking opportunities or to obtain career relevant experiences, and volunteering to gain greater *understanding* or knowledge about a problem, cause, or set of people. Other motivations for volunteerism include *personal development* concerns (e.g., developing skills and testing oneself), ego or *esteem enhancement* (e.g., to feel better about oneself or bring stability to one's life), and *social* concerns (e.g., a desire to build one's social network and to meet new people and make new friends). People may seek out opportunities to volunteer for one or more of these motivations (e.g., Kiviniemi, Snyder, & Omoto, 2002), but the simple point is that they engage in their work in an effort to meet personal and specific needs. These personal needs or motivations differ across persons, but they can also differ within the same individual over time or life circumstances.

A second central tenet of functionalist theorizing in the context of volunteerism is that these motivations are related to more general agendas for action, and that these motivations guide and direct the unfolding course of people's pursuit of these agendas. In fact, in support of this functional principle, research has documented how the motivations that bring people to engage in volunteerism foreshadow and get played out over the entire life history of their service as volunteers, predicting their experiences as volunteers and the outcomes of their volunteer service, including their contributions to the well-being of others and the benefits, costs, and other personal outcomes that accrue to volunteers themselves (e.g., Omoto & Snyder, 1995; Penner & Finkelstein, 1998). When it comes to initiating involvement in

volunteerism, people are particularly likely to become involved when circumstances suggest that engaging in social action can and will serve their own motivations. Thus, in studies of persuasive messages to recruit volunteers, the messages that have been found to be particularly effective are those that target specific motivations of individual recipients—in other words, messages that are fine-tuned to individually based needs and motives (e.g., Clary et al., 1998; Omoto, Snyder, & Smith, 1999; Smith, Omoto, & Snyder, 2001a). Volunteers also gravitate toward tasks with benefits that match their personally relevant motives (e.g., Houle, Sagarin, & Kaplan, 2005).

Moreover, once in volunteer service, matching between motivations, expectations, and experiences is predictive of greater satisfaction and lesser burnout (e.g., Crain, Omoto, & Snyder, 1998; Omoto et al., 2000). Similarly, when it comes time to decide whether to continue to be involved in volunteering, volunteers are particularly likely to do so in circumstances in which their own motivations are being fulfilled (e.g., Clary et al., 1998; O'Brien, Crain, Omoto, & Snyder, 2000; Williamson, Snyder, & Omoto, 2000). This principle of the "matching" of motivations and experiences and the facilitating effects of such matches on events related to the initiation and sustaining of social action have been observed in other social action contexts as well, including studies of the determinants of voting behavior (e.g., Burgess, Haney, Snyder, Sullivan, & Transue, 2000; Lavine & Snyder, 1996). The functional theoretical perspective guiding research on volunteer motivations and the motivations identified in related research guided by it serve as reminders that volunteers act both on behalf of *others* (e.g., volunteering to alleviate the problems of homelessness, poverty, etc.) *and* on behalf of *themselves* (e.g., volunteering to make friends, acquire new skills, and affirm personal values). This joining of concern for others and self is, of course, highly congruent with the nature of social action as individual action taken on behalf of larger societal concerns (i.e., the descriptive principle of "an individual phenomenon, but also a social phenomenon") and the bridging of individual and collective concerns that is characteristic of social action.

Civic and Political Participation

An emphasis on the role of motivations for social action, and on the importance of recognizing a diversity of motives that may be differentially important across individuals who engage in social action, is also evident in research on the motivations for civic engagement and political participation. Thus, in their attempts to understand individuals who had become active in civic and political causes, Verba and colleagues (1995) identified four categories of benefits that people may seek from civic participation: selective *material benefits* (e.g., furthering one's own career), selective *social gratifications* (e.g., being with other people), selective *civic gratification* (e.g., making the community or nation a better place), and *collective outcomes* (e.g., influencing government policy). These benefits clearly span a continuum from outcomes of relatively

specific benefit to individual activists to outcomes that benefit larger collectives of individuals in communities, states, and nations.

Based on interviews with political activists, Teske (1997) has also identified a diversity of motives that may underlie political participation, including the affirmation of one's principles, the good feelings derived from doing the right thing, growth and development as a person, increased self-esteem, and gaining community. Finally, adopting an explicitly functional approach to motivation, Miller (2004) has examined the role of several motives (including value expressive, social, ego defensive, self-interest, and collective interest) in predicting political participation.

Social Movements

The role of motivation has also been central to understanding participation in social movements. It is well known and documented that members of certain social groups (ethnic minorities, women, gays and lesbians, etc.) are often targets of prejudice and discrimination, have limited economic and employment opportunities, and do not enjoy full access to education and health care. Sometimes, attempts to change these disadvantaged conditions take on a collective form, as when people decide that the only way to change these disadvantageous conditions is to join together with other members of their group to take collective action for the good of the group. The activities of social movements can include public discussions, lobbying, petition drives, boycotts, protests, and civil disobedience. The list of social movements is a long one, with movements dedicated to issues of race, class, gender, sexual orientation, peace, the environment, labor, nuclear power, to name but a few, attracting legions of participants around the globe.

Motivational issues have been a central concern in attempts to understand why people participate in social movements and engage in collective action (Klandermans, 1997; Klandermans & Oegema, 1987; Simon et al., 1998, 2000; Stryker, Owens, & White, 2000; Tropp & Brown, 2004). Thus, in his theory of social movements, Klandermans (1997) has proposed three classes of motivation for social movement participation, each originating from different types of expected costs and benefits of participation. The *collective motive* involves the shared benefits that the social movement seeks (e.g., equal rights and higher wages). The *normative motive* involves the expected reactions of significant others to one's participation in a social movement (e.g., approval or disapproval, praise or criticism). And, the *reward motive* involves the individual and personal costs and benefits of participation (e.g., time taken away from work or new friends made in the movement). The role of these three motives has been documented in a variety of social movements, including (among others) the labor and peace movements in the Netherlands (Klandermans, 1984, 1997; Klandermans & Oegema, 1987); for related studies of social movements, see also Simon and colleagues (1998); Stuermer (2000); Stuermer and Simon (2004b); Stuermer, Simon, Loewy, and Joerger (2003).

Organizational Citizenship

The role of motives for participation has also been examined in another domain of social action, namely, organizational citizenship behavior. It has been suggested that for firms and organizations to operate successfully and to encourage employee satisfaction and camaraderie, workers must do more than the formally specified technical aspects of their jobs; toward those ends, it has been suggested that workers engage in prosocial behaviors directed at helping others and the organization itself, phenomena referred to as organizational citizenship behaviors (Borman & Penner, 2001; Brief & Weiss, 2002; Cropanzano & Byrne, 2000; Katz, 1964; Organ, 1988). Most conceptualizations of organizational citizenship behaviors have drawn distinctions between prosocial behaviors that are directed at helping specific individuals or groups within the organization and prosocial behaviors directed at the organization itself (e.g., Organ & Ryan, 1995; Smith, Organ, & Near, 1983). And, a variety of personality, attitudinal, and motivational variables have been linked to organizational citizenship behaviors (for reviews, see Borman, Penner, Allen, & Motowidlo, 2001; Motowidlo, Borman, & Schmitt, 1997; Organ & Ryan, 1995).

Of particular relevance to our current concerns with motives and social action, Penner, Midili, and Kegelmeyer (1997) have conceptualized organizational citizenship as proactive behaviors, consciously chosen by individuals and engaged in to meet their needs and to satisfy their motives. Drawing from functionalist theorizing proposed and developed in the context of volunteerism, Penner and colleagues have suggested that the same acts of organizational citizenship could reflect different motives for different individuals. Guided by and in support of this theorizing, Rioux and Penner (2001) have identified and developed measures of a set of motives for organizational citizenship behaviors—specifically *prosocial values*, *organizational concern*, and *impression management*. Moreover, they have found that prosocial values motives were most strongly predictive of organizational citizenship directed at individuals, whereas organizational concern motives were most strongly associated with organizational citizenship behaviors directed at the organization (see also Finkelstein & Penner, 2004). For additional perspectives on the motivational bases of organizational citizenship behaviors, see Bolino (1999) and Folger (1993).

Summary

Thus, in four domains of social action—volunteerism, civic and political participation, social movements, and organizational citizenship—we have seen the important role that motives play in disposing people to become involved. In each domain, we have seen that there is a diversity of motives potentially in play. Moreover, it appears that the sets of motivations, although they may involve some degree of specificity for particular forms of social action, include some recurring themes and interrelated themes. Specifically, there is clear emphasis on the

importance of recognizing a diversity of potential motivations, as well as the necessity of identifying which motive or motives are particularly salient for which individuals at any given time in promoting and sustaining social action.

The importance of motivations is evident across many of the unfolding stages of participation in social action, as motivations set the stage for and foreshadow the expectations that people form in anticipation of getting involved in social action, their choices to get involved, the experiences that they have while participating, their satisfaction with their participation, and their ultimate decisions to continue their involvement over time or to terminate their participation. That is, motivations set the stage for the agendas for action that individuals pursue in the context of social action.

The Role of Identity

Not only are motivations implicated in social action, so too are there indications of an important role for identity and identification concerns in understanding social action. That is, for some people and under some circumstances, involvement in social action appears to derive from, as well as become a part of, one's identity, part of that which defines who one is, both as a matter of *individual identity* and as a matter of *social identity*. It is, so to speak, not just that one *"does"* social action but that one *"is"* a social actor. Identity concerns are thus another explanatory principle of social action.

Individual Identity

In the case of individual identity, roles and habits have been implicated in the initiation and persistence of social action. Consider the case of blood donation as a form of social action. In their attempts to explain why people become blood donors and especially why some people become regular and habitual blood donors, Piliavin and Callero (1991; Callero, 1985; Callero et al., 1987; Charng, Piliavin, & Callero, 1988; Piliavin, 1989) have noted that for some blood donors, there develops over time a "role person merger" in which what one does as a blood donor becomes a defining part of who one is as a person. This identity as a blood donor is thought to be important in sustaining blood donation over time (see Piliavin, Grube, & Callero, 2002, for further discussion of roles and social action).

Extending this perspective, the role of individual role identity in other forms of social action have also been examined, including the giving of time as a volunteer and the giving of money as a volunteer (e.g., Grube & Piliavin, 2000; Lee, Piliavin, & Call, 1999; Martino, Snyder, & Omoto, 1998), organizational citizenship behavior (e.g., Finkelstein & Penner, 2004; Krueger, 2004), and whistle blowing and other activities of "principled organizational dissent" (Piliavin et al., 2002). In each case, involvement in social action seems to be not just a behavior but also an identity. By extension, and more generally, taking on the identity as a participant in social action may be important for sustaining involvement in social action activities.

Social Identity

At the level of social identity, students of social movements have considered the role of collective identification in motivating and sustaining social movement participation. Across a variety of social movements, *collective identification* with a social movement and the groups that benefit from social movement activities have been found to constitute one pathway to social movement participation, including predicting who becomes involved, extensiveness of involvement, and persistence in social movement activities (e.g.. Simon et al., 1998, 2000; Stuermer & Kampmeier, 2003; Stuermer & Simon, 2004b; Stuermer et al., 2003). Moreover, collective identification predicts social movement participation independently of another pathway to involvement made up of considerations of the rewards and punishments, costs, and benefits associated with social movement participation. For a review of theory on the dual-pathway model of social movement participation and evidence in support of it coming from investigations of a variety of social movements (including those of gay people, old people, and fat people), see Stuermer, 2000, and Stuermer and Simon, 2004a.

In a related vein, there have been examinations of the role that social identity plays as the "social glue" that builds group loyalty and that holds groups together, such that group members develop extremely positive impressions of their groups and stay invested in them even when they could obtain better outcomes by leaving the group (e.g., Van Vugt & Hart, 2004). This state of affairs is one that would facilitate social action within groups, with individual group members translating their social identities as group members into working loyally on behalf of the group and the betterment of its members. In fact, research suggests that those who identify most strongly with their social groups are likely to invest more of their personal resources in the group, work harder for the group, and show greater self-restraint in consuming the group's resources (e.g., Barreto & Ellemers, 2000, 2002; De Cremer & Van Vugt, 1999; Kramer & Brewer, 1984).

Considerations of Self and Considerations of Others

A recurring theme across diverse domains of social action, and captured in another of our explanatory principles, is that considerations of *self* and considerations of *others* are invoked and implicated in the initiation and maintenance of social action.

Volunteerism

In research on the motivations that underlie volunteerism as a form of social action, for example, it has been suggested (e.g., Omoto & Snyder, 1995) that much of the variability in motivations for volunteering is captured in a two-category classification system in which motivations are grouped into those that focus on *others* who are the beneficiaries of volunteerism (e.g., values and community concern motivations) and those that focus on

the *self* and the benefits that accrue to the self from volunteering (e.g., career advancement and esteem enhancement). As well, the distribution of motives for volunteerism into relatively self-serving and relatively other-oriented has been discussed by Bierhoff (2002) and Chambré (1987), and the variation in the balancing of these motives across nations has been examined by Van de Vliert, Huang, and Levine (2004). In a related vein, Miller (1994) has proposed that the moral foundations of caring and helping may vary across cultures, especially with respect to the extent that caring and helping reflect personal and individual considerations versus the extent to which these prosocial actions reflect interpersonal and social obligations.

Social Movements

Similarly, in theory and research on social movement participation (e.g., Stuermer & Simon, 2004a), the multiple pathways to involvement that have been identified tend to group into those that focus on the individual actor (e.g., calculations of personal rewards and costs of joining and participating in a social movement) and those that focus on others (e.g., identification with a larger group that will benefit from the actions of the movement). In fact, research suggests that these different pathways, one focused on the self and one on others, have independent predictive value in accounting for participation in social movements (for a review, see Stuermer & Simon, 2004a).

Social Dilemmas

A similar distinction between a focus on the self and a focus on others is found in theory and research on social dilemmas and their resolution. Specifically, this work has attempted to understand in the context of social dilemmas when and why individuals focus more heavily on the collective concerns of all members of a community or a society and when and why they focus more on their own individual concerns. The focus on collective concerns is a prosocial orientation that tends to lead to greater involvement in collective actions for the common good, whereas the focus on individual concerns seems to be an orientation that tends to inhibit social action on behalf of the greater good, instead leading people to emphasize more selfish considerations of personal gain. In fact, research on social dilemmas, including studies of energy conservation and the use of public transportation (e.g., De Cremer & Van Vugt, 1999; Van Lange, Van Vugt, & De Cremer, 2000; Van Lange, Van Vugt, Meertens, & Rutter, 1998; Van Vugt, Meertens, & Van Lange, 1995), suggests that the resolution of social dilemmas in ways that involve acting on behalf of a collective good stems from prosocial orientations that involve beliefs, feelings, and motivations in which concern for others figures prominently. (For a review of this work on social dilemmas, see Van Lange, 2000.) At the same time, some of the theoretical and empirical literature on social dilemmas also argues that prosocial action is undertaken in the service of selfish concerns about personal gain, both at an individual level (as a manifestation of a form of "Machiavellian intelligence," e.g., Orbell, Morikawa, Hartwig, Hanley, & Allen, 2004) and as a result of evolutionary pressures and concerns (as a manifestation of a "selfish gene" [e.g., Dawkins, 1976], and the evolution of cooperation [e.g., Axelrod, 1984], and reciprocal altruism [e.g., Trivers, 1971]; see also Van Vugt & Van Lange, 2006).

Leadership

Finally, a related distinction between a focus on the self and a focus on others can be found in studies of leadership, specifically the sources of motivations that leaders can use to promote social action. For example, Tyler (2002) has proposed two sources of motivations that leaders can use to promote cooperation in groups: rewards/punishments and attitudes/values. In the case of the former, rewards/punishments can be interpreted as reflecting extrinsic considerations and focus on the other people who have the ability to dispense rewards and punishments. For attitudes/values, meanwhile, the focus is on relatively more intrinsic and self-considerations, including the influence of attributes of the self who holds the guiding attitudes and values. Furthermore, it is suggested that leaders who appeal to attitudes and values (as internal motivations coming from the self) rather than using promises of rewards and threats of punishments will ultimately be more effective in gaining voluntary followers and cooperation within groups (Tyler, 2002; see also Tyler & Blader, 2000). This work is relevant to social action in that leaders can encourage and influence group members to behave cooperatively, generally benefiting the group as a whole and aiding in finding effective solutions to social dilemma problems. In fact, Tyler and his colleagues have examined internal motivations for cooperative behavior and how authorities can gain support and deference from group members through the use of fair procedures (i.e., procedural justice). These procedures convey and foster respect and pride and also inspire identification with the community on the part of group members. Consequently, group members are more likely to accept the decisions of authorities and to show restraint when faced with social dilemmas involving conflicts between individual- and group-serving actions (e.g., Tyler, 2000; Tyler & Degoey, 1995).

Implications for Initiation and Maintenance of Social Action

The importance of recognizing that considerations of self and considerations of others are both involved in social action is underscored by the fact that these considerations may be differentially involved in the initiation and the maintenance of social action. Across diverse domains of social action, there are indications that the factors that are important in accounting for the initiation of social action are not necessarily the same as the factors that are important in accounting for the maintenance of social action. For example, consider the case of volunteerism. Although considerations of values and related other-oriented motivations figure prominently in the motiva-

tions reported by new volunteers, such motivations have relatively little predictive power in accounting for the ultimate duration of service as a volunteer. By contrast, although self-oriented motivations such as esteem enhancement are relatively rarer among the motivations that volunteers claim to have brought them to volunteer in the first place, such motivations have relatively greater predictive power in forecasting just how long volunteers will remain active in service to others (Omoto & Snyder, 1995; but see Penner & Finkelstein, 1998). More generally, it may be that when it comes to understanding the initiation of social action, other-oriented considerations may be important, but when it comes to understanding the maintenance of social action over sustained periods of time, the critically important factors are related to self-oriented considerations (see Snyder, Omoto, & Smith, in press, for further discussion of this point). This contention is speculative at this point; it remains for future research to more fully explore differences between the factors that are crucial to the initiation versus the maintenance of social action, and especially how these factors might be differentially focused on the self or on others.

Implications for the Locus of the Impetus for Social Action

It also should be noted that in drawing the distinction between a focus on the self and a focus on others, it is not a question of where the motivations are located. That is, it is not that the motivations for social action are thought to vary between those that reside within the self and those that reside within others. In all cases, the motivations that lead to social action are properties of the individual social actors and are thought to reside within them; however, what varies is the focus of attention to be on the actor or on others, who often are the targets of the action. Thus, the humanitarian concerns and sense of societal obligation that lead people to try to improve the welfare of others through social action and the desires for social recognition and career advancement that also can motivate social action both reside within the individual actor; however, what varies is whether the driving force behind social action is the quest for benefits to the self or benefits to others through social action.

Implications for the Selfish/Selfless Nature of Social Action

Finally, we point out that there is some conceptual overlap between this social action explanatory principle of considerations of self and considerations of others and distinctions that have been made between "egoism and altruism" or "selfish and selfless" motivations in the literature on helping and prosocial behavior (e.g., Batson, 1991, 1998). To be sure, there does seem to be something selfish and egoistic about social action that is prompted by the quest for benefits to the actor him- or herself; similarly, social action that is undertaken to generate benefits to others and to society at large does seem to have a rather selfless and altruistic quality to it. How-

ever, even social action that is motivated by desires for benefits to the self delivers benefits to others and to society at large and is, in that sense, prosocial and altruistic in its consequences. And, social action that is undertaken out of concern for others and for society at large does deliver benefits to the individual actor as a member of society.

Nevertheless, individual actors faced with decisions about whether or not to get involved in social action may face tensions and trade-offs between potential benefits to self and potential benefits to others. In fact, research on social movements suggests that individuals may calculate the rewards and punishments, costs, and benefits associated with joining a social movement (e.g., Klandermans, 1984, 1997); among the benefits and costs to be weighed are the benefits of the social movement to society should the movement succeed and the costs to the self of the time and energy taken away from other personal pursuits by involvement in the social movement. As well, research on volunteerism indicates that, at the same time as volunteers clearly deliver services of benefit to others, they often experience personal costs, including, at times, being made to feel embarrassed, uncomfortable, or otherwise stigmatized as a result of their volunteer service (e.g., Snyder, Omoto, & Crain, 1999). Furthermore, for a general cost–benefit approach to understanding involvement in volunteer organizations, see Chinman and Wandersman, 1999, and Chinman, Wandersman, and Goodman, 2005.

The Role of Personal Connections to Other People

Theory and research on diverse forms of social action, conducted from diverse conceptual orientations, point to an important role of connections to other people—whether direct and behavioral or indirect and symbolic—in understanding social action. For, connections to other people are involved in prompting people to become involved in social action and in sustaining their continued involvement in social action, and particularly social action that directly benefits others. Although connections to others can take many forms, we focus on research on empathy and common group memberships, two of the more extensively researched forms of connections to other people, as exemplars of this explanatory principle of social action.

Empathy

The connections between individuals that facilitate social action may be feelings of empathy and bonds of identification. That is, to the extent that individuals feel empathy and identify with other people, they become more likely to engage in various forms of social action that potentially benefit those with whom they empathize and those with whom they identify. The important role of empathy has been revealed in studies of prosocial action involving one-to-one helping and volunteerism.

Researchers in personality, social, and developmental psychology have provided evidence for the role of empathy—an emotional reaction including feelings of

compassion, concern, and tenderness—in spurring individuals to help people in need. For example, to the extent that an individual feels empathy for another person, he or she will be more likely to offer help (e.g., Batson, 1998; Dovidio, Allen, & Schroeder, 1990; Schroeder, Dovidio, Sibicky, Mathews, & Allen, 1988), even in situations in which helping is relatively demanding and may not bring direct benefits to the helper.

Moreover, people who have a general tendency to react to the plight of other people with feelings of empathy (i.e., people who are high in *dispositional* empathy) are particularly likely to help others (Davis, 1983, 1996; Davis et al., 1999; Penner, Fritzsche, Craiger, & Freifeld, 1995). More specifically, people who are generally disposed to feel empathy are also more likely to feel empathy for a specific person in need, which, in turn, fosters helping and related forms of social action (Carlo, Eisenberg, Troyer, Switzer, & Speer, 1991; Davis et al., 2003; Eisenberg et al., 1989).

The role of empathic connections with other people in promoting and sustaining involvement in social action has been examined in the specific context of volunteerism, where Davis has traced the role of empathy over the course of volunteering (for a review, see Davis, 2005). In his work, he has shown that individual differences in empathy are related to initial decisions to volunteer and to the subjective experiences of volunteers during their service (e.g., Davis et al., 1999), although less so to sustaining their involvement over time (e.g., Davis et al., 2003).

Common Group Membership

While empathic connections seem to facilitate the provision of help, and perhaps other forms of social action, it also seems to be the case that the effects of empathic connections with others may depend on whether those others are members of one's own ingroup or whether they are members of an outgroup. Thus, for example, in research on volunteerism, it has been demonstrated, in both field and laboratory studies, that *empathy* is a critical ingredient in accounting for volunteering to help ingroup members whereas another type of connection, *liking*, is a critical ingredient in accounting for volunteering to help outgroup members (e.g., Stuermer & Siem, 2005; Stuermer, Snyder, Kropp, & Siem, 2005; Stuermer, Snyder, & Omoto, 2005).

More generally, there are growing indications that quite different motivations are involved in helping ingroup and outgroup members. For example, although people typically help ingroup members because of their identification with the common ingroup and a concern for their collective welfare (e.g., Simon et al., 2000), outgroup helping often reflects self-serving considerations, including the desire to appear unprejudiced or to avoid feeling guilty (Dutton & Lake, 1973; Gaertner & Dovidio, 1977).

Although not necessarily conducted in social action contexts, other research similarly suggests that people may have different reasons for helping ingroup and outgroup members, and especially when group member-

ships are made salient by contextual cues, structural differences, or histories of asymmetrical intergroup relations. For example, changing group conceptualizations or boundaries so that former outgroup members are now viewed as members of one's ingroup leads to increases in helping those members (Dovidio et al., 1997). The provision of help to others might also be a means for asserting group dominance or an attempt to renegotiate the power structure between groups (e.g., Nadler, 2002). Additional research is needed to carefully examine both of these intriguing possibilities. For now, our point is that there is suggestive evidence that group memberships and conceptualizations of ingroups and outgroups may play important roles in determining when, why, and on whose behalf social action is enacted.

Other Personal Connections

Finally, there is considerable support from research derived from a social exchange perspective for the principle that connections to others facilitate social action. Illustrative of the findings in this domain, it has been frequently demonstrated that increases in the social action tendency to cooperate are promoted by stronger ties between the parties to the exchange (e.g., Granovetter, 1973; Macy & Skvoretz, 1998; Orbell, van de Kragt, & Dawes, 1988).

The Role of Community

In addition to the importance of personal connections with other people, there are growing indications, across diverse forms of social action, that a sense of *connection with a larger community*, including a psychological sense of community (e.g., McMillan, 1996; McMillan & Chavis, 1986; Sarason, 1974), can facilitate social action. This explanatory principle has been revealed in studies on the resolution of social dilemmas, volunteerism, and civic participation.

Social Dilemmas

For example, it has been demonstrated that fostering a community orientation can promote the resolution of social dilemmas associated with, among other things, the use of public transportation and energy conservation (e.g., De Cremer & Van Vugt, 1999; Ostrom, 1990; Schlager, 2002; Van Vugt, 2002; Van Vugt & Samuelson, 1999), thereby averting the "tragedy of the commons" first described by Hardin (1968). Similarly, Tyler (2000; Tyler & Blader, 2000) has observed that promoting identification with and commitment to community can promote collective solutions to social dilemmas and compliance with authorities in regulating social dilemmas.

Volunteerism

Similarly, research on the processes of volunteerism has yielded recurring indications that connection to community can draw people into volunteerism and sustain their involvement over time. In reciprocal fashion, moreover,

involvement in volunteerism seems to strengthen and build connections to community. Specifically, community concern and the influences of other community members figure prominently in the motivations of new volunteers (Omoto & Snyder, 1995; Stuermer & Kampmeier, 2003). Moreover, over the course of their service, volunteers become increasing connected with their surrounding communities, including the communities defined by the volunteers, staff, and clients associated with their volunteer service organizations (Omoto & Snyder, 2002). And, their effectiveness as volunteers is enhanced by a sense of connection to a relevant community (Omoto & Snyder, 2002).

Reversing the causal order, volunteering also appears to build and foster a sense of community. For example, as a consequence of their work, volunteers are increasingly surrounded by a community of people who are somehow connected to their volunteer service, including people they have recruited to be volunteers (Omoto & Snyder, 2002). In addition, volunteering can and does contribute to the creation of bonds of social capital (e.g., Stukas, Daly, & Cowling, 2005), and it even has been considered a central indicator or measure of social capital itself (Putnam, 2000). Research also suggests that as connections to a community of shared concerns increase, participation in the community, including in forms of social action other than volunteerism (such as giving to charitable causes, attending fund-raising events, and engaging in social activism), also increases (Malsch, 2005; Omoto & Malsch, 2005; Omoto & Snyder, 2002).

Civic Participation

Furthermore, beyond social dilemmas and volunteerism, cross-sectional research has demonstrated positive associations between sense of community and civic participation. Thus, for example, individuals who report a stronger psychological sense of community are more likely to be registered voters and to be active in their neighborhoods (Brodsky, O'Campo, & Aronson, 1999), to engage in neighboring behaviors such as lending their neighbors food or tools (Kingston, Mitchell, Forin, & Stevenson, 1999), and to participate in community organizations (Chavis & Wandersman, 1990; Perkins & Long, 2002; Wandersman, 1980; Wandersman, Florin, Friedmann, & Mier, 1987) and political activities (Davidson & Cotter, 1989). Moreover, bonds of connection within communities and the social capital associated with them have been implicated in the provision of public goods (Anderson, Mellor, & Milyo, 2004), the reduction of crime within communities (Saegert, Winkel, & Swartz, 2002), and the promotion of the health of community members (Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997). Finally, residential stability has been implicated in identification with one's community, which in turn manifests itself in diverse forms of helping behaviors, procommunity involvement, collective efficacy, and social action (Kang & Kwak, 2003; Kasarda & Janowitz, 1974; Oishi et al., 2005; Sampson, Raudenbush, & Earls, 1997).

Thus, there is substantial evidence that connections to a community promote social action, often for the benefit of the community, as well as broader involvement in society. However, we also note that there are hints of an association between individualism and broad indicators of social capital. Thus, in the United States, the states with higher levels of social capital, as indexed by greater amounts of civic engagement and political participation, are also the most individualistic; similarly, there is a positive association between individualism and social capital across different countries (Allik & Realo, 2004). These associations may suggest that as much as bonds of connection with community may promote social action, the apparent liberation from social bonds that may come with individualistic cultural views may also make people dependent on being or staying involved with society (as suggested over a century ago by Durkheim, 1893/1984; see also Sarason, 1974). For further discussion of, and evidence for, cross-cultural variation in social capital and bonds of trust and reciprocity, see Buchan, Croson, and Dawes (2002); Fukuyama (1995); and Yamagishi and Yamagishi (1994).

As reviewed in this section, then, connections to communities (and often psychological communities rather than geographically defined ones; Omoto & Malsch, 2005; Omoto & Snyder, 2002) are related to diverse forms of social action. In fact, it appears likely that there is a cyclical process at work here, one in which connections to community lead individuals to engage in social action that, in turn, further builds community connections and social capital. As a result of this self-perpetuating and accretionary process, social action becomes more likely and sense of community is increased. More generally, it may be that social action begets social action *via* a mechanism or principle of community such that one of the more significant consequences of social action is the creation and perpetuation of a culture of service, participation, and involvement.

Summary

To summarize, we have described a set of explanatory principles of social action that provide explanations for why social action occurs and that give some indications for why social action takes the forms it does. Specifically, we began by examining research from diverse domains that illustrates motivational approaches to social action, and while motivational concerns were central in each of these domains, there was not perfect agreement on the specific motivations most critical for understanding and predicting social action. Next, we discussed how individual and social identities seem to be related to the initiation, persistence, and some of the effects of social action activities. Another explanatory principle that we identified, and one for which there is precedent for and links to other literature, is the relative importance of considerations of the self relative to considerations of others. Finally, we described two types of connections that appear to be important for social action, namely, personal connections to specific others (including empathy and common group memberships) and connections to broader, more diffuse communities. To the extent that connections are salient or strong—whether they are cog-

nitive, emotional, or behavioral connections—social action, and especially helping, appears likely to result.

AN INTEGRATIVE MODEL OF SOCIAL ACTION

Thus, at this point, we have accomplished our goals of providing both descriptive and explanatory principles of social action. Our set of descriptive principles are useful for characterizing and comparing different forms of social action, as well for identifying their dimensions of variability and indicating what is to be included in and what is to be excluded from the category of social action. Meanwhile, our set of explanatory principles helps us to understand how and why social action occurs, including the factors that provide the impetus for social action, that guide the forms it takes, and that influence how it unfolds over time. To conclude this chapter, then, we now offer a heuristic model of social action, a model that we believe will be useful for organizing and perhaps guiding future work on these important and wide-ranging topics. As with our principles, we see this model as a starting point, as an initial attempt to bring order to a diverse set of topics and principles. The ultimate utility of this model can and will be determined in future theoretical and empirical work.

In reflecting on the principles we have identified for characterizing and understanding social action, it is noteworthy and impressive that these principles recur across research in diverse domains, conducted by researchers working in different countries, coming from a multiplicity of disciplines, and guided by a range of theoretical predilections and perspectives. But, perhaps these recurring themes are no accident. The principles may recur because they may reflect the operations of a common, integrative heuristic model for theory, research, and application on social action. In this model of social action, which builds on and extends earlier models of volunteerism as one form of social action (e.g., Omoto & Snyder, 1995, 2002; Snyder et al., in press), social action is conceptualized as a process that unfolds over time and that can be viewed from multiple levels of analysis. As such, there is both a "horizontal" and a "vertical" organization to this heuristic model of social action.

Horizontal and Vertical Organization of the Model

With respect to the *horizontal* organization or structure of the model, one of the key features of social action is that it unfolds over time. That is, there are the *antecedents* of social action, which specify those factors that bring people to social action, including the personal motivations that dispose people to get involved and the social circumstances that call for and prompt action. Then, there is the social action itself (which may itself involve sustained, ongoing, recurring action) and the *experiences* of those involved in social action, including the perceived and actual costs and benefits that accrue to social actors involved in social actions. Finally, there are the *consequences* of social action for those who take social action, for those who are affected by it, and for society at large.

This horizontal structure defines the unfolding history of social action, taking us through three successive stages in the processes of social action, allowing us to see what leads to social action, what happens during social action, and what results from social action.

With respect to the *vertical* organization or structure of the model of social action, there are various levels of analysis that run from the individual to the cultural, such that at each stage of the processes of social action, it is possible to articulate principles of social action from the perspective of the individuals involved in social action, the interpersonal relationships of which they are a part, the groups and organizations to which they belong, and the larger community, societal, and cultural contexts in which social action takes place.

In conceptualizing social action, therefore, we find it useful to view it from the vantage point of multiple levels of analysis. At an *individual* level, the model calls attention to the activities and psychological processes of individual actors and the recipients or targets of social action. At an *interpersonal* level, many forms of social action occur in the context of or because of people's interpersonal relationships. At a *group* level, many forms of social action are enacted by groups and collectives, take place in the context of community-based organizations and institutions, or are carried out for the purpose of addressing the conditions of members of groups and institutions. Finally, at a broader *societal* level, the model considers the linkages between individuals and their societies as well as cultural dynamics associated with the emergence and evolution of traditions of social action and community involvement.

Putting the horizontal and vertical structures of the model together results in a "matrix," depicted in Figure 41.1. The columns of this matrix represent stages of the process of social action and the rows reflect the different levels of analysis or perspectives that can be brought to bear on understanding social action.

Thus, with respect to the first (antecedents) stage of the process of social action, which addresses the question of what brings people to social action, diverse features of persons (e.g., their motivations and reasons for participation, their helping dispositions or personalities, and their identities as social actors) are the focus and the search is for relevant traits, motives, and characteristics that dispose people to take social action. In addition, interpersonal relationships and social networks, as well as groups and formal organizations, can and do play important roles in recruiting people to participate in social action. Moreover, differing cultural orientations may set the stage for determining how certain forms of social action will be construed, including whether causes will be judged "worthy" of action. In fact, these cultural orientations are likely to influence whether solving a particular societal problem is seen as a matter of personal choice and individual responsibility or whether it is construed as one of normative obligation and collective concern.

Similarly, with respect to the second (experiences) stage of the process of social action, which focuses on the question of what happens during social action, there is value in regarding social action in terms of the pursuit of

Level of Analysis	Stages of Social Action		
	Antecedents	Experiences	Consequences
Individual			
Interpersonal			
Group			
Societal			

FIGURE 41.1. A heuristic model of social action.

agendas. Thus, not only do individuals have and pursue goals, but agendas for social action can also be seen at other and multiple levels of analysis, including those pursued by and affecting targets of social action, the separate and shared group memberships of social actors and targets, the organizations that support social action, and the larger community and society.

And, finally, at the third (consequences) stage of the process of social action, which is concerned with the question of what results from social action, social action can and does have outcomes and consequences for those who engage in it, as well as consequences for other people, including relationship partners, groups, organizations, communities, and the larger society. Consideration of these consequences, in fact, brings us back to where we started in our analysis, that is, to the realization that social action involves people working alone and working together, often in attempts to benefit themselves and society.

Generativity of the Model

This model, it should be emphasized, is not a theory of social action per se but a broad framework for identifying conceptual issues for empirical investigation. And, by virtue of its deliberate generality, it is applicable to many (if not most) forms of social action. Moreover, we view it as a generative model; that is, it provides a way of asking questions about *moderator*s and *qualifiers* of the principles of social action and it can point the way toward identifying unstudied and understudied questions that have the potential to generate *extensions* of scientific understanding of social action.

Many of the moderators, qualifiers, and extensions can be identified in the various "rows" of the model and may operate to change the unfolding processes of social action. Briefly, and to give a feel for how this model may help to organize and generate knowledge of social action, let us consider some representative examples of moderators at the different levels of analysis. First, at the level of individual social actors, differences in personality (e.g., Davis et al., 1999), interpersonal orientation (e.g.,

Lavine & Snyder, 1996), and age (e.g., Boling, 2005; Okun & Schultz, 2003; Omoto et al., 2000) appear to moderate the processes of social action. At the levels of interpersonal relationships and group memberships, there are indications that ingroup versus outgroup status importantly moderates the resolution of social dilemmas (e.g., Dawes & Messick, 2000), the giving of aid to a person in need (e.g., Levine, Prosser, Evans, & Reicher, 2005), and the processes of volunteerism (e.g., Stuermer, Snyder, & Omoto, 2005). Finally, at the broader level of societies and cultures, there are growing indications of important differences between societies and cultures in the construal of helping, social participation, and civic engagement; these differences, moreover, may moderate the forms that social action takes across cultures (e.g., Allik & Realo, 2004; Buchan et al., 2002; Miller, 1994; Van de Vliert et al., 2004).

Taken together, then, the moderators identified in the rows of the model may involve features of the persons who engage in social action (including individual differences in personality, attitudes, values, and motivations), features of the social circumstances and environments in which these persons function (including their memberships in relationships, groups, and organizations), and the features of their surrounding communities that either promote or impede social action (including their culturally based conceptions of the nature of involvement and participation in society).

In addition to the moderators that can be defined and situated in the "rows" of our heuristic model of social action, there are also moderators that can be identified across the model's "columns." These moderators can be thought of as specifying differences in the ways that constructs defined with respect to the rows of the matrix operate in social action. One such illustrative moderator is the distinction between the initiation and the maintenance of social action and the findings from research on various forms of social action that the factors that dispose people to become involved in social action in the first place may be quite different from the factors that sustain involvement in social action over time (e.g., Omoto & Snyder, 1995). Moving across the columns, ad-

ditional moderators or qualifiers emerge, including intriguing (and as yet untested) possibilities about the ways in which societal consequences of social action produce changed social and cultural environments for future social action—in short, the ways by which social actors and their coordinated social actions change society.

Finally, over and above the ways that the heuristic model of social action helps to organize the existing research on social action and to make clearer the operation of the different principles of social action that we have identified in this chapter, an additional benefit of our proposed model is its generativity in uncovering topics for future research. That is, the "cells" of the model (defined by the conjoint vertical and horizontal structures) that are relatively "underpopulated" in terms of theoretical statements and research findings would seem to be excellent candidates for future theoretical inquiry and empirical investigation. The body of work on social action, in all of its diverse forms, can only benefit and grow from such attention.

Challenges and Prospects

Our heuristic model of social action also serves to underscore some of the challenges posed by the study of social action (see also Omoto, 2005b). For example, to the extent that investigators seek to understand social action *in vivo*, they may find themselves chronically "behind the curve," only able to identify and study instances of social action (e.g., the emergence of social movements in response to problems in society and the formation of grassroots community organizations to deliver needed volunteer services) *after* they have already occurred. In such circumstances, investigators may find themselves shut out of the early stages of the processes of social action, unable, for example, to study the features of individual social actors *before* they become involved in social action.

Moreover, because involvement in social action may extend over long periods, practical constraints associated with following research participants over time may make it difficult to follow the unfolding dynamics of social action to their natural conclusions. Accordingly, the research literature on social action may be more likely to include studies of short-term rather than long-term social action. Nevertheless, studies of long-term social action may be well worth the effort, especially because they are likely critical in establishing a relatively complete and comprehensive understanding of these important phenomena.

In addition, a focus on studying instances of social action as they actually occur may lead to an overrepresentation of studies of successful instances of social action, as ones that fail to meet their goals may not stay in existence long enough to capture the attention of researchers. Yet, a full understanding of social action will surely need to come to grips with the differences between those instances of social action that succeed and those that fail, as challenging as it may be to identify and investigate instances of failed social action.

CONCLUSION

Reflecting on social action, we cannot help but be struck by the recurring themes and principles, both descriptive and explanatory, that emerge from investigations of diverse domains of social action and their ability to be encompassed in a broad and integrative model. The study of social action simultaneously brings into sharp relief, and also connects, research on a wide range of topics, from volunteerism to social movements to political participation to social dilemmas. In addition, and particularly striking, are the bridges that theoretical and empirical work on social action builds between the psychology of *individuals*, their concerns for the *interpersonal relationships* of which they are part and the *groups* to which they belong, and their participation in their larger *societies*.

As such, the study of social action provides new perspectives on the linkages between individuals and society, highlighting the mutual interplay and influence of individuals and the larger society and opening the door for theoretical, as well as practical, advances across a broad spectrum of social scientific domains. In the end, though, it is through their involvement in social action that individuals contribute to the functioning of society, and the changes wrought in society by social action affect the lives of its individual members.

Moreover, in the study of social action, researchers can engage in diverse forms of "action research" (e.g., Chein, Cook, & Harding, 1948; Lewin, 1946, 1947; Sanford, 1970). They can intertwine basic and applied research foci and methodologies in the interests both of theoretical advance and of solving pressing social problems. In addition, they can engage real people in real contexts who are seeking real solutions to the real challenges they face. Thus, just as social action connects individuals and society, the study of social action holds vast potential for connecting psychological science, social policy, and public good. This end is exciting, indeed, and, certainly worthy of vigorous pursuit.

REFERENCES

Allik, J., & Realo, A. (2004). Individualism–collectivism and social capital. *Journal of Cross-Cultural Psychology*, *35*, 29–49.

Anderson, L. R., Mellor, J. M., & Milyo, J. (2004). *Social capital and contributions in a public goods experiment.* Unpublished manuscript, College of William and Mary, Williamsburg, VA.

Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Barreto, M., & Ellemers, N. (2000). You can't always do what you want: Social identity and self-presentational determinants of the choice to work for a low-status group. *Personality and Social Psychology Bulletin*, *26*, 629–639.

Barreto, M., & Ellemers, N. (2002). The impact of respect versus ne-

glect of self identities on identification and group loyalty. *Personality and Social Psychology Bulletin*, *28*, 629–639.

Batson, C. D. (1991). *The altruism question: Toward a social psychological answer*. Hillsdale, NJ: Erlbaum.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 282–316). New York: McGraw-Hill.

Berkman, L. F., & Glass. T. (2000). Social integration, social networks, and health. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 137–173). New York: Oxford University Press.

Bierhoff, H. W. (2002). *Prosocial behavior*. London: Psychology Press.

Boling, A. (2005). *Motivations for volunteerism over the lifespan*. Unpublished doctoral dissertation, Claremont Graduate University, Claremont, CA.

Bolino, M. C. (1999). Citizenship and impression management: Good soldiers or good actors? *Academy of Management Review*, *24*, 82–98.

Bono, J., Snyder, M., & Duehr, E. (2005). *Types of community involvement: The role of personality and motives*. Paper presented at the annual meetings of the Society for Personality and Social Psychology, New Orleans, LA.

Borman, W. C., & Penner, L. A. (2001). Citizenship performance: Its nature, antecedents, and motives. In B. W. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 45–61). Washington, DC: American Psychological Association.

Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, *9*, 52–69.

Boyte, H. C., & Kari, N. N. (1996). *Building America: The democratic promise of public work*. Philadelphia: Temple University Press.

Brief, A. P., & Weiss, H. M. (2002). Organizational behavior: Affect in the workplace. *Annual Review of Psychology*, *53*, 279–307.

Bringle, R. G., & Duffy, D. K. (Eds.). (1998). *With service in mind: Concepts and models for service learning in psychology*. Washington, DC: American Association of Higher Education.

Brodsky, A. E., O'Campo, P. J., & Aronson, R. E. (1999). PSOC in community context: Multi-level correlates of a measure of psychological sense of community in low-income, urban neighborhoods. *Journal of Community Psychology*, *27*, 659–679.

Brown, S. L., House, J. S., Brown, R. M., & Smith, D. M. (2004). *Coping with spousal loss: The buffering effects of giving social support to others*. Unpublished manuscript, Institute for Social Research, University of Michigan.

Brown, S. L., Nesse, R. M., Vinokur, A. D., & Smith, D. M. (2003). Providing social support may be more beneficial than receiving it: Results from a prospective study of mortality. *Psychological Science*, *14*, 320–327.

Buchan, N. R., Croson, R. T. A., & Dawes, R. M. (2002). Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology*, *108*, 168–206.

Burgess, D., Haney, B., Snyder, M., Sullivan, J., & Transue, J. (2000). Rocking the vote: Using personalized messages to motivate voting among young adults. *Public Opinion Quarterly*, *64*, 29–52.

Callero, P. L. (1985). Role-identity salience. *Social Psychology Quarterly*, *48*, 203–214.

Callero, P. L., Howard, J. A., & Piliavin, J. A. (1987). Helping behavior as role behavior: Disclosing social structure and history in the analysis of prosocial action. *Social Psychology Quarterly*, *50*, 247–256.

Carlo, G., Eisenberg, N., Troyer, D., Switzer, G., & Speer, A. L. (1991). The altruistic personality: In what contexts is it apparent? *Journal of Personality and Social Psychology*, *61*, 450–458.

Chambré, S. M. (1987). *Good deeds in old age: Volunteering by the new leisure class*. Lexington, MA: Lexington Books.

Chambré, S. M. (1989). Kindling points of light: Volunteering as public policy. *Nonprofit and Voluntary Sector Quarterly*, *18*, 249–268.

Charng, H., Piliavin, J. A., & Callero, P. L. (1988). Role identity and reasoned action. *Social Psychology Quarterly*, *51*, 303–317.

Chavis, D. M., & Wandersman, A. (1990). Sense of community in the urban environment: A catalyst for participation and community development. *American Journal of Community Psychology*, *17*, 119–125.

Chein, I., Cook, S. W., & Harding, J. (1948). The field of action research. *American Psychologist*, *3*, 43–50.

Chinman, M., & Wandersman A. (1999). The benefits and costs of volunteering in community organizations: Review and practical implications. *Nonprofit and Voluntary Sector Quarterly*, *28*, 46–64.

Chinman, M. J., Wandersman, A., & Goodman, R. M. (2005). A benefit-and-cost approach to understanding social participation and volunteerism in multilevel organizations. In A. M. Omoto (Ed.), *Processes of community change and social action* (pp. 105–125). Mahwah, NJ: Erlbaum.

Clary, E. G., & Orenstein, L. (1991). The amount and effectiveness of help: The relationship of motives and abilities to helping behavior. *Personality and Social Psychology Bulletin*, *17*, 58–64.

Clary, E. G., & Snyder, M. (1991). A functional analysis of altruism and prosocial behavior: The case of volunteerism. In M. Clark (Ed.), *Review of personality and social psychology* (Vol. 12, pp. 119–148). Newbury Park, CA: Sage.

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J. T., Stukas, A. A., Haugen, J. A., et al. (1998). Understanding and assessing the motivations of volunteers: A functional approach. *Journal of Personality and Social Psychology*, *74*, 1516–1530.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.

Crain, A. L., Omoto, A. M., & Snyder, M. (1998, April). *What if you can't always get what you want?: Testing a functional approach to volunteerism*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Cropanzano, R., & Byrne, Z. S. (2000). Workplace justice and the dilemma or organizational citizenship. In M. Van Vugt, M. Snyder, T. Tyler, & A. Biel (Eds.), *Cooperation in modern society: Promoting the welfare of communities, states, and organizations* (pp. 142–161). London: Routledge

Curtis, J. E., Grabb, E., & Baer, D. (1992). Voluntary association membership in fifteen countries: A comparative analysis. *American Sociological Review*, *57*, 139–152.

Davidson, W. B., & Cotter, P. R. (1991). Sense of community and political participation. *Journal of Community Psychology*, *17*, 119–125.

Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality*, *51*, 167–184.

Davis, M. H. (1996). *Empathy: A social psychological approach*. Boulder, CO: Westview.

Davis, M. H. (2005). Becoming (and remaining) a community volunteer: Does personality matter? In A. M. Omoto (Ed.), *Processes of community change and social action* (pp. 67–82). Mahwah, NJ: Erlbaum.

Davis, M. H., Hall, J. A., & Meyer, M. (2003). The first year: Influences on the satisfaction, involvement, and persistence of new community volunteers. *Personality and Social Psychology Bulletin*, *29*, 248–260.

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality*, *67*, 469–503.

Dawes, R. M., & Messick, D. M. (2000). Social dilemmas. *International Journal of Psychology*, *35*, 111–116.

Dawkins, R. (1976). *The selfish gene*. Oxford, UK: Oxford University Press.

De Cremer, D., & Van Vugt, M. (1999). Social identification effects in social dilemmas: A transformation of motives. *European Journal of Social Psychology*, *29*, 871–893.

de Tocqueville, A. (1969). *Democracy in America* (G. Lawrence,

Trans., & J. P. Mayer, Ed.). New York: Anchor Books, Double-day. (Original work published 1835)

Dovidio, J. F., Allen, J. L., & Schroeder, D. A. (1990). Specificity of empathy-induced helping: Evidence for altruistic motivation. *Journal of Personality and Social Psychology*, *59*, 249–260.

Dovidio, J. F., Gaertner, S. L., Validzic, A., Matoka, K., Johnson, B., & Frazier, S. (1997). Extending the benefits of recategorization: Evaluation, self-disclosure, and helping. *Journal of Experimental Social Psychology*, *33*, 401–420.

Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2006). *The social psychology of prosocial behavior*. Mahwah, NJ: Erlbaum.

Durkheim, E. (1984). *The division of labor in society*. London: Macmillan. (Original work published 1893)

Dutton, D. G., & Lake, R. A. (1973). Threat of own prejudice and reverse discrimination in interracial situations. *Journal of Personality and Social Psychology*, *28*, 94–100.

Eisenberg, N., Miller, P. S., Schaller, M., Fabes, R. A., Fultz, J., Shell, R., et al. (1989). The role of sympathy and altruistic personality traits in helping: A reexamination. *Journal of Personality*, *57*, 41–67.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*, 203–210.

Finkelstein, M., & Penner, L. A. (2004). Predicting organizational citizenship behaviour: Integrating functional and role identity approaches. *Personality and Social Behavior*, *32*, 383–398.

Folger, R. (1993). Justice, motivation, and performance beyond role requirements. *Employee Responsibility and Rights Journal*, *6*, 239–248.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Free Press.

Gaertner, S. L., & Dovidio, J. F. (1977). The subtlety of white racism, arousal, and helping behavior. *Journal of Personality and Social Psychology*, *35*, 691–707.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360–1380.

Grube, J., & Piliavin, J. A. (2000). Role-identity, organizational experiences, and volunteer performance. *Personality and Social Psychology Bulletin*, *26*, 1108–1119.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Houle, B. J., Sagarin, B. J., & Kaplan, M F. (2005). A functional approach to volunteerism: Do volunteer motives predict task preference. *Basic and Applied Social Psychology*, *27*, 337–344.

House, J. (2001). Social isolation kills, but how and why? *Psychosomatic Medicine*, *63*, 273–274

House, J. S., Robbins, C., & Metzner, H. L. (1982). The association of social relationships and activities with morality: Prospective evidence from the Tecumseh Community Health Study. *American Journal of Epidemiology*, *116*, 123–140.

Independent Sector. (1988). *Giving and volunteering in the United States: Findings from a national survey*. Washington, DC: Author.

Independent Sector. (1999). *Giving and volunteering in the United States: Findings from a national survey*. Washington, DC: Author.

Independent Sector. (2002). *Giving and volunteering in the United States: Findings from a national survey*. Washington, DC: Author.

Kang, N., & Kwak, N. (2003). A multilevel approach to civic participation: Individual length of residence, neighborhood residential stability, and their interactive effects with media use. *Communication Research*, *30*, 80–106.

Kasarda, J., & Janowitz, M. (1974). Community attachment in mass society. *American Sociological Review*, *39*, 328–339.

Katz, D. (1964). The motivational basis of organizational behavior. *Behavioral Science*, *9*, 131–133.

Kawachi, I., & Berkman, L. F. (2000). Social cohesion, social capital, and health. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 174–190). New York: Oxford University Press.

Kawachi, I., Kennedy, B. P., Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health*, *87*, 1491–1498.

Kingston, S., Mitchell, R., Forin, P., & Stevenson, J. (1999). Sense of community in neighborhoods as a multi-level construct. *Journal of Community Psychology*, *27*, 384–394.

Kiviniemi, M. T., Snyder, M., & Omoto, A. M. (2002). Too many of a good thing? The effects of multiple motivations on task fulfillment, satisfaction, and cost. *Personality and Social Psychology Bulletin*, *28*, 732–743.

Klandermans, B. (1984). Mobilization and participation: Social psychological explanations of resource mobilization theory. *American Sociological Review*, *49*, 583–600.

Klandermans, B. (1997). *The social psychology of protest*. Oxford, UK: Blackwell.

Klandermans, B., & Oegema, D. (1987). Potentials, networks, motivations, and barriers: Steps toward participation in social movements. *American Sociological Review*, *52*, 519–531.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*, 1044–1057.

Krueger, T. (2004). *The influence of an organizational citizen role identity on sustained organizational citizenship behavior*. Unpublished doctoral dissertation, University of South Florida.

Kymlicka, W., & Norman, W. (1994). Return of the citizen: A survey of recent work on citizen theory. *Ethics*, *104*, 352–381.

Lavine, H., & Snyder (1996). Cognitive processing and the functional matching effect in persuasion: The mediating role of subjective perceptions of message quality. *Journal of Experimental Social Psychology*, *32*, 580–604.

Lee, L., Piliavin, J. A., & Call, Y. R. A. (1999). Giving money, time, and blood: Similarities and differences. *Social Psychology Quarterly*, *62*, 276–290.

Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, *31*, 443–453.

Levine, R. V., Norenzayan, A., & Philbrick, K. (2001). Cross-cultural differences in helping strangers. *Journal of Cross-Cultural Psychology*, *32*, 543–560.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, *2*, 34–46.

Lewin, K. (1947). Group decision and social change. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology* (pp. 330–341). New York: Holt.

Macy, M. W., & Skvoretz, J. (1998). The evolution of trust and cooperation between strangers: A computational model. *American Sociological Review*, *63*, 638–660.

Malsch, A. M. (2005). *Prosocial behavior beyond borders: Understanding a psychological sense of global community*. Unpublished doctoral dissertation, Claremont Graduate University, Claremont, CA.

Martino, S. C., Snyder, M., & Omoto, A. M. (1998, April). *Predicting sustained volunteer service: Intentions and identity concerns*. Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

McMillan, D. W. (1996). Sense of community. *Journal of Community Psychology*, *24*, 315–325.

McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition of theory. *Journal of Community Psychology*, *14*, 6–23.

Miller, J. G. (1994). Cultural diversity in the morality of caring: Individually oriented versus duty-based interpersonal moral codes. *Cross-Cultural Research*, *28*, 3–39.

Miller, J. M. (2004). *What motivates political participation?* Paper presented at the annual meetings of the American Political Science Association, Chicago.

Motowidlo, S. J., Borman, W. C., & Schmitt, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, *10*, 71–84.

Musik, M. A., Herzog, A. R., & House, J. S. (1999). Volunteerism and mortality among older adults: Findings from a national sample. *Journal of Gerontology*, *54B*, S173–S180.

Nadler, A. (2002). Inter-group helping relations as power relations: Maintaining or challenging social dominance between groups through helping. *Journal of Social Issues*, *58*, 487–502.

O'Brien, L. T., Crain, A. L., Omoto, A. M., & Snyder, M. (2000, May). *Matching motivations to outcomes: Implications for persistence in service.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Oishi, S., Rothman, A. J., Snyder, M., Su, J., Zehm, K., Hertel, A., et al. (2005). *Residential stability and the well-being of community: The socio-ecological model of self and prosocial behavior.* Unpublished manuscript, University of Virginia and University of Minnesota.

Okun, M. A., Barr, A., & Herzog, A. R. (1998). Motivation to volunteer by older adults: A test of competing measurement models. *Psychology and Aging*, *13*, 608–621.

Okun, M. A., & Schultz, A. (2003). Age and motives for volunteering: Testing hypotheses derived from socioemotional selectivity theory. *Psychology and Aging*, *18*, 231–239.

Omoto, A. M. (Ed.). (2005a). *Processes of community change and social action.* Mahwah, NJ: Erlbaum.

Omoto, A. M. (2005b). Understanding social change: Introduction to the volume. In A. M. Omoto (Ed.), *Processes of community change and social action* (pp. 1–10). Mahwah, NJ: Erlbaum.

Omoto, A. M., & Malsch, A. (2005). Psychological sense of community: Conceptual issues and connections to volunteerism-related activism. In A. M. Omoto (Ed.), *Processes of community change and social action* (pp. 83–102). Mahwah, NJ: Erlbaum.

Omoto, A. M., & Snyder, M. (1990). Basic research in action: Volunteerism and society's response to AIDS. *Personality and Social Psychology Bulletin*, *16*, 152–166.

Omoto, A. M., & Snyder, M. (1995). Sustained helping without obligation: Motivation, longevity of service, and perceived attitude change among AIDS volunteers. *Journal of Personality and Social Psychology*, *68*, 671–686.

Omoto, A. M., & Snyder, M. (2002). Considerations of community: The context and process of volunteerism. *American Behavioral Scientist*, *45*(5), 846–867.

Omoto, A. M., Snyder, M., & Martino, S. C. (2000). Volunteerism and the life course: Investigating age-related agendas for action. *Basic and Applied Social Psychology*, *22*, 181–198.

Omoto, A. M., Snyder, M., & Smith, D. M. (1999). Unpublished data. Lawrence: University of Kansas.

Orbell, J., Morikawa, T., Hartwig, J., Hanley, J., & Allen, N. (2004). "Machiavellian" intelligence as a basis for the evolution of cooperative dispositions. *American Political Science Review*, *98*, 1–15.

Orbell, J., van de Kragt, A., & Dawes, R. M. (1988). A "cognitive miser" theory of cooperators' advantage. *American Political Science Review*, *85*, 515–528.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome.* Lexington, MA: Lexington Books.

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behaviors. *Personnel Psychology*, *48*, 775–802.

Oskamp, S. (2000). A sustainable future for humanity? How can psychology help? *American Psychologist*, *55*, 496–508.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action.* New York: Cambridge University Press.

Ouellette, S. C., Cassel, B., Maslanka, H., & Wong, L. M. (1995). GMHC volunteers and the challenges and hopes for the second decade of AIDS. *AIDS Education and Prevention*, *7*, 64–79.

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005). Prosocial behavior: Multilevel perspectives. *Annual Review of Psychology*, *56*, 365–392.

Penner, L. A., & Finkelstein, M. A. (1998). Dispositional and structural determinants of volunteerism. *Journal of Personality and Social Psychology*, *74*, 525–537.

Penner, L. A., Fritzsche B. A., Craiger, J. P., & Freifeld, T. R. (1995). Measuring the prosocial personality. In J. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 10, pp. 147–163). Hillsdale, NJ: Erlbaum.

Penner, L. A., Midili, A. R., & Kegelmeyer, J. (1997). Beyond job at-titudes: A personality and social psychology perspective on the causes of organizational citizenship behaviors. *Human Performance*, *10*, 111–132.

Perkins, D. D., & Long, D. A. (2002). Neighborhood sense of community and social capital: A multi-level analysis. In A. Fisher, C. Sonn, & B. Bishop (Eds.), *Psychological sense of community: Research, application, and implications* (pp. 291–318). New York: Kluwer.

Piliavin, J. A. (1989). The development of motives, self-identities, and values related to blood donation: A Polish-American comparison study. In N. Eisenberg, J. Reykowski, & E. Staub (Eds.), *Social and moral values: Individual and societal perspectives* (pp. 253–276). Mahwah, NJ: Erlbaum.

Piliavin, J. A. (2003). Doing well by doing good: Benefits for the benefactor. In C. L. M. Keyes & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 227–247). Washington, DC: American Psychological Association.

Piliavin, J. A. (2004a). Feeling good by doing good. In A. M. Omoto (Ed.), *Processes of community change and social action*. Mahwah, NJ: Erlbaum.

Piliavin, J. A. (2004b). *I get more than I give: Volunteering, well-being, and health.* Paper presented at the annual meetings of the Eastern Sociological Association.

Piliavin, J. A. (2005, July). *Lifelong voluntary participation, well-being, and health.* Paper presented at the 14th general meeting of the European Association of Experimental Social Psychology, Wuerzburg, Germany.

Piliavin, J. A., & Callero, P. L. (1991). *Giving blood: The development of an altruistic identity.* Baltimore: Johns Hopkins University Press.

Piliavin, J. A., Grube, J. A., & Callero, P. L. (2002). Role as a resource for action in public service. *Journal of Social Issues*, *58*, 469–485.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, *24*, 1–24.

Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, *6*, 65–78.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community.* New York: Simon & Schuster.

Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology*, *81*, 5–16.

Reeder, G. D., Davison, D. M., Gipson, K. L., & Hesson-McInnis, M. S. (2001). Identifying the motivations of African American volunteers working to prevent HIV/AIDS. *AIDS Education and Prevention*, *13*, 343–354.

Rioux, S. M., & Penner, L. A. (2001). The causes of organizational citizenship behavior: A motivational analysis. *Journal of Applied Psychology*, *36*, 1306–1314.

Saegert, S., Winkel, G., & Swartz, C. (2002). Social capital and crime in New York City's low-income housing. *Housing Policy Debate*, *13*, 189–226.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*, 918–927.

Sanford, N. (1970). Whatever happened to action research? *Journal of Social Issues*, *26*, 3–23.

Sarason, S. B. (1974). *The psychological sense of community: Prospects for a community psychology.* San Francisco: Jossey-Bass.

Schondel, C., Shields, G., & Orel, N. (1992). Development of an instrument to measure volunteers' motivation in working with people with AIDS. *Social Work in Health Care*, *17*, 53–71.

Schroeder, D. A., Dovidio, J. F., Sibicky, M. E., Matthews, L. L., & Allen, J. L. (1988). Empathy and helping behavior: Egoism or altruism. *Journal of Experimental Social Psychology*, *24*, 333–353.

Schlager, E. (2002). Rationality, cooperation, and common pool resources. *American Behavioral Scientist*, *45*, 801–819.

Schwartz, C., Meisenhelder, J. B., Ma, Y., & Reed, G. (2003). Altruistic social interest behaviors are associated with better mental health. *Psychosomatic Medicine*, *65*, 778–785.

Simon, B., Loewy, M., Stuermer, S., Weber, U., Freytag, P., Habig, C., et al. (1998). Collective identification and social movement participation. *Journal of Personality and Social Psychology*, *74*, 646–658.

Simon, B., Stuermer, S., & Steffens, K. (2000). Helping individuals or group members? The role of individual and collective identification in AIDS volunteerism. *Personality and Social Psychology Bulletin*, *26*, 497–506.

Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, *68*, 653–663.

Smith, D. M., Omoto, A. M., & Snyder, M. (2001a, June). *Motivation matching and recruitment of volunteers: A field study*. Paper presented at the annual meeting of the American Psychological Society, Toronto.

Smith, D. M., Omoto, A. M., & Snyder, M. (2001b, February). *Stigmatization within a helping context*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, San Antonio, TX.

Snyder, M. (1993). Basic research and practical problems: The promise of a "functional" personality and social psychology. *Personality and Social Psychology Bulletin*, *19*, 251–264.

Snyder, M., & Cantor, N. (1998). Understanding personality and social behavior: A functionalist strategy. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 635–679). Boston: McGraw-Hill.

Snyder, M., Clary, E. G., & Stukas, A. A. (2000). The functional approach to volunteerism. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate: Functions of attitudes* (pp. 365–393). Hillsdale, NJ: Erlbaum.

Snyder, M., & Omoto, A. M. (2000). Doing good for self and society: Volunteerism and the psychology of citizen participation. In M. Van Vugt, M. Snyder, T. Tyler, & A. Biel (Eds.), *Cooperation in modern society: Promoting the welfare of communities, states, and organizations* (pp. 127–141). London: Routledge.

Snyder, M., & Omoto, A. M. (2001). Basic research and practical problems: Volunteerism and the psychology of individual and collective action. In W. Wosinska, R. Cialdini, D. Barrett, & J. Reykowski (Eds.), *The practice of social influence in multiple cultures* (pp. 287–307). Mahwah, NJ: Erlbaum.

Snyder, M., Omoto, A. M., & Crain, A. L. (1999). Punished for their good deeds: Stigmatization of AIDS volunteers. *American Behavioral Scientist*, *42*, 1175–1192.

Snyder, M., Omoto, A. M., & Smith, D. M. (in press). The role of persuasion strategies in motivating individual and collective action. In E. Borgida, J. Sullivan, & E. Reidel (Eds.), *The political psychology of democratic citizenship*. New York: Cambridge University Press.

Stryker, S., Owens, T., & White, R. (Eds.). (2000). *Self, identity, and social movements*. Minneapolis: University of Minnesota Press.

Stuermer, S. (2000). *Soziale Bewegungsbeteiligung: Ein psychologisches Zwei-Wege Modell* [Social movement participation: A psychological dual pathway model]. Unpublished doctoral thesis, Faculty of Philosophy, Christian-Albrechts-Universitat zu Kiel, Germany.

Stuermer, S., & Kampmeier, C. (2003). Active citizenship: The role of community identification in community volunteerism and local participation. *Psychologica Belgica*, *43*, 103–122.

Stuermer, S., & Siem, B. (2005, July). *A group-level perspective on the role of empathy and interpersonal attraction in helping*. Paper presented at the 14th general meeting of the European Association of Experimental Social Psychology, Wuerzburg, Germany.

Stuermer, S., & Simon, B. (2004a). Collective action: Towards a dual-pathway model. *European Review of Social Psychology*, *15*, 59–99.

Stuermer, S., & Simon, B. (2004b). The role of collective identification in community volunteerism: A panel study in the context of the German gay movement. *Personality and Social Psychology Bulletin*, *30*, 363–377.

Stuermer, S., Simon, B., Loewy, M., & Joerger, H. (2003). The dual pathway model of social movement participation: The case of the fat acceptance movement. *Social Psychology Quarterly*, *66*, 71–82.

Stuermer, S., Snyder, M., Kropp, A., & Siem, B. (2005). *Empathy-motivated helping: The moderating role of group membership*. Unpublished manuscript, University of Kiel, Germany.

Stuermer, S., Snyder, M., & Omoto, A. M. (2005). Prosocial emotions and helping: The moderating role of group membership. *Journal of Personality and Social Psychology*, *88*, 532–546.

Stukas, A. A., Clary, E. G., & Snyder, M. (1999). Service learning: Who benefits and why. *Social Policy Report: Society for Research on Child Development*, *13*, 1–19.

Stukas, A. A., Daly, M., & Cowling, M. J. (2005). Volunteerism and the creation of social capital: A functional approach. *Australian Journal of Volunteering*, *10*, 35–44

Stukas, A. A., Snyder, M., & Clary, E. G. (1999). The effects of "mandatory volunteerism" on intentions to volunteer. *Psychological Science*, *10*(1), 59–64.

Teske, N. (1997). *Political activists in America: The identity construction model of political participation*. Cambridge, UK: Cambridge University Press.

Thoits, P. A., & Hewitt, L. N. (2001). Volunteer work and well-being. *Journal of Health and Social Behavior*, *42*, 115–131.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Tropp, L. R., & Brown, A. C. (2004). What benefits the group can also benefit the individual: Group-enhancing and individual-enhancing motives for collective action. *Group Processes and Intergroup Relations*, *7*, 267–282.

Tschirhart, M., Mesch, D. J., Perry, J. L., Miller, T. K., & Lee, G. (2001). Stipended volunteers: Their goals, experiences, satisfaction, and likelihood of future service. *Nonprofit and Voluntary Sector Quarterly*, *30*, 422–443.

Tyler, T. R. (2000). Why do people cooperate in groups? Support for structural solutions to social dilemma problems. In M. Van Vugt, M. Snyder, T. Tyler, & A. Biel (Eds.), *Cooperation in modern society: Promoting the welfare of communities, states, and organizations* (pp. 64–82). London: Routledge.

Tyler, T. R. (2002). Leadership and cooperation in groups. *American Behavioral Scientist*, *45*, 769–782.

Tyler, T. R., & Blader, S. (2000). *Cooperation in groups: Procedural justice, social identity, and behavioral engagement*. Philadelphia: Psychology Press.

Tyler, T. R., & Degoey, P. (1995). Collective restraint in a social dilemma situation: The influence of procedural justice and community identification on the empowerment and legitimacy of authority. *Journal of Personality and Social Psychology*, *69*, 482–497.

Van de Vliert, E., Huang, X., & Levine, R. V. (2004). National wealth and thermal climate as predictors of motives for volunteer work. *Journal of Cross-Cultural Psychology*, *35*, 62–73.

Van Lange, P. A. M. (2000). Beyond self-interest: A set of propositions relevant to interpersonal orientations. *European Review of Social Psychology*, *11*, 297–331.

Van Lange, P. A. M., Van Vugt, M., & De Cremer, D. (2000). Choosing between personal comfort and the environment. In M. Van Vugt, M. Snyder, T. Tyler, & A. Biel (Eds.), *Cooperation in modern society: Promoting the welfare of communities, states, and organizations* (pp. 45–63). London: Routledge.

Van Lange, P. A. M., Van Vugt, M., Meertens, R. M., & Ruiter, R. (1998). A social dilemma analysis of commuting preferences: The roles of social value orientation and trust. *Journal of Applied Social Psychology*, *28*, 796–820.

Van Vugt, M. (2002). Central, individual, or collective control? Social dilemma strategies for natural resource management. *American Behavioral Scientist*, *45*, 783–800.

Van Vugt, M., & Hart, C. M. (2004). Social identity as social glue:

The origins of group loyalty. *Journal of Personality and Social Psychology*, *86*, 585–598.

Van Vugt, M., Meertens, R. M., & Van Lange, P. A. M. (1995). Car versus public transportation? The role of social value orientations in a real-life social dilemma. *Journal of Applied Social Psychology*, *25*, 258–278.

Van Vugt, M., & Samuelson, C. D. (1999). The impact of personal metering in the management of a natural resource crisis: A social dilemma analysis. *Personality and Social Psychology Bulletin*, *25*, 731–745.

Van Vugt, M., & Snyder, M. (2002). Cooperation in society: Fostering community action and civic participation. *American Behavioral Scientist*, *45*, 765–768.

Van Vugt, M., Snyder, M., Tyler, T., & Biel, A. (Eds.). (2000). *Cooperation in modern society: Promoting the welfare of communities, states, and organizations*. London: Routledge.

Van Vugt, M., & Van Lange, P. A. M. (2006). The altruism puzzle: Psychological adaptations for prosocial behavior. In M. Schaller, J. Simpson, & D. Kenrick (Eds.), *Evolution and social psychology* (pp. 237–261). New York: Psychology Press.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

Wandersman, A. (1980). Community and individual difference characteristics as influences on initial participation. *American Journal of Community Psychology*, *8*, 217–228.

Wandersman, A., Florin, P., Friedmann, R. R., & Meier, R. B. (1987). Who participates and who does not, and why?: An analysis of voluntary neighborhood organizations in the United States and Israel. *Sociological Forum*, *2*, 534–555.

Williamson, I., Snyder, M., & Omoto, A. M. (2000, May). *How motivations and reenlistment frames interact to predict volunteer attitudes and intentions: A test of the functional matching effect*. Paper presented at the annual meetings of the Midwestern Psychological Association, Chicago.

Wilson, J. (2000). Volunteering. *Annual Review of Sociology*, *26*, 215–240.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, *18*, 129–166.

Author Index

Subject Index

Page numbers followed by *f* indicate figure, *n* indicate note, and *t* indicate table