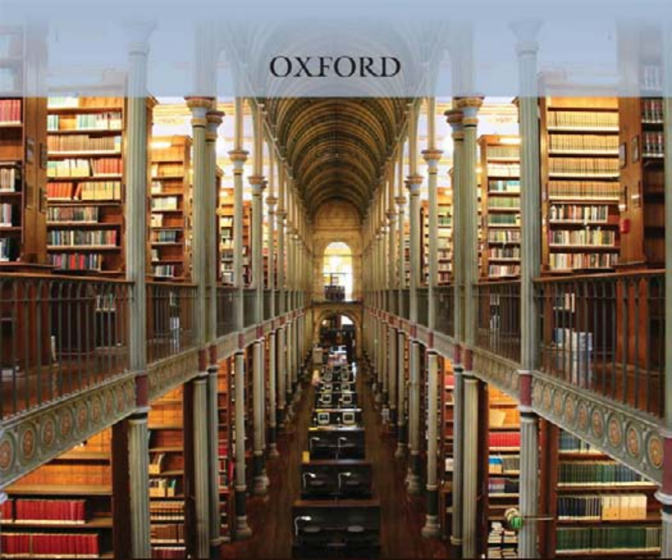
OXFORD



Arlene Rubin Stiffman

# The Field Research Survival Guide

# The Field Research Survival Guide

# The Field Research Survival Guide

**Edited by**

## Arlene Rubin Stiffman

New York　　Oxford
OXFORD UNIVERSITY PRESS
2009

OXFORD
UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Library of Congress Cataloging-in-Publication Data

The field research survival guide / edited by Arlene Rubin Stiffman.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-19-532552-2 (pbk. : alk. paper)
1. Social service—Fieldwork. 2. Psychology—Fieldwork.
3. Psychiatry—Research—Fieldwork. 4. Public health—Fieldwork.
I. Stiffman, Arlene Rubin, 1941-
HV11.F454 2009
361.0072'3—dc22 2008027003

9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

# Preface

Those who are new to a research career have taken courses on research methodology and statistics and likely served as a research assistant on a professor's project. But are they really ready to conduct their own field research? This author certainly was not ready—although, of course, she had no idea of how woefully unprepared she really was. Oh, she could plan a scientifically sound study on paper, but no one prepared her for the many practical problems and choice points that occur during research projects and across the life of a research career. Field research, whether survey or intervention research, is that in which the researcher is doing his or her study in the real world rather than a laboratory situation. As such, it involves many potential hazards to the scientific control needed for valid research. When discussing the situation with colleagues, they also mentioned similar initial experiences. Furthermore, there is very little literature to guide any researcher who must make numerous practical field decisions.

As researchers gain experience, their role begins to include mentoring junior faculty in writing proposals and planning their research careers. To this author's great joy, a number of her mentees' research proposals were funded. However, to her great consternation, some of those research projects foundered sadly when being mounted. For example, one researcher hired a data entry company but did not make it clear that the case number on the front of the packet also had to be entered or how to code skipped questions. This slowed his progress by months when he finally found out what was wrong.

One of the most important tasks performed by researchers in social service, psychiatry, psychology, sociology, and anthropology is managing a field research project. Yet field research training largely focuses on theories and formal methods, leaving researchers with no guidance in actual implementation. As explained in the following pages, this book is composed of chapters written by leading experienced field researchers. The chapters together provide a concise, inclusive narrative about the experiences of leading researchers in field research and information about the skills needed for successful field research management.

The chapters each address one or more of the research stress points that many researchers have found concerning during their research careers and are loosely organized by the naturally unfolding series of choice points and problems confronted during research projects and careers. Unlike the standard research methods text, each chapter has practical import for the researcher, ties together extant literature, and illustrates the issues with concrete examples from the chapter authors' own experience. Each chapter also closes with a brief outline of its main points. All the chapters are authored by nationally—and often internationally—recognized leaders in different aspects of field research. Because of their expertise, each author was asked to write in his or her own "voice," detailing his or her personal experience. Although the chapters progress naturally through the stages of research, the reader will note some overlap in principles and approaches that may be applied to the various research stages.

The book allows both new and more experienced researchers to learn from the shared experiential knowledge garnered by seasoned, widely respected field researchers. Thus, it should help both new and experienced researchers make the choices and decisions that will enhance the outcome and productivity of their research.

The first chapter in this book (Hoagwood) deals with the initial implementation of a funded project that involves developing the support and stakeholder networking necessary to conduct research. True, some of this was developed before funding, but it always needs reinforcing and extension. Field research differs dramatically from laboratory research in more than just locale. Life happens, organizations evolve, interveners' priorities and pressures change, policies respond to different laws and regulations, and agency finances do not remain static. The field world in which the research actually gets mounted may be very different from the field world that existed 1, 2, or 3 years prior when the research proposal was first developed.

The second chapter (Horwitz) addresses an issue common to all projects and typically deals with at two points: proposal development and project initiation. A consistent issue is developing the project instrumentation by finding, adapting, and gathering assorted measures to best tap a research project's goals. Unlike textbook situations, in real-life fieldwork one can seldom find adequate, established measures for each concept being researched. Often measures need to be adapted to the needs or limitations of particular subject populations or truncated to fit a reasonable time limitation for the interview.

The third chapter (Arnold) addresses an aspect of field research that is never addressed in research training and yet is the foundation of all successful field research: the business aspect of conducting field research. Hiring, training, and maintaining an interview pool requires a set of skills that differs dramatically from that of the conceptual work of research. Yet without doing this part well, obtaining reliable valid data and meeting the time and cost goals of the project become impossible.

The fourth (Doré) and fifth chapters (Smith) address the spectrum of issues related to data management. Together they cover the major aspects of the entire trajectory of data management from instrumentation through archival documentation. The fourth chapter focuses on the development of instruments and management of interviewers. It deals exclusively with the practicalities of how to manage data to increase and maintain reliability and accuracy. The chapter focuses primarily on the issues of setting up the printed or computerized interview; structuring it to minimize human error; and maximizing a short timeline to a clean, usable, archivable dataset suitable for analysis. The fifth chapter also addresses data management concerns but focuses primarily on data standards for the processes of data entry choices, documentation, hiring, and variable naming.

The sixth chapter (Stiffman) discusses the tension between human subject considerations, basic ethics, cultural approaches to the research process, and researchers' need for objective data not influenced by the research project itself. It outlines potential conflicts and presents examples of how these issues have been managed by other research projects.

The seventh chapter (Landsverk) addresses the need to organize collaborative research projects that will benefit from a coordinated variety of skills. It discusses how to gather and use the collaboration of experts, mentors, and consultants as well as how to move into the mentor/consultant role as one's experience grows.

The eighth chapter (Walters) addresses the issues and difficulties in doing research on ethnic minorities and other discriminated against groups in a sensitive and respectful manner that will ultimately benefit the ethnic community. One of the unusual aspects of this chapter is that in its example it addresses a particular ethnic minority (Native Americans) and that the first author and many of the other authors are themselves Native Americans. Thus, as researchers, they are exceptionally and uniquely aware of the past misuse of research concerning their peoples. Their sensitivity to and pain from the historical trauma their people had suffered permeates the chapter, forcing the reader to become aware of the depth of the issues confronting a field researcher.

The ninth chapter (Bickman) addresses every researchers' nightmare: what to do when you have null results. What does it mean? How can an investigator extract value from the data and move on in one's career from that point?

The tenth chapter (Sherraden) lets the author's field work on individual development accounts (IDAs) serve as an illustration of how he used and is using research to influence national and international policies concerning asset development. The chapter also explicitly deals with the critical, but never discussed, issue of planning a career that will make a difference in the world.

The last chapter, Chapter 11, (Howard) addresses an issue of absolutely prime importance to academics, yet is one that is not addressed formally in any training situation. To function in an academic research career, one must learn the practicalities of publishing and disseminating research results in such a way as to maximize impact on the field and on one's own research career trajectory.

This book is NOT a research methods text. It IS a companion to such a text that will enable researchers to anticipate problems and issues at various stages of the field research process. Readers will gain knowledge of the experiences of experts to solve similar problems in their own field research. None of the chapter authors intends to say that the pattern of his or her work, or the solutions arrived at, are the only right ones. Instead, the authors allow their experiences to serve as a base for creative problem solving.

# Acknowledgments

Success in any endeavor depends to a great extent on mentors, colleagues, and family. I have been lucky in each department.

Although throughout one's career, many people serve as mentors, I would like to especially thank my first mentors: Dr. Ronald Feldman, Dr. Felton Earls, and Dr. Lee Robins.

I have also been lucky enough to have wonderful colleagues who have made my work a pleasure every day for several decades. Two have been especially supportive: Dr. Enola Proctor and Dr. John Landsverk.

Last, but really first in importance, is the support of my family: my husband, Rabbi Jeffrey Stiffman, who gave me confidence and love; my children, Dr. Michael Stiffman, Marti Stiffman, and Cheryl Maayan, who make me so proud; and my grandchildren, Alex, George, Sam, Eliana, Maya, Mickey, Gabe, and Ari, who bring me endless joy.

# Contents

Contributors *xiii*

1 Balancing Science and Services: The Challenges and Rewards of Field Research *3*
*Kimberly Eaton Hoagwood and Sarah McCue Horwitz*

2 Developing Questions when the Perfect Instrument Is Not Available *23*
*Sarah McCue Horwitz and Kimberly Eaton Hoagwood*

3 Hiring, Training, and Retaining Research Staff and Interviewers *38*
*Elizabeth Mayfield Arnold and Mary Jane Rotheram-Borus*

4 Managing the Data from Survey Development through Archiving *59*
*Peter Doré and Arlene Rubin Stiffman*

5 Data Preparation and Data Standards: The Devil Is in the Details *82*
*Catherine M. Smith, Carolyn Breda, Tonya Simmons, Ana Regina Vides de Andrade, and Leonard Bickman*

6 Cultural Sensitivity and Cultural Disparities: Ethical Dilemmas, Legal Issues, and IRB Requirements *104*
*Arlene Rubin Stiffman*

7 Creating Interdisciplinary Research Teams and Using Consultants *127*
*John Landsverk*

8 "Indigenist" Collaborative Research Efforts in Native American Communities *146*
*Karina L. Walters, Antony Stately, Teresa Evans-Campbell, Jane M. Simoni, Bonnie Duran, Katie Schultz, Erin C. Stanley, Chris Charles, and Deborah Guerrero*

9 The Worst of all Possible Program Evaluation Outcomes *174*
*Leonard Bickman and Michele Athay*

10 The Influence of Research on Policy and Practice: Lessons from Studies of Asset Building and Low-Income Families *205*
*Michael Sherraden and Trina Williams Shanks*

11 Disseminating Results and Sharing Data and Publications *234*
*Matthew Owen Howard and Michael G. Vaughn*

Index *257*

# Contributors

**Ana Regina Vides de Andrade, Ph.D.**
Center for Evaluation and Program Improvement,
Vanderbilt University,
Nashville, TN

**Elizabeth Mayfield Arnold, Ph.D.**

Associate Professor, Department of Psychiatry and Behavioral Medicine,
Wake Forest University School of Medicine,
Winston-Salem, NC

**Michele Athay**

Center for Evaluation and Program Improvement,
Vanderbilt University,
Nashville, TN

**Leonard Bickman, Ph.D.**
Professor of Psychology, Psychiatry and Public Policy; Associate Dean for Research; and Director, Center for Evaluation and Program Improvement,
Vanderbilt University,
Nashville, TN

**Carolyn Breda, Ph.D.**
Center for Evaluation and Program Improvement,
Vanderbilt University,
Nashville, TN

**Chris Charles, BSW**
Media Relations Assistant, Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Peter Doré, M.S.**
Director, Data Management Unit, Center for Mental Health Services Research,
George Warren Brown School of Social Work,
Washington University,
St. Louis, MO

**Bonnie Duran, Dr. P.H.**
Director, Center for Indigenous Health Research,
Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Teresa Evans-Campbell, M.S.W., Ph.D.**
Associate Professor and Director, The Center for Indigenous Child Welfare and Family Wellness,
Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Deborah Guerrero, BA**

**Kimberly Eaton Hoagwood, Ph.D.**
Professor of Clinical Psychology in Psychiatry at Columbia University, and Bureau Director,
Child & Adolescent Services Research,
New York State Office of Mental Health,
New York, NY

**Sarah McCue Horwitz, Ph.D.**
Professor of Pediatrics,
Stanford University,
Stanford, CA

**Matthew Owen Howard, Ph.D.**
Frank A. Daniels Distinguished Professor, School of Social Work,
University of North Carolina,
Chapel Hill, NC

**John Landsverk, Ph.D.**
Child and Adolescent Services Research Center,
Rady Children's Hospital,
San Diego, CA

**Mary Jane Rotheram-Borus, Ph.D.**
Bat-Yaacov Professor of Child Psychiatry and Behavioral Sciences, UCLA Department of Psychiatry; Director, UCLA Global Center for Children and Families; and Director, Center for HIV Identification, Prevention, and Treatment Services (CHIPTS) at the UCLA Semel Institute for Neuroscience and Human Behavior,
University of California, Los Angeles,
Los Angeles, CA

**Katie Schultz, MSW**
Administrative Director, Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Trina Williams Shanks, Ph.D.**
Assistant Professor, University of Michigan School of Social Work,
Ann Arbor, MI

**Michael Sherraden, Ph.D.**
Youngdahl Professor of Social Development and Director, Center for Social Development, George Warren Brown School of Social Work,
Washington University,
St. Louis, MO

**Tonya Simmons**
Center for Evaluation and Program Improvement,
Vanderbilt University,
Nashville, TN

**Jane M. Simoni, Ph.D.**
Professor of Psychology,
University of Washington,
Seattle, WA

**Catherine M. Smith, Ph.D.**
Deputy Director and Research Associate, Center for Evaluation & Program Improvement,
Vanderbilt University,
Nashville, TN

**Erin C. Stanley, MSW**
Former Assistant Director, Honor Project, Native Wellness Center in the Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Antony Stately, Ph.D.**
Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

**Arlene Rubin Stiffman, Ph.D.**
Senior Scholar and Barbara A. Bailey Professor of Social Work, Emeritus, at the George Warren Brown School of Social Work,
Washington University in St. Louis,
St. Louis, MO

**Michael G. Vaughn, Ph.D.**
Assistant Professor, School of Social Work,
The University of Pittsburgh,
Pittsburgh, PA

**Karina L. Walters, M.S.W., Ph.D.**
Director, Indigenous Wellness Research Institute,
University of Washington School of Social Work,
Seattle, WA

# The Field Research Survival Guide

# 1

# Balancing Science and Services: The Challenges and Rewards of Field Research

*Kimberly Eaton Hoagwood and Sarah McCue Horwitz*

> Science is inevitably and inextricably bound to social processes . . . To study social reality requires a theory of profoundly transformative processes.
> —M. Horkheimer and T. Adorno, 1998

Fitting research aims and expectations into the demanding conditions of real-world practice, although a laudable goal, can create a host of problems for researchers trying to generate usable and practical findings to improve routine practice. The fundamental conflicts between science and practice come into sharp relief. The designs for high-quality scientific studies must control for many sources of bias if they are going to yield valid findings. On the other hand, many practical and necessary constraints are made on service delivery practices, and these can and do conflict with scientific goals. Although closing the research-to-practice gap is agreed to be a highly significant and important public health goal (National Advisory Mental Health Council, 1999; New Freedom Commission on Mental Health, 2003), in reality doing so requires enormous patience, skills in working with many different individuals representing diverse interests, flexibility, tolerance for constant change, and an unflinching commitment to scientific principles, the practicalities of grounded fieldwork, and one's own ethical compass. Managing the demands of science and practice, in short, requires attention to a host of issues that are simultaneously pragmatic, ethical, and scientific.

The purpose of this chapter is to describe steps that the field researcher can take to position his or her studies strategically to respond to the challenges that inevitably arise. Case examples from the authors' studies are provided in these steps. In addition, a model for accelerating the applicability of service studies is described to create a better fit between the demands of rigorous science and the exigencies of practice.

Undertaking field research in children's mental health is not for the faint hearted. Many fail. Yet when the effort is successful, the result can be the generation of findings with tremendous impact. High-quality field research can produce valid, ethically derived, and immediately relevant scientific findings that can be integrated into practice settings and that can—sometimes with astonishing haste—improve the lives of children and families seeking services. The personal rewards of this work and its public health significance have no parallel.

## Steps toward Launching Scientifically Valid Field Research

Outlined in the following sections are seven steps that a field researcher can take to deal with the challenges that arise when bringing science and practice together. Case examples from the authors' experiences are woven into the steps.

### *Step 1: Openly acknowledge the differing purpose between the goals of the research and the goals of the practice setting; then identify the shared aims for both*

The integration of high-quality research within field settings can undoubtedly enhance science by generating useful and usable knowledge and improve practice. However, recognizing that the purposes of each are distinct and at times incompatible is also important.

Scientific pursuits are characterized by the application of rigorous methods to clearly specified questions for the purpose of deriving reliable and valid answers that are not predetermined. Studies are undertaken not to "prove" that a particular strategy, intervention, or program works but rather to examine whether and under what conditions the particular strategy may exert an effect—positive, negative, or neutral.

This issue about the purpose of scientific inquiry is often confusing to members of the practice community, particularly if funding may be in jeopardy if the study does not "prove" the value of a particular service program. The fact that there can be no predetermined answers to the primary research questions must be made clear from the outset. Doubt is the driver. It is important to emphasize that the purpose of research is to answer study questions in as valid and reliable a manner as possible. Thus a range of results is possible.

Second, recognizing that science is inherently conservative, slow, and incremental is important. In fact, the findings from most scientific studies are quite modest. Large effects are the exception, not the rule. Furthermore, scientific studies often are the most important to the field when they identify nuances—findings that occur under one condition but not another. In this way, science is self-correcting and proceeds incrementally. The process of empirical inquiry itself characterizes scientific endeavors (Habermas, 1990; Horkheimer & Adorno, 1998).

Practice environments, such as those in which most field research described by the authors in this volume occur, are characterized by a different set of purposes. Real-world service systems exist for the purpose of serving the public interest by spending public dollars for particular populations and services. They are organized to administer public health services to eligible populations, often defined by statute, law, or federal or state administrative authorities. The priorities for the practices may be defined by political agendas. In addition, most public service systems are overwhelmed by demands that are nearly impossible to meet. These include inadequate numbers of staff, large turnover rates, long waiting lists, and often daily crises that drive service priorities and demand immediate attention. As a consequence, most practicing clinics are managed by decision making that is expedient, opportunistic, and often unpredictable.

Given these dual (if not dueling) purposes between research and practice, improving the linkage between them automatically places demands on both the science and the practicing environment. New research paradigms are needed to better fit research studies and their findings into what Weisz, Weiss, and Donenberg (1992) called "the crucible of real life." Further discussions of models for doing so are discussed in the following text.

However, improving the linkage also places demands on community practice. Theoretical models and empirical studies of innovation and effectiveness in service environments have identified several practice parameters that contribute to service delivery (Glisson, 2002; Glisson & James, 2002). One is the treatment technology itself—the specific therapies, assessment tools, monitoring systems—provided to clients seeking services. The second consists of organizational factors that reflect the contexts of practice. These include issues such as staff turnover, stress, resistance or openness to change, opportunities for rewards and advancements, and work attitudes. The third parameter represents the attitudes, beliefs, and expectations of the families, youth, or consumers seeking services. Lack of information about mental health treatment, distrust of service providers, and dissatisfaction with services have been found to lead to limited family support for treatment and low treatment completion and retention (McKay & Bannon, 2004).

The implication of studies on effective practices is that embedding new therapies or service models in work environments does not, by itself, improve delivery of care. Attention to all three aspects of service delivery is necessary if improvements in the quality of service delivery are to occur.

The common ground that connects the aims of the research study should be connected to the needs of the practice setting. Often this may involve aspects of the treatment technology, the organizational context that might be improved, or the levels of consumer or family involvement that may be modified. Identifying areas of commonality among key partners in the study and at the site can help launch the study with a set of shared goals among all partners.

### *Step 2: Create an advisory board to guide the research project*

One of the most important practical steps that a field researcher can undertake at the outset of a project is the creation of an advisory group composed of primary stakeholders within the practice setting. This may include family members or representatives of the family or consumer perspective; clinicians who represent on-the-ground field workers; middle management or supervisors who attend to the practical realities of case assignment, workload, and so forth; and community leaders who have a stake or interest in the outcomes of the study. One group (Hatch, Moss, Saran, Presley-Cantrell, & Mallory, 1993) identified a range of collaborative models that can be useful in organizing the collective efforts of stakeholders. For example, low-intensity collaborations may begin with informal meetings at which researchers consult key stakeholders representing agencies or institutions within a specific community for advice or guidance. A more intensive collaboration exists when researchers identify key informants from the community and seek their input on key aspects of the research project. Community leaders can be invited to sit on advisory boards and provide direct input to the project.

The creation of formal mechanisms for soliciting input is a valuable way to maintain openness and communication between the researcher and community partners. It is also helpful in ensuring that the demands of the scientific agenda are understood and that the interests of the practice setting are communicated openly. In these meetings the boundaries between science and practice can often be openly discussed and the areas where negotiation and compromise are possible can be identified, as well as those areas where change would compromise the integrity of the project (e.g., such as cutoff designs that require use of a predetermined assessment measure, random assignment to intervention arms).

Numerous other models of collaborative arrangements are possible. For example, in the authors' work on parent empowerment and engagement in New York, a Family Advocacy and Research (FAR) Board has been created to deal with specific research problems, including revising the training protocols, incorporating specific motivational and engagement techniques into the program, responding to unexpectedly high levels of parental depression, and clarifying parent advisor roles within their own work setting. Other examples of collaborative partnership models abound, especially in the work of Atkins and colleagues (1998, 2003); Elliott, Koroloff, Koren, & Friesen (1998); Frazier, Cappella, and Atkins (2007), Jensen and Hoagwood (2008); Koroloff, Elliott, Koren, and Friesen (1994); Koroloff et al. (in press); McCormick, McKay, Wilson, McKinney, Paikoff, Bell, et al. (2000); McKay, Stoewe, McCadam, and Gonzales (1998); and McKay (in press). These models emphasize a participatory action research framework (Singer et al., 1993) with the goal of effecting transfer of research skills to all participants, while contributing to public education, advocacy, and empowerment of the community. The underlying principles of this collaboration include: (1) agreement and investment in shared goals; (2) equitable distribution of power, including fair involvement in decision making and opportunities to change aspects of the research process; (3) recognition of skills and expertise associated with both university training and community experience; (4) ongoing opportunities for communication based upon commitment to honest exchanges and willingness to raise concerns without blame and; (5) trust (McKay, in press).

### *Step 3: Be clear about roles and boundaries. Set up organizational structures to respect those roles and manage the project, but remain flexible and be prepared to change these structures if new needs arise*

Conducting research in field settings can sometimes create confusion about the roles of research staff versus clinical or service staff. Being explicit about roles and expectations is especially important early in the project, but the need to clarify roles can occur at any time in a study. For example, if the study requires that clinicians collect data, then the extent to which this is incorporated into ongoing practice or is kept separate from it and only part of the research protocol is an important distinction. Explicit direction about the roles that individuals in the practice setting will have regarding the study is necessary, and the burden on practitioners should be kept to a minimum. Will staff be paid extra for data collection or will this be considered part of their ongoing clinical responsibilities? Will research staff be assigned to the setting to assist in data collection? If so, how can they be folded in so as not to disrupt normal activities?

For example, in one large study in New York City involving 173 clinicians and supervisors across nine provider organizations in 45 sites, clinicians were asked to gather clinical assessment measures for the evaluation of the project during intake. The data collection was incorporated into ongoing clinical procedures and placed some burden on practicing therapists. The directors and middle management (i.e., supervisors) were included in the discussions about this design element and were fully supportive of it before implementation. To reduce the paperwork burden, research staff were hired to handle data entry and create and manage the database itself. This was paid for with the funds from the evaluation grant. The research assistants came into the clinics and abstracted the data directly from the files. This arrangement protected the collection of the data, minimized some of the paperwork burden on the clinicians, and reduced unnecessary interference with clinical operations. This negotiated arrangement enabled the research and clinical goals to be protected (CATS Consortium, 2007).

No matter how carefully a study is planned, expect conflicts between the demands of clinical settings and those of research. In a study examining factors related to the diagnosis and management of psychosocial problems in pediatric primary care practices, research assistants sat in the waiting rooms of participating practices and invited the parents of children visiting the practices to participate in the study. One practice, keenly interested in psychosocial issues, believed that it would be more appropriate to have practice physicians invite families to participate. The response rate in this practice initially was as good as or better than response rates in practices at which the research assistants did the recruiting. However, as the demands of clinical practice accelerated during high-illness months, physicians no longer had time to recruit families, and response rates dropped sharply (Horwitz, Leaf, Leventhal, Forsyth, & Speechley, 1992).

Being explicit about research versus practice roles and working closely with upper and middle management staff to ensure that normal operating procedures are not disrupted can help offset potential problems in fitting the study into the service setting.

### *Step 4: Structure the research project to reflect collaboration between research and field staff. Be flexible*

One of the ways to incorporate research projects into service sites is to structure the project so that it reflects and supports the administrative hierarchy within the practice setting where the study will be conducted. This can be accomplished by providing remuneration to the clinic or its staff for work related to the study. It can also occur by creating co-leadership positions that reflect the shared responsibilities for service delivery and research.

For example, in a study of empowerment among family advocates in a network of parent advisors in New York City, the authors structured the project so that it had two co-directors. One was an experienced family advocate with respect for research; the other was an experienced clinician/researcher with respect for advocacy. These two individuals, along with the principal investigator, a team of trainers, and an advisory board, made all the decisions about the project, oversaw all activities and interpreted the data and findings together. Having two project directors enabled the study to fit more directly within the family advocacy network in New York and enabled the findings from the study to inform the city's efforts directly to restructure family advocacy programs (Jensen & Hoagwood, 2008; Olin et al., in press).

### *Step 5: Be prepared for the funding priorities of the grantor to shift. If they do, remain flexible and responsive but do not compromise the integrity of the study*

Funding for most services research projects comes from federal, state, or foundation sources. These granting agencies are subject to changes in their research or evaluation priorities as legislative, administrative, political, or policy shifts occur. This is not unusual. Researchers working on a project within the context of shifting funding priorities can be asked by granting agencies—sometimes when a study is already underway—to modify the project by adding additional questions or measures, expanding the population being examined, or even changing the basic design.

Although remaining responsive and flexible to these requests is important so that the findings from the study can have the greatest impact, protecting the integrity of the study is equally important. Will the requested additions or modifications enhance or jeopardize the core questions that the principal investigator is responsible for addressing? The principal investigator and his or her core team have the responsibility to think through the implications of any requested changes carefully and ensure that the integrity of the scientific project is not compromised. This may entail negotiating with the granting agency, discussing with experts in the field about the implications of any changes for the project, and sometimes, simply saying no.

### *Step 6: Make the budget, the staffing, and the research goals explicit and transparent*

One of the surest ways to unravel a partnership quickly is to arouse suspicions about the fairness of funding by obscuring details about the budget. The requirements of the funding agency, the intended distribution of funds, and the major activities that will be supported by the funds are issues that should be discussed openly among the leadership team of the project. This often includes advisory board leaders and project directors. Although the ultimate decision about funding rests with the principal investigator, creating a transparent decision-making process can be essential to establishing trust among the partners.

For example, in a multisite study of child trauma after September 11 in New York, an executive steering committee composed of all site directors, co-directors, and the principal investigators and co-investigators from the evaluation team (including one of the authors) was created, and they met weekly by phone to structure the project; establish an organizational hierarchy; set up subcommittees to manage the selection of assessment measures, treatment modalities, and training protocols; and manage institutional review board issues. One of the issues that arose midway in the project was an opportunity for additional funding for some of the sites to provide enhanced services. This opportunity, although potentially beneficial for some sites, put the evaluation design in jeopardy because it provided services to participants in the comparison condition that were similar to the experimental group. In addition, it created the potential for unevenness in funding across the sites. The authors brought this issue fully into the open in discussions during the steering committee calls. The budget and fiscal officers from each of the sites were assembled to discuss how, when, and under what conditions the enhanced services would be provided. The authors found ways to equalize the funding across the sites and integrate this new service model into the design of the study. This was a laborious undertaking that consumed many hours of research staff and the principal investigator's time, yet failure to do so would have certainly undermined the entire project (CATS Consortium, 2007).

### *Step 7: Share the credit and do not surprise colleagues*

Although academic credit in the form of manuscripts, book chapters, and invited presentations is most valuable to the scientists on the research team, clinical and community partners deserve acknowledgment and often equal intellectual credit for their contributions to the success of the research endeavor.

Engaging clinical and community partners in the development of research products in a meaningful, not token, way and ensuring that everyone on the research team understands the findings are critical for any future partnerships. These kinds of substantive collaborations around the products of research studies may strongly influence whether the results from the research result in long-term service change.

## New Research-to-Practice Paradigms

Connecting research and practice is fraught with stumbling blocks. Some are inevitable, and the best that can be done is muddle through them as they arise. Some can be foreseen and, as noted above, some steps can be taken to minimize their impact.

But some of the difficulties are inherent in the scientific paradigms themselves that guide research training. One reason that the findings from many efficacy studies of treatment outcomes, for instance, have not been readily deployed into service settings is that the research model used to guide the development, refinement, and testing of those treatments does not mesh with the practical constraints of clinic or community-based care (Weisz, Jensen, & McLeod, 2005b).

Students typically are taught to conduct scientific studies first within controlled settings such as university laboratories. Over the past 40 years, a large number of controlled clinical trials and within-group studies have been published on the impact of psychosocial treatments examined in this way. By and large, these studies have demonstrated that specific treatments are efficacious for approximately two dozen clinical conditions in children. In contrast, studies of conventional treatments delivered in clinics and clinical programs have demonstrated much weaker effects (Weisz, Doss, & Hawley, 2005a; Weisz, Weiss, Han, Granger, & Morton, 1995).

An implicit assumption has been made that once the laboratory studies of the efficacy of treatments have been completed, the results will be usable and relevant outside these laboratories. But as Weisz and colleagues have noted time and time again (Kazdin & Weisz, 1998; Weisz et al., 1992, 1995, 2003; 2005a), numerous differences exist between the conditions of most research and the conditions in which everyday service is delivered. The discrepancies between research findings, as tested within controlled environments, and service delivery, as provided in real-world clinics or community settings, constitute the single most important impediment to improving the quality of care in routine practice (Burns, Hoagwood, & Mrazek, 1999; Burns & Hoagwood, 2004; Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001; Jensen, Hoagwood & Petti, 1996; Jensen, Hoagwood, & Trickett, 1999; Schoenwald & Hoagwood, 2001).

The problem is that many of the real-world factors that laboratory researchers consider "nuisance variables" and therefore rule out or control experimentally are precisely those variables that need to be examined in field studies (Bickman, 1990; Hohmann & Shear, 2002). Therefore, new models for connecting research and practice are needed to avoid the current situation in which research findings are largely irrelevant to practice, and practices are largely unexamined.

## The Community Intervention Development Model

To provide a framework for conceptualizing how new treatment or service interventions can be developed, tested, and deployed within practice settings, Hoagwood, Burns, and Weisz (2002) described a community intervention development model (CID). This model was conceptualized as an extension of Weisz' deployment-focused model (Weisz 2003; Weisz et al., 2005b). The CID model describes a series of steps that can accelerate the pace at which the science base for mental health services can be developed, adapted, and refined; importantly, it describes steps for taking research-based practices to scale and rapidly disseminating findings. The model outlines a series of steps that begin and end with the practice setting (e.g., clinic, school, health center) where the treatment or service will ultimately be delivered. The goal is to enhance the probability that the end product—a scientifically valid treatment or service—will be grounded, useable, and relevant to the practice context for which it is ultimately intended. The steps are briefly outlined here but the reader is directed to Hoagwood et al. (2002) for elaboration. Steps 1 to 6 are thoroughly described by Weisz and collaborators (2003, 2005b) and are summarized here, with minor modifications, to make them adaptable to a range of interventions involving treatment, service, or prevention. Steps 7 and 8 extend the model to the dissemination and implementation of the intervention into a variety of practice settings or communities.

### *Step 1: Developing and creating a manual for the protocol*

As with any new therapy or service, the first step is to develop, pilot test, and refine a manual for the protocol. The scientific and theoretical literature is useful in identifying the constructs and the rationale for the intervention and the hypothesized explanation for intended outcomes, but the experiences of clinicians, practitioners, or other mental health providers within the setting in which the treatment or service is to be developed are essential for ensuring that the protocol reflects the needs and capacities of those who will ultimately deliver it.

### *Step 2: Efficacy trial*

An initial test of the protocol occurs in this step under controlled conditions and with children who have significant problems but not severe pathology. These children typically will be recruited for the study and will not have been referred for treatment in service settings. This step is intended to assess whether the service, compared with a control condition, results in positive outcomes among children, who are usually volunteers. The purpose of this step is to ensure that the protocol does no harm and has the potential to be beneficial. The controlled conditions also enable the investigator to test whether the hypothesized relation between the intervention and the consequences is supported.

### *Step 3: Single-case applications*

This step involves a series of pilot tests with cases referred from the practice setting to research-trained practitioners or clinicians familiar with the protocol. Across the series of individual cases, adaptations are made to the protocol to reflect what is learned about individual variations. The objective is to keep the development and refinement of the protocol closely connected to practice with increasing involvement of real-world cases in real-world practice settings. At the same time, an equally important goal is to maintain sufficient scientific control over the testing of the new protocol such that the intervention is developed in an ethical and scientific manner and no harm is done to the participants. This third step therefore should involve supervision from both the research team that has developed the manual of protocol and the practitioner or clinical staff to ensure confidence that the protocol is appropriate for the needs of the clients, the clinical staff, and the practice setting.

### *Step 4: Initial effectiveness trial*

This step involves a trial of the newly adapted protocol within the practice setting itself. Clients who typically are seen in the practice are randomly assigned to receive the new protocol or services as usual. The protocol is delivered by research-trained staff, generally not by actual practice staff, to assess the impact of the new protocol under slightly more controlled conditions than are normal in practice settings.

### *Step 5: Full effectiveness trial*

This step entails a randomized field trial in which the protocol is tested within the practice setting itself. Clients are randomly assigned to receive either the protocol or the services as usual; the actual clinic or practice setting staff are randomly assigned to deliver the new protocol or deliver services as usual. Outcomes are tracked for their long-term (i.e., at least 12 months) impact, and a range of outcomes is assessed, including cost effectiveness. Embedded within this trial can be tests of moderators or mediators hypothesized to be related to outcome variations.

### *Step 6: Effectiveness of moderators and mediators*

A variety of studies can be launched in this step to address factors that impinge on outcomes; for example, tests of outcome moderators and tests of variations in the treatment or service, such as differential impact of treatment for children alone versus children and their parents. Also included here are tests of the mediators of child improvement (i.e., the change processes that potentiate outcomes). Tests of treatment variations and mediators are especially important in the effort to keep treatments streamlined and efficient. Findings are relevant circumstances for which pared-down versions of the treatment and service may be requested. Outcomes attained or not attainable through such minimized versions must be understood. Tests of augmented models of the protocol can fit within this step as well.

### *Step 7: Goodness of fit within the organizational or practice context*

This step involves a series of studies to assess organizational characteristics that may influence the willingness of practitioners to adopt or use the protocol or the ability of the institution, agency, or practice setting to sustain the service with fidelity to improved outcomes. Studies in this step include examination of features of the organizational culture or context that impede or facilitate the uptake or adoption of the new protocol. Fortunately, measures for assessing organizational context within mental health clinics exist and can be used to develop profiles of key dimensions of context (e.g., climate, culture, work attitudes) that can target areas for intervention (Glisson et al., 2008). Variables such as levels of family activation in treatment planning, workplace flexibility, practitioner autonomy, leadership style, productivity requirements, incentive structure, workplace staff turnover, practitioner motivation, and attitudes toward change may influence the ability of clinicians or providers to use the new service and the ability of the organization or practice environment to sustain it (Glisson & James, 2002). Models for attending to implementation processes have also been developed and can be used to guide conceptually the goodness of fit within various practice contexts (Fixsen et al., 2005). In particular, the role—either structured or unstructured—for family participation can be a key contextual indicator of organizational readiness (Hoagwood, 2005; Hoagwood et al., 2008).

### *Step 8: Dissemination and quality in a variety of organizational or practice contexts*

The difference between this final step and the previous one is that this involves a series of studies about the ability to disseminate the effective treatment or service to multiple agencies, organizations, clinics, or communities. The purpose of this step is to examine the range of variations in organizational culture or context across multiple practice settings that facilitate the uptake of the effective service, generate long-term outcomes, and sustain the service within the sociology of improved practice. Improvements in practice depend on knowledge of those structural elements of agencies, clinics, schools, or other practice environments that interfere with or support the quality of care and the delivery of quality services (e.g., fiscal structure; extent to which families are engaged in treatment development, selection, or delivery; clinician autonomy or motivation; workplace flexibility). Studies of variations in practice environments that enable providers to deliver the service with fidelity to the protocol will lead to improvements in the quality of care across a range of delivery agencies.

## Conclusion

As the gap between need for mental health care and its delivery to children and families continues to be painfully large, and because most real-world services are either ineffective (Bickman, 1998; Weisz et al., 1992) or unexamined (Kazdin, 2005), accelerated models for development and testing of service delivery practices and effective interventions must be made available to guide the field.

Creating a usable science requires constant attention to the factors that may undercut or diminish their impact and persistent focus on understanding why and under what conditions services can attain their intended outcomes and can be sustained within communities. Factors that may influence the delivery of services come from many different sources; they may arise from the stressors that surround the family; the barriers to care that family experiences; and policy changes, organizational restructuring, or even elections that remove or replace service programs. The constantly changing contexts of service delivery create uncertainties that must be accepted as normative for those undertaking field research. The rewards of such work for the researcher, however, and for the families and children who stand to benefit from a stronger science base relevant to their lives, far outweigh the challenges.

---

POINTS TO REMEMBER

- Steps toward launching scientifically valid field research
  - Step 1: Openly acknowledge the differing purpose between the goals of research and goals of the practice setting, then identify the shared aims for both.
  - Step 2: Create an advisory board to guide the research project.
  - Step 3: Be clear about roles and boundaries. Set up organizational structures to respect those roles and manage the project, but remain flexible and be prepared to change these structures if new needs arise.
  - Step 4: Structure the research project to reflect collaboration between research and field staff. Be flexible.
  - Step 5: Be prepared for the funding priorities of the grantor to shift. If they do, remain flexible and responsive but do not compromise the integrity of the study.
  - Step 6: Make the budget, staffing, and research goals explicit and transparent.
  - Step 7: Share the credit and do not surprise colleagues.
- The community intervention development model
  - Step 1: Developing and creating a manual for the Protocol.
  - Step 2: Efficacy trial.
  - Step 3: Single-case applications.
  - Step 4: Initial effectiveness trial.
  - Step 5: Full effectiveness trial.
  - Step 6: Effectiveness of moderators and mediators.
  - Step 7: Goodness of fit within the organizational or practice context.
  - Step 8: Dissemination and quality in a variety of organizational or practice contexts.

## REFERENCES

Atkins, M. S., Graczyk, P., Frazier, S. L., & Adil, J. A. (2003). Toward a new model for school-based mental health: Accessible, effective, and sustainable services in urban communities. *School Psychology Review, 12,* 503–514.

Atkins, M. S., McKay, M., Arvanitis, P., London, L., Madison, S., Costigan, C., Haney, P., Zevenbergen, A., Hess, L., Bennett, D., & Webster, D. (1998). An ecological model for school-based mental health services for urban low income aggressive children. *The Journal of Behavioral Health Services & Research, 5,* 64–75.

Bickman, L. (Ed.). (1990). *Advances in program theory. New directions for program evaluation, No. 47.* San Francisco: Jossey-Bass.

Bickman, L. (1998). A continuum of care: More is not always better. In M. E. Hertzig & E. A. Farber (Eds.), *Annual progress in child psychiatry and child development* (pp. 395–413). Philadelphia, PA: Brunner/Mazel.

Burns, B. J., & Hoagwood, K. (2004). Evidence-based practice, part I: Research update. *Child and Adolescent Psychiatric Clinics of North America, 13,* xi–xiii.

Burns, B. J., Hoagwood, K., & Mrazek, P. J. (1999). Effective treatment for mental disorders in children and adolescents. *Clinical Child and Family Psychology Review, 2,* 199–254.

CATS Consortium. (2007). Implementing CBT for traumatized children and adolescents after September 11: Lessons learned from the Child and Adolescent Trauma Treatments and Services (CATS) project. *Journal of Clinical Child and Adolescent Psychology, 36,* 581–592.

Elliott, D. J., Koroloff, N. M., Koren, P. E., & Friesen, B. J. (1998). Improving access to children's mental health services: The family associate approach. In M. H. Epstein, H. Michael, K. Kutash, & A. Duchnowski (Eds.), *Outcomes for children and youth with emotional and behavioral disorders and their families: Programs and evaluation best practices* (pp. 581–609). Austin, TX: PRO-ED.

Fixsen, D. L., Naoom, S. F., Blasé, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Frazier, S. L., Cappella, E., & Atkins, M. S (2007). Linking mental health and after school systems for children in urban poverty: Preventing problems, promoting possibilities. *Administration and Policy in Mental Health and Mental Health Services Research, 34,* 389–399.

Glisson, C. (2002). The organizational context of children's mental health services. *Clinical Child and Family Psychology Review, 5*(4), 233–253.

Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior, 23*, 767–794.

Glisson, C., Landsverk, J., Schoenwald, S., Kelleher, K., Hoagwood, K. E., Mayberg, S., et al., & The Research Network on Youth Mental Health (2008). Assessing the Organizational Social Context (OSC) of mental health services: Implications for research and practice. *Administration and Policy in Mental Health and Mental Health Services Research, 35*, 98–113.

Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge, MA: MIT Press.

Hatch, J., Moss, N., Saran, A., Presley-Cantrell, L., & Mallory, C. (1993). Community research: Partnership in black communities. *American Journal of Preventive Medicine, 9*(6, Suppl.), 27–31.

Hoagwood, K. E. (2005). Family-based services in children's mental health: A research review and synthesis. *Journal of Child Psychology and Psychiatry: Annual Research Review, 46*(7), 690–713.

Hoagwood, K. E., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services, 52*, 1179–1189.

Hoagwood, K. E., Burns, B. J., & Weisz, J. R. (2002). A profitable conjunction: From science to service in children's mental health. In B. J. Burns & K. Hoagwood (Eds.), *Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders* (pp. 327–338). New York: Oxford University Press.

Hoagwood, K. E., Green, E., Kelleher, K., Schoenwald, S., Rolls-Reutz, J., Landsverk, J., et al., & The Research Network on Youth Mental Health. (2008). Family Advocacy, Support and Education in Children's Mental Health: Results of a National Survey. *Administration and Policy in Mental Health and Mental Health Services Research, 35*, 73–83.

Hoagwood, K. E., Vogel, J. M., Levitt, J. M., D'Amico, P. J., Paisner, W. I., & Kaplan, S. J. (2007). Child and Adolescent Trauma Treatments and Services Consortium. Implementing an evidence-based trauma treatment in a state system after September 11: The CATS project. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(6), 773–779.

Hohmann, A. A., & Shear, M. K. (2002). Community-based intervention research: Coping with the noise in real life in study design. *American Journal of Psychiatry, 159*, 201–207.

Horkheimer, M., & Adorno, T. W. (1998). *Dialectic of enlightenment*. New York: Continuum Publishing.

Horwitz, S. M., Leaf, P. J., Leventhal, J. M., Forsyth, B., & Speechley, K. N. (1992). Identification and management of psychosocial and developmental problems in community-based, primary care pediatric practices. *Pediatrics, 89*, 480–485.

Jensen, P., & Hoagwood, K. E. (Eds.). (2008). *Improving children's mental health through parent empowerment: A guide to assisting families*. New York: Oxford University Press.

Jensen, P. S., Hoagwood, K. E., & Petti, T. (1996). Outcomes of mental health care for children and adolescents II: Literature review and application of a comprehensive model. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 1064–1077.

Jensen, P. S., Hoagwood, K. E., & Trickett, E. J. (1999). Ivory towers or earthen trenches? Community collaborations to foster real-world research. *Applied Developmental Science, 3*(4), 206–212.

Kazdin, A. E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology: Science and Practice, 12*(2), 184–188.

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology, 66*(1), 19–36.

Koroloff**,** N. M., Elliott, D. J., Koren, P. E., & Friesen, B. J. (1994). Connecting low-income families to mental health services: The role of the family associate. *Journal of Emotional and Behavioral Disorders, 2(4), 240–246.*

Koroloff, N. M., Osher, T., Jivanjee, P., Pullman, M. D., Sofich, K., Guthrie, L., et al. (in press). Under new management: Research collaboration with family members and youth. In K. Hoagwood, P. S. Jensen, M. McKay, & S. Olin (Eds.), *Community collaborations in child mental health research: A casebook*. New York: Oxford University Press.

McCormick, A., McKay, M. M., Wilson, M., McKinney, L., Paikoff, R., Bell, C., et al. (2000). Involving families in an urban HIV preventive intervention: How community collaboration addresses barriers to participation. *AIDS Education & Prevention, 12*(4), 299–307.

McKay, M. M. (in press). Collaborating with consumers, providers, systems and communities to enhance child mental health services research. In K. Hoagwood, P. S. Jensen, M. McKay, & S. Olin (Eds.), *Community collaborations in child mental health research: A casebook*. New York: Oxford University Press

McKay, M. M., & Bannon, W. M. (2004). Engaging families in child mental health services. *Child and Adolescent Psychiatric Clinics of North America, 13*, 905–921.

McKay, M., Hibbert, R., Hoagwood, K., Rodriguez, J., Legerski, J., Fernandez, D., & Murray, L. (2004). Integrating evidence-based engagement interventions into ‘real world’ child mental health settings. *Journal of Brief Treatment and Crisis Intervention, 4* (2), 177–186.

McKay, M. M., Stoewe, J., McCadam, K., & Gonzales, J. (1998). Increasing access to child mental health services for urban children and their care givers. *Health and Social Work, 23*, 9–15.

National Advisory Mental Health Council. (1999). *Bridging science and service: A report by the National Advisory Mental Health Council's Clinical Treatment and Services Research Workgroup*. Retrieved February 1, 2007, from www.nimh.nih.gov/research/bridge.pdf (NIH Publication No. 99-4353).

New Freedom Commission on Mental Health. (2003). *Achieving the promise: Transforming mental health care in America*. Retrieved February 1, 2007, from www.mentalhealthcommission.gov/reports/FinalReport/toc.html (DHHS Publication No. SMA-03-3832).

Olin, S., Hoagwood, K., Rodriguez, J., Burton, G., Ramos, B., Radigan, M., et al. (in press). The application of behavior change theory to family-based services: Improving parent empowerment in children's mental health.

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52*(9), 1190–1197.

Singer, G., Irvin, L., Irvine, B., Hawkins, N., Hegreness, J., & Jackson, R. (1993). Helping families adapt positively to disability: Overcoming demoralisation through community supports. In G. Singer & L. Powers (Eds.), *Families, disability, and empowerment: Active coping skills and strategies for family interventions*. Baltimore: Paul H. Brookes Publishing.

Weisz, J. R. (2003). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples*. Cambridge, UK: Cambridge University Press.

Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005a). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology, 56*, 337–363.

Weisz, J. R., Jensen, A. L., & McLeod, B. D. (2005b). Development and dissemination of child and adolescent psychotherapies: Milestones, methods, and a new deployment-focused model. In E. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent* *disorders: Empirically based strategies for clinical practice* (2nd ed.) (pp. 9–39). Washington, DC: American Psychological Association.

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47*(12), 1578–1585.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117*, 450–468.

## 2

# Developing Questions when the Perfect Instrument Is Not Available

*Sarah McCue Horwitz and Kimberly Eaton Hoagwood*

> Far better an approximate answer to the right question than the exact answer to the wrong question, which can always be made precise.
> —John Tukey, 1962, p. 13

Once an important research question is identified and a strong study design to investigate that question is developed, the issue of choosing the best measures for key variables must be confronted. Numerous valid and reliable instruments are available, but often they do not adequately capture important concepts or must be modified for use in the subjects intended for the study. This chapter briefly addresses some of the principles of measurement and, by using the development of a measure of children's mental health services utilization, illustrates the process of instrument development and some of the challenges faced when designing a new instrument.

## Background

Measurement is the process of specializing and putting a concept into operation. Specifying the concept involves refining the key factors and finding or developing a clear, consistent definition for them (Shi, 1997).

Specification of concepts often is guided by the theoretical frame of the study as well as the perspective or approach of the investigator. Operationalization makes the conceptual definition more precise by linking it to one or more concrete indicators (Connor, 2006; Shi, 1997). Once the indicators have been selected (e.g., physical health status) and the sources for the indicators have been selected (e.g., adult self-report), an investigator usually searches for instruments to measure the indicators. As pointed out in an excellent article on research instruments by Switzer and colleagues (1999), two issues are central for instrument selection: context and psychometrics (Switzer, Wisniewski, Belle, Dew, & Schultz, 1999). Context is generally thought of as factors outside the actual assessment tool. Some important contextual factors include characteristics of the study subjects, goals of the research, and constraints on data gathering (Switzer et al., 1999). Context is important because it may frame subjects' responses to the selected instrument. Given that most instruments are developed on middle-class, Western-European or North-American assumptions, values, and norms, they may have different meanings for individuals from different cultures. Similarly, as pointed out by Switzer et al., historical context may be important because societies change regarding knowledge, beliefs, values, and attitudes (Switzer et al., 1999). For example, how knowledge is acquired (through observation, experimentation, intergenerational transmission) and whether it is validated (through consensus, standardized measurement, inferential reasoning) constitute core characteristics of societies or communities because they represent different ways of consolidating and transmitting knowledge. These ways of knowing vary across cultures and across historical periods and are important to acknowledge when defining the questions of interest for any particular study.

Finally, there are the real practical issues of administration. The costs associated with obtaining the information of interest, including purchasing the instrument, choosing the best respondent, and recognizing the limitations of the data-gathering strategy (e.g., young children usually cannot read a questionnaire) are real constraints and must be taken into account before undertaking a study (Shi, 1997; Switzer et al., 1999).

Another core step in selecting appropriate instrumentation is examining properties of the instrument itself. For example, the psychometric properties of an instrument indicate whether the instrument measures the concept of interest (validity) and whether it measures the concept consistently (reliability) (Carmines & Zeller, 1979; Connor, 2006;

Shi, 1997; Switzer et al., 1999). The psychometric evaluation of instruments is a well-developed science. Reliability of an instrument consists of two central issues: (1) do the items on a scale measure a single construct? (2) Do scales produce consistent estimates across multiple measures (Connor, 2006)? The measurement of a single construct usually is evaluated by the internal consistency of the scale using Cronbach's alpha. The consistency across multiple measurements can be evaluated in several ways, including test-retest (assesses the same individuals at a second time point), alternate form (assesses the same individuals with a similar version of the same instrument), split-half (assesses the same individuals with an instrument that has been divided in half), and inter-rater (assesses the same individuals with two different raters; Carmines & Zeller, 1979; Connor, 2006; Shi, 1997; Switzer et al., 1999). Most consistency ratings are measured by the family of Pearson correlations.

Validity is another property of an instrument. Validity measures the extent to which an instrument measures what it claims to measure or the "truth" of the instrument. It can be determined in several ways. Three common ways are face or content, criterion, and construct validity. Face, or content, validity asks whether the instrument makes sense and measures the scope of the construct. For criterion validity, the question is how well the instrument agrees with any gold standard. It can be measured concurrently (instrument and gold standard are measured at the same time) or predictively (the instrument under study is used to predict the gold standard or criterion). The third and most complex type of validity is construct. Here the question is whether the instrument measures the underlying construct of interest (Carmines & Zeller, 1979; Connor, 2006; Switzer et al., 1999).

Measures usually fall into one of four types: established, modified, hybrid, or new (Connor, 2006). Established measures are those commonly used to measure the construct of interest. The key question with established measures is whether they have been used in the population under investigation (Connor, 2006). Modified instruments have been changed in some way. The important questions to ask about modified instruments are whether the rationale for modifying the instrument was clear and whether the psychometric properties of the modified instrument have been established (Connor, 2006). Hybrid instruments are created by combining items from established scales or by adding items to an established scale. Again, the rationale for the change and the psychometric properties of the hybrid scale are important questions (Connor, 2006).

Finally, the development of a new instrument can be undertaken. This is a long, difficult, and sometimes expensive process whose success is not guaranteed. The example that follows illustrates the process for developing a new instrument and the amount of work necessary to establish its psychometric properties.

## Example

Knowledge concerning the services children receive for mental health problems is critical because many (14% to 25%) children in the United States meet criteria for a psychiatric diagnosis and are functionally impaired but report receiving no services for these problems (American Psychiatric Association, 1994; Anderson, Williams, McGee, & Silva, 1987; Brandenberg, Friedman, & Silver, 1990; Costello, 1989; Costello et al., 1996; Costello, Messer, Bird, Cohen, & Reinherz, 1998; Offord et al., 1987; Shaffer et al., 1996). Furthermore, little information is available about the few services that are received (Burns et al., 1995; Leaf et al., 1996; Stiffman, Chen, Elze, Dore, & Cheng, 1997; Wang, Berglund, Olfson, Pincus, Wells, & Kessler, 2005; Wang, Lane, Olfson, Pincus, Wells, & Kessler, 2005; Zahner, Pawel, DeFrancesco, & Adnopoz, 1992). Because of the paucity of information about children's mental health services, a 12-member working group of services researchers, epidemiologists, and child mental health experts was formed by the National Institute of Mental Health in 1996 under the leadership of the senior author. The purpose of the working group was to review the existing literature and instruments measuring mental health service utilization for children and adolescents. This work was done in preparation for a large national study of children's mental health.

### *Issue 1: Refining the key factors, developing clear definitions, and selecting the level of detail*

The review of mental health services literature identified several gaps in knowledge about children's utilization of services. First, to examine the predictors of onset and cessation of utilization, services data must be collected in meaningful units (Kessler, Steinwachs, & Hankin, 1980).

Episodes of care are considered the boundaries for summing all inputs into the care process by specifying beginning and ending points (i.e., the units) and the course of services used during treatment (Hornbrook & Berki, 1985, Keeler, Manning, & Wells, 1988; Keeler & Rolph, 1988; Wells, Keeler, & Manning, 1990). However, few studies had examined episodes of care or their applicability to children's mental health problems.

A second gap involved service trajectories. Predictors of and pathways into, between, and out of mental health services is an important issue that has received relatively little attention in the children's mental health field. Although earlier studies documented the need for additional or more effective mental health services (Knitzer, 1982; Stroul & Friedman, 1986), the exploration of pathways has been and continues to be inadequate (Burns et al., 1995; Costello et al., 1988; Costello et al., 1998; Farmer, Stangl, Burns, Costello, & Angold, 1999; Stiffman et al., 1997; U.S. Office of Inspector General, 1991; Weisz & Weiss, 1991).

A third issue relevant to service utilization involves the type, intensity, content, and coordination of care received. A number of studies have concluded that coordination of services is necessary, but not sufficient, for better outcomes (Bickman, 1996; Glisson, 2002; Glisson & Hemmelgarn, 1998; Glisson & James, 2002; Goldman, Morrissey, & Ridgely, 1994; Lehman, Postrado, Roth, & McNary, 1994; U.S. Office of Inspector General, 1991; U.S. General Accounting Office, 1977). These studies, along with those that examine the effectiveness of protocol-driven services, point to the need to understand the type, intensity, and content of care received (Weiss, Donenberg, Han, & Weiss, 1995; Weisz, 2004; Weisz, Huey, & Weersing, 1998; Weisz & Jensen, 1999; Weisz, Weiss, & Donenberg, 1992).

Finally, because few studies have examined how racial identity, acculturation, and cultural mistrust affect help-seeking patterns, these issues constitute another large gap in knowledge about service use. Increasingly, these issues are recognized as central to the development of valid measures (Costello et al., 1997). For example, programs for children with serious emotional disturbances must attend to the unique, culturally related characteristics of these children, their families, and their communities because several studies have found that ethnic minorities under use particular services, even when barriers related to income and availability are reduced or eliminated (Broman, 1987; Chung & Snowden, 1990; McMiller & Weisz, 1996; Scheffler & Miller, 1989; Takeuchi, Leaf, & Kuo, 1988; Terrell & Terrell, 1984; Weisz & McMiller, 1997).

To determine whether key information about children's use of mental health services could be obtained by using instruments available at the time, the working group next evaluated extant instruments. The National Institutes of Mental Health (NIMH) sponsored Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) service use questions (the SURF) could not identify episodes; provide any information on pathways; or disaggregate provider from setting, content, or assessment of care (Leaf et al., 1996).

The Child and Adolescent Service Assessment (CASA) demonstrated good to excellent parent and child 3-month test-retest reliability and similar validity but could not disaggregate type of provider from place of service nor pathways into care (Farmer, Angold, Burns, & Costello, 1994; Ascher et al., 1996).

The Services for Children and Adolescents, Parent Interview (SCA-PI) developed for the Multisite Treatment Study for Attention Deficit Hyperactive Disorder was also evaluated. When reviewed, the instrument had no youth version and could not address either lifetime or 1-year use.

The Referral Sequence and Problems Interview (RSPI) details the process of entry into mental health services (Weisz, 1996). However, it was not designed to disaggregate type of service from provider and setting or to assess the duration, intensity, or content of services received.

Given that no instrument available at the time of the review could assess all the domains identified in the literature review, the committee concluded that a new instrument was necessary. The strategy was to modify the CASA with additions from other instruments. The new instrument was constructed in modules by service setting and was structured so that, with longitudinal assessment over defined, consistent periods, episodes of care could be constructed.

As the instrument was to be used across the United States, the committee selected general service system terminology to encompass the range of diversity in the service system. Similarly, determining the level of detail parents were likely to know about their children's services use also posed a challenge, particularly for cost data. Only well-structured pretests could determine whether the language chosen and level of detail was appropriate.

### *Issue 2: Structuring the services questions independently or tying them to another core construct–in this case measurement of psychopathology*

A common strategy when assessing health or mental health services is to link those services to a particular problem. Two issues arise with using a linked strategy. First, respondents quickly learn that when they respond positively to a gate question, multiple additional questions follow but if they say "No," fewer questions are asked. This can produce attenuated responses. The second issue is somewhat more subtle. The assumption that parents seek services for particular mental health issues might artificially constrain parents' responses because research has shown that parents often seek services for more general issues, such as school problems, rather than for a specific problem such as depression (Yeh et al., 2005).

### *Issue 3: Pretests serve multiple purposes, and more than one may be needed*

Pretests are useful for a host of issues. They help refine questions, reveal better ways of structuring the order of questions, demonstrate feasibility, shed light on the acceptability of instruments, and establish the psychometrics of instruments. When developing a new instrument, multiple pretests may often be needed and serve multiple purposes. For example, the initial draft of the instrument (Services Assessment for Children and Adolescents [SACA]) was first pretested with parent and child volunteers from one inner-city child outpatient mental health clinic to determine feasibility and acceptability. However, based on the responses of the volunteers, the instrument was completely restructured to allow the historical unfolding of a child's mental health services utilization.

The formal reliability and validity work was undertaken for an interview schedule that included the SACA and was carried out in two sites with two different goals. In one site, the pretest focused on establishing the reliability and validity of the instruments using subjects recruited from clinics, public schools and a day-care facility. In the second site, the focus was on both psychometrics and pretesting of field procedures thus demanding a different recruitment strategy. Clearly, the recruitment procedures in the two pretest sites were different, yet each was necessary to provide the full range of information necessary to judge the strength of the newly developed instrument (Horwitz et al., 2001).

### *Issue 4: Validity: Often no gold standard exists*

As previously mentioned, validity is the truth of the instrument: does it measure what it is supposed to measure? Ideally, the new instrument will demonstrate good to excellent agreement with some established gold standard.

The gold standard for utilization is the medical record. Ideally, a medical record captures the encounter with sufficient detail to recreate the details of the care delivered within the visit. However, medical records are notoriously inaccurate. They are lost, incomplete, and sometimes illegible. Although electronic medical records are likely to improve the quality of recorded information, they were not in use where the validity work for the SACA was undertaken. Thus, to determine validity, all service agencies reported by individuals at one of the pretest sites were contacted. Although records were usually available, considerable amounts of data were missing, thus preventing validation of some of the constructs of interest (Hoagwood et al., 2000; Horwitz et al., 2001; Stiffman et al., 2000).

### *Issue 5: Competing demands sometimes prevent the timely dissemination of results*

The development of the SACA occurred within the planning for a large national study of children's mental health. Therefore the data from the pretest were quickly reviewed and the study investigators, including the working group of services researchers, turned their attention to fielding the national study. Consequently, although the initial analyses were quite promising, thorough data analyses and manuscript preparation did not occur in a timely manner.

### *Issue 6: Continued funding for the dissemination of the new instrument is not ensured*

The SACA represents an interesting example of the issues associated with dissemination. Considerable time and effort went into the development and testing of the instrument. The psychometric properties of the instrument proved to be quite good, and it was acceptable to parents. However, because it was developed by a team as part of a larger study, the responsibility and timeline for dissemination was unclear. The dissemination issue was further obscured by the fact that the large national study was aborted before data collection. Consequently, no one was responsible for disseminating the SACA, and no policy at NIMH governed the dissemination of instruments developed with NIMH funding. Until she left the NIMH, the second author assumed the responsibility for dissemination, and the first author currently handles dissemination.

This "hit or miss" dissemination is inefficient, does not promote a clear line of instrument development, and does not promote an efficient use of federal funds. Alternatively, investigators who develop instruments can support them by making them proprietary and selling the instrument (e.g., the Child Behavior Checklist) or through ongoing grant or contract support from federal and private funds (e.g., the Diagnostic Interview Schedule for Children). All three strategies have limitations. In the absence of a clearinghouse or a central repository of instruments that can review, update, and disseminate instruments as well as the scoring instructions, manuals, normative data, and so forth, the field researcher is left to fend for himself or herself. This often means contacting individual instrument developers; tracking down updates that may as yet be unpublished; and trying to identify overlaps and redundancies in items, constructs, or measures. This is a laborious undertaking.

Publishing companies, notably Oxford University Press, among others, are taking advantage of newer Web-based technologies to provide this kind of centralized repository for guidebooks, manuals, and other publications. This entails providing ongoing updates, status reports, and revisions for their publications in certain areas. In the near future, other publishing companies likely will provide similar technologic supports for researchers.

## Conclusion

Identifying core constructs and reliable and valid instruments that target those constructs are essential steps in undertaking scientific studies. The selection of appropriate instrumentation requires that investigators consider the context, including the setting, in which the study will take place; the psychometric properties of the instruments available for use; and the practical exigencies (e.g., costs, types of respondents, time, burden) that may constrain administration of the selected measures.

Given the challenges inherent in developing an instrument, using or modifying an existing instrument should be considered over developing a new one. Furthermore, given the uncertainty of federal funding for the support of instruments once they are developed, investigators must be willing to support dissemination through other means.

---

## POINTS TO REMEMBER

- Principles of measurement
  - Specializing and putting a concept into operation.
    - Clear, consistent definition.
    - Guidance by the theoretical frame.
- Searches for instruments to measure the indicators
  - Two issues central for instrument selection: context and psychometrics.
  - Real practical issues of administration: costs, types of respondents available, time burden.
  - Psychometric properties of an instrument: reliability and validity.
  - Measures usually fall into one of four types: established, modified, hybrid, or new.
- Challenges faced when designing a new instrument
  - Long, difficult, and sometimes expensive process whose success is not guaranteed.
- Issues in designing new instrument
  - Issue 1: Refining the key factors, developing clear definitions, and selecting the level of detail.
  - Issue 2: Structuring the services questions independently or tying them to another core construct—in this case measurement of psychopathology.
  - Issue 3. Pretests serve multiple purposes, and more than one may be needed.
  - Issue 4: Validity: often no gold standard exists.
  - Issue 5: Competing demands sometimes prevent the timely dissemination of results.
  - Issue 6: Continued funding for the dissemination of the new instrument is not ensured.

---

## REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (4th ed.), Washington, DC: American Psychiatric Association.

Anderson, J. B., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in pre-adolescent children. *Archives of General Psychiatry, 44,* 69–76.

Ascher, B. A., Farmer, E. M. Z., Burns, B. J., & Angold, A. (1996). The Child and Adolescent Services Assessment (CASA): Description and psychometric. *Journal of Emotional and Behavioral Disorders, 4,* 12–20.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychology, 51,* 689–701.

Brandenberg, N. A., Friedman, R. M., & Silver, S. E. (1990). The epidemiology of childhood psychiatric disorders: Prevalence findings from recent studies. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 76–83.

Broman, C. L. (1987). Race differences in professional help seeking. *American Journal of Community Psychology, 15,* 473–489.

Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., et al. (1995). Children's mental health service use across service sectors. *Health Affairs (Millwood), 14,* 147–159.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment. Sage series: Quantitative applications in the social sciences.* Beverly Hills, CA: Sage Publications.

Chung, F. K., & Snowden, L. R. (1990). Community mental health and ethnic minority populations. *Community Mental Health Journal, 26,* 277–291.

Connor, T. (2006). *Measurement in research,* http://faculty.ncuc.edu/toconnor/308/308lect04.htm, accessed July 11, 2006.

Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 836–841.

Costello, E. J., Angold, A., Burns, B. J., Stangl, D. K., Tweed, D. L., Erkanli, A., et al. (1996). The Great Smoky Mountains Study of youth: Goals, design, methods and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry, 53,* 1129–1136.

Costello, E. J., Burns, B. J., Costello, A. J., Edelbrock, C., Dulcan, M., & Brent, D. (1988). Service utilization and psychiatric diagnosis in pediatric primary care: The role of the gatekeeper. *Pediatrics, 82,* 435–441.

Costello, E. J., Farmer, E. M., Angold, A., Burns, B. J., & Erkanli, A. (1997). Psychiatric disorders among American Indian and white youth in Appalachia: The Great Smoky Mountains Study. *American Journal of Public Health, 87,* 827–832.

Costello, E. J., Messer, S. C., Bird, H., Cohen, P., & Reinherz, H. Z. (1998). The prevalence of serious emotional disturbance: a re-analysis of community studies. *Journal of Child and Family Studies, 7,* 411–432.

Farmer, E. M. Z., Angold, A., Burns, B. J., & Costello, E. J. (1994). Reliability of self-reported service use: Test-retest consistency of children's responses to the Child and Adolescent Services Assessment (CASA). *Journal of Child and Family Studies, 3,* 307–325.

Farmer, E. M., Stangl, D. K., Burns, B. J., Costello, E. J., & Angold, A. (1999). Use, persistence and intensity: Patterns of care for children's mental health across one year. *Community Mental Health Journal, 35,* 31–46.

Glisson, C. (2002). The organizational context of children's mental health services. *Clinical Child and Family Psychological Review, 5,* 233–253.

Glisson, C., & Hemmelgarn, A. (1998). The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse and Neglect, 22,* 1–21.

Glisson, C., & James, L. R. (2002). The cross-level effects of culture and climate in human service teams. *Journal of Organizational Behavior, 23,* 767–794.

Goldman, H. H., Morrissey, J. P., & Ridgely, M. S. (1994). Evaluating the Robert Wood Johnson Foundation Program in chronic mental illness. *Milbank Quarterly, 72,* 37–47.

Hoagwood, K. E., Horwitz, S. M., Stiffman, A. R., Weisz, J., Bean, D., Rae, D., et al. (2000). Concordance between parent reports of children's mental health services and service records: The Services Assessment for Children and Adolescents (SACA). *Journal of Child and Family Studies, 9,* 315–331.

Hornbrook, M. C., & Berki, S. E. (1985). Practice mode and payment method: Effects on use, costs, quality and access. *Medical Care, 23,* 484–511.

Horwitz, S. M., Hoagwood, K. E., Stiffman, A. R., Summerfeld, T., Weisz, J. R., Costello, et al. (2001). Reliability of the Services Assessment for Children and Adolescents. *Psychiatric Services, 52,* 1088–1094.

Keeler, E. B., Manning, W. G., & Wells, K. B. (1988). The demand for episodes of mental health services. *Journal of Health Economics, 7,* 369–392.

Keeler, E. B., & Rolph, J. E. (1988). The demand for episodes of treatment in the health insurance experiment. *Journal of Health Economics, 7,* 337–367.

Kessler, L. G., Steinwachs, D. M., & Hankin, J. R. (1980). Episodes of psychiatric utilization. *Medical Care, 18,* 1219–1227.

Knitzer, J. (1982). *Unclaimed children: The failure of public responsibility to children and adolescents in need of mental health services*. Washington, DC: Children's Defense Fund.

Leaf, P. J., Alegria, M., Cohen, P., Goodman, S. H., Horwitz, S. M., Hoven, C. W., et al. (1996). Mental health service use in the community and schools: Results from the four-community MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 889–897.

Lehman, A. F., Postrado, L. T., Roth, D., & McNary, S. W. (1994). Continuity of care and client outcomes in the Robert Wood Johnson Foundation Program in Chronic Mental Illness. *Milbank Quarterly, 72*, 105–122.

McMiller, W. P., & Weisz, J. P. (1996). Help-seeking proceeding mental health clinic intake among African-American, Latino and Caucasian youth. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 1086–1094.

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N.I., Links, P.S., Cadman, D.T., et al. (1987). Ontario Child Health Study II: Six month prevalence of disorders and rates of service utilization. *Archives of General Psychiatry, 44*, 832–836.

Scheffler, R. M., & Miller, A. B. (1989). Demand analysis of mental health service use among ethnic subpopulations. *Inquiry, 26*, 202–215.

Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., et al. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates and performance in the MECA Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 865–877.

Shi, L. (1997). *Health services research methods*. Albany, NY: Delmar Publishers.

Stiffman, A. R., Chen, Y. W., Elze, D., Dore, P., & Cheng, L.C. (1997). Adolescents' and providers' perspectives on the need for and use of mental health services. *Journal of Adolescent Health, 21*, 335–342.

Stiffman, A. R., Horwitz, S. M., Hoagwood, K. E., Compton, W., Cottler, L., Bean, D.L., et al. (2000). The Services Assessment for Children and Adolescents (SACA): Adult and child reports. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 1032–1039.

Stroul, B., & Friedman, R. M. (1986). *A system of care for children and youth with severe emotional disturbances* (rev. ed.). Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center.

Switzer, G. E., Wisniewski, S. R., Belle, S. H., Dew, M. A., & Schultz, R. (1999). Selecting, developing, and evaluating research instruments. *Social Psychiatry and Psychiatric Epidemiology, 34*, 399–409.

Takeuchi, D. T., Leaf, P. J., & Kuo, H. S. (1988). Ethnic differences in the perception of barriers to help-seeking. *Social Psychiatry and Psychiatric Epidemiology, 23*, 273–280.

Terrell, F., & Terrell, S. L. (1984). Race of counselor, client sex, cultural mistrust level and premature termination from counseling among black clients. *Journal of Counseling Psychology, 31*, 371–375.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics 33*(1), 1–67.

U.S. Office of the Inspector General. (1991). *Services integration: A twenty-year retrospective*. Washington, DC: U.S. Department of Health and Human Services, Office of the Inspector General.

U.S. General Accounting Office. (1977). *Returning the mentally disabled to the community: Government needs to do more* (HRD-76-152). Washington, DC: Government Printing Office.

Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 603–613.

Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: Results from the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 629–640.

Weiss, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting Clinical Psychology 63*, 688–701.

Weisz, J. R. (1996). *Studying clinic-based child mental health care*. Los Angeles: University of California.

Weisz, J. R. (2004). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples*. Cambridge, UK: Cambridge University Press.

Weisz, J. R., Huey, S. J., & Weersing, V. R. (1998). Psychotherapy outcome research with children and adolescents. *Advances in Clinical Child Psychology, 20*, 49–91.

Weisz, J. R., & Jensen, P. S. (1999). Efficacy and effectiveness of psychotherapy and pharmacotherapy with children and adolescents. *Mental Health Services Research, 1*, 125–158.

Weisz, J. R., & McMiller, W. P. (1997). Minority help-seeking. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 444–445.

Weisz, J. R., & Weiss, B. (1991). Studying the referability of child clinical problems. *Journal of Consulting Clinical Psychology, 59*, 266–273.

Weisz, J. R, Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychology, 47*, 1578–1585.

Wells, K. B., Keeler, E., & Manning, G. W. (1990). Patterns of outpatient mental health care over time: Some implications for estimates of demand and for benefit design. *Health Services Research, 24*, 773–789.

Yeh, M., McCabe, K., Hough, R.L., Lau, A., Fakhry F., & Garland, A. (2005). Why bother with beliefs? Examining relationships between race/ethnicity, parental beliefs about causes of mental health problems and mental health service use. *Journal of Consulting Clinical Psychology, 75*, 800–807.

Zahner, G. E. P., Pawel, K. W., DeFrancesco, J., & Adnopoz, J. (1992). Children mental health services needs and utilization patterns in an urban community: An epidemiological assessment. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 951–960.

# 3

# Hiring, Training, and Retaining Research Staff and Interviewers

*Elizabeth Mayfield Arnold and Mary Jane Rotheram-Borus*

> I am convinced that nothing we do is more important than hiring and developing people. At the end of the day you bet on people, not on strategies.
> —Larry Bossidy, www.quotatio.com

When conducting field research, one of the most important decisions that the investigator has to make is whom to hire to staff the project. Without capable and dedicated staff, the best research projects may not achieve their intended goals. The decision starts with the hiring process, but training and retaining quality staff members are of critical importance. This chapter focuses on these aspects of field research and provides suggestions about how to make hiring, training, and retaining research staff and interviewers a successful endeavor.

## History

In the past several decades, with the growth of field research, the importance of the role of staff and interviewers has increased. Clinical research conducted in an office setting requires strong research skills and attention to detail. However, when research takes place in the community, additional skills are required from staff and interviewers. Not only must these individuals have research skills and be attuned to the various aspects of the data collection but they also must have other characteristics not always deemed critical to the success of other types of research. The staff members are the "face" of the project to both the community and participants. When these relationships are not strong or when problems occur, the study will be in jeopardy. The significance of hiring the best individuals cannot be overestimated. As others have found in longitudinal studies, the characteristics of the field interviewer can have an impact on retention of participants (Cotter, Burke, Loeber, & Navratil, 2002). Despite the importance of these individuals, little research exists about how to identify, train, and retain quality research staff members who work in the field. Most researchers know very little about hiring practices; thus most of the knowledge is derived from the business world (Kang, Davis, Habermann, Rice, & Broome, 2005).

## Background

This chapter draws on the authors' experiences in implementing human immunodeficiency (HIV) prevention (both authors) and intervention studies (Rotheram-Borus) in community settings. Research on HIV is fraught with numerous complex issues, including issues regarding confidentiality, access to care, and risk behaviors. Highly qualified staff with good interpersonal skills and clear boundaries who can work well with other team members are essential. The "field" in this type of research runs the gamut from community agencies, shelters, clinics, and the streets; thus employees must feel comfortable working with individuals from diverse backgrounds who may present with many problems that can affect the dynamics of the interviews and the data collection process.

### *Hiring*

The first step in the hiring process for identified applicants for field research positions is in-depth, multiperson interviews with existing staff and the investigator(s). These interviews can be conducted in small groups or one-on-one. The initial interview is a great way to assess the appropriateness of the candidate for the position. The interview should provide an opportunity to get to know the individual and assess his or her strengths as a candidate. Although the interviewer should provide information about the position, the emphasis of the interview should be on the interviewee. The interviewer typically should ask open-ended questions that address the applicant's suitability for the current position (Lewis & Gardner, 2000) and allow an opportunity for the candidate to talk about himself or herself and highlight particular accomplishments. The interviewer should also use examples of potential situations to assess the interviewee's skills at handling challenges that will occur on the job (Bozell, 2001). In field research, asking the interviewee to respond to real-life examples from the study can provide excellent insight into the individual's problem-solving abilities.

Paying attention to the interviewee's verbal communication is important both in terms of the content as well as his or her ability to communicate effectively. In addition, the interviewee's nonverbal communication is an indicator of comfort level and personality. If the individual does not make eye contact and appears overly nervous among strangers, the candidate may not be well suited for a job that involves interviewing people he or she does not know.

In describing the position, be sure to describe the role and expectations clearly. As Kang et al. (2005) assert, do not overstate the duties associated with the position: that can lead to job dissatisfaction later. The individual may think he or she is being hired to do one thing and then feel disappointed by the realities of the actual position. Similarly, Kang and colleagues (2005) note that, at times, the opposite phenomenon occurs when the interviewer underplays the position. Later, when the individual is hired, he or she may not be able to live up to the expectations of the position.

For applicants who are interested in working in research but have not done so in the past, understanding their motivations is important. Some individuals may like the idea of doing research but not be at ease with the tasks associated with the project. Others (Levkoff & Sanchez, 2003) have noted that with research studies, the interviewers' own agendas may affect the data if they are uncomfortable with the researcher role. Some interviewers may even unofficially serve as gatekeepers of the research by keeping certain types of individuals out of the study (Levkoff & Sanchez, 2003). These types of behaviors cannot be tolerated, and individuals who might allow their own behaviors to affect the study should not be hired.

Lastly, checking references and verifying that the individual has the credentials he or she claims to have is also important. In particular, one should verify that the individual has the education and/or license that he or she claims to have (Bozell, 2001). For positions in which a clinical license is desired or required, it should be verified that the licensee is in good standing with the regulatory board that governs the particular field.

### *Selecting Staffing Patterns*

In hiring staff for funded research projects, one of the main issues to consider is the best way to use project funds. Full-time staff members are costlier, but in many instances projects need full-time staff. For the employer, the benefit of hiring a full-time staff member is the knowledge that someone is always available. You do not have to wait until the next day that person is scheduled to work. You can take care of problems as they arise. In research, problems in the field can occur at any time, and the ability to respond to situations quickly is of the utmost importance. Furthermore, full-time staff members tend to be more loyal to their jobs because the position likely is their main source of income. They do not have to prioritize the commitment to the project with the demands from another part-time job. From an employee perspective, the main draw for a full-time position tends to be benefits, particularly health care coverage. In some cases, retirement plans or other specific benefits may be a desirable aspect of full-time employment.

Despite the benefits to both the employer and employee, certain costs are associated with full-time employment. Full-time employees are expensive because they typically receive the full package of benefits offered by the institution or agency. Given the ebb and flow of conducting research interviews or performing other research-related tasks, full-time employees will sometimes be less productive. Downtime may occur while waiting for Institutional Review Board (IRB) approval or during the holidays, when participants are less willing to engage in study-related activities. Full-time employees may become bored during the slower times. For staff members who thrive on staying busy, the times with less activity may be difficult and have a negative impact on their motivation.

In contrast, contract interviewers typically are paid only for the time they are actively involved in a study-related task. Interviewers may be paid more than they would be for similar positions in the community and are paid for travel time and mileage as well (Coen, Patrick, & Shern, 1996). However, in almost all instances they do not receive health care or other benefits associated with employment. Thus very little paid time is being used for tasks unrelated to the study, but if the person is not doing interviews on a regular basis, his or her interest in being available to conduct the interviews may lessen. If the person is counting on a minimal amount of income each pay period and no interviews are scheduled, the likelihood of retaining the interviewer long term is not good. Graduate students can be outstanding interviewers, especially if they are highly motivated and organized and are compensated on a per-interview basis. The reimbursement per interview is high, little downtime occurs, and the person is likely to bond well and establish a good relationship in the initial stages. If the bond is accomplished, few calls and tracking will be necessary to contact participants longitudinally.

### *Finding Good Applicants*

Creativity is sometimes required to find the right person for the job. Options for identifying candidates include advertising in newspapers, journals, or professional publications or on the Internet; networking with colleagues at meetings or conferences; and posting positions in the office in case staff may be looking to transfer or know of someone interested in applying (Feuer, 2000). Many times investigators feel more comfortable hiring someone who is recommended by a colleague or employee. Although this method of hiring can be beneficial, someone who performed well in an office-based role may not perform as effectively in the field. In addition, tasks and roles can differ dramatically by study, so the individual must be open to new types of experiences.

### *Selection factors*

*Positive factors.* A variety of desirable attributes should be assessed when checking an applicant's references. Others have noted that some "personal

*Editors note: On one of my projects I ended up dealing with an interviewer pool (unchangeable due to community demands) whose primary motivation was money, not the work itself. I suspected widespread padding of hours and had few interview completions. Switching to a generous per completed interview fee with a bonus of an additional interview few for each 5th, 10th, and 50th completed interview totally solved the problem. The most notorious suspected "padders" quit, and the others were very motivated toward completing more and more interviews. Eventually this motivation even bled through to a real interest in the project itself. ARS* characteristics are hard to quantify but are absolutely essential" (Coen et al. [1996, p. 313]), and we have found some specific qualities in our most successful research staff members. Some of these aspects of the applicant can be hard to assess in an interview but are necessary to be effective in working on a field research study.

First, the individual must enjoy working with people and have *solid social skills* that transcend the setting. The employee likely will be interacting with a variety of individuals from diverse backgrounds, including study participants, other staff members, and community collaborators. The staff member must be able to interact well with others and make others feel at ease. If the applicant is overly anxious in the interview and cannot make conversation, then he or she is not the right candidate. Often someone whose resume lists social skills as strength is so anxious in the interview that communication is tiring and awkward. For more introverted individuals, working on field research project generally is not a good fit. Those who report that they enjoy being "out in field" and meeting new people typically are well suited for these positions.

*Organizational skills*. Organizational skills and attention to detail are also vital to working on a field research project. Conducting interviews or meeting with community partners involves a high level of organization. Meetings or study-related sessions must be highly organized so as not to waste participants' or collaborators' time. Research staff members should come to meetings in the community with a clearly defined agenda that is communicated in advance of the meeting. When implementing the study, all aspects of the protocol must be followed as indicated. Research projects seldom allow for creativity in the implementation of the intervention or the completion of study-related interviews. The individual must understand the importance of the study protocol and not view the project as an opportunity to test new and innovative methods of data collection that are not consistent with the project.

Those who make outstanding employees in any type of work have a *positive attitude* (Feuer, 2000). As Jeffcoat (2002) notes, skills can be taught and improved with training, but a bad attitude typically is not amenable to change through a seminar. Individuals who have a positive attitude in general, but particularly toward their job, likely will have a positive impact on the morale of others in the office. Staff members who are upbeat and excited about their jobs are essential in recruitment and may aid in participant retention if participants view their interactions in a positive manner. In particular, staff members must not negatively label the behaviors of study participants or use pejorative language to describe participants or their problems (Prinz et al., 2001). Staff members who are negative may reduce study participants' enthusiasm for the project or cause participants to drop out because they do not like their interactions with the staff member.

Lastly, the employee must be *committed to making a difference* in the lives of others. For field research, individuals who tend to make the best employees typically feel some sense of commitment to the population or problem being studied. The individual need not feel that the project is the most important issue in his or her life, but should believe that the issue is important and the people being studied matter. The employee should see a sense of value in the work and speak positively about the study, both to other employees and to those outside the agency.

In addition, be sure to hire staff who will *fit well* with the goals of a particular project. Staff members' values should be consistent with those of the project. For example, in Coen and colleagues' (1996) longitudinal study of individuals with severe mental illness, three project values guided the work: (1) respect for the participant; (2) confidentiality; and (3) an emphasis on community relations. Staff working on this project clearly should have these same values. Dress codes often are an issue with interviewers. To bond with high-risk populations and reach participants in their settings, the interviewer should dress casually, not hide tattoos, and so forth. However, having good boundaries is critical and does not require that dress be exactly the same as the participants to bond with them. We have a dress code that states no torn jeans, no shorts, and no t-shirts with slogans on the shirts. Each team must identify its own code.

*Negative Factors.* Some of the negative factors that would make an applicant undesirable for a field research position are not unique to research, and others are more specific to the setting. For instance, someone with *a criminal record* may not be desirable as an employee depending on the type of past criminal behavior and the tasks involved with the job. Although this rule has exceptions, those with histories of criminal behavior may not be well suited to work with participants who may be engaging in risky behaviors. The temptations may be too great for them to be effective in the position. Certainly, this does not mean that these individuals are not employable, but rather that field research jobs might not be the best fit. Those who have been involved in the legal system for filing lawsuits also may make problematic employees. Although having filed suit against someone is not an indication of any wrongdoing, those who are overly litigious may be difficult to supervise or tend to use outside resources to resolve job-related conflicts. These are things to be aware of during the interview and hiring process but individual factors must be taken into account, as well as institution hiring policies (e.g., some institutions have restrictions on hiring individuals with felony convictions).

*Poor personal or social boundaries*. The hiring of individuals with *poor personal or social boundaries* is also problematic. Given the nature of field research, research staff will be going to locations such as community agencies, treatment centers, participants' homes, or even street corners. Scott and White (2005) describe relationship boundaries in longitudinal research as "an arena of ethical ambiguity" (p. S95) for those in the field.

For those who lack clear boundaries, being on someone else's turf may soften or even erase the appropriate professional boundaries. Drinking a cup of tea in someone's home may be the culturally appropriate response to a request to "please sit down while we talk and have something to drink." However, sharing a meal and staying to watch a movie with dinner would be outside of the realm of proper behavior as an employee on a study. This may be obvious, yet finding the boundary while in the field is a struggle for many individuals who may be well meaning but have never been in this type of situation. Because most good interviewers are good listeners, participants may mistakenly believe that an interviewer is their counselor (Scott & White, 2005). Staff members who fail to make the boundaries clear in a professional and caring manner pose problems in field research.

*Trying to save the world*. In addition, individuals who are interested in *trying to save the world* usually are not good employees on research studies. Being out on the field, employees will hear many sad stories from participants about their lives. In some instances, these participants may tell these stories simply to have someone to listen to them talk without expecting a solution. Listening to the participant is appropriate, but so is providing a resource for obtaining services outside the study. The employee should not try to "rescue" the participant, but serve as a resource for obtaining assistance. In other situations, the participant may be trying to persuade the employee to do something for him or her that is not consistent with the protocol. Participants may ask for additional financial compensation to pay for an unexpected expense or hardship. The employee must again realize that the participant cannot be saved from his or her situation or be given any resource not consistent with the study protocol.

*Unusual or erratic past employment.* A history of *unusual or erratic past employment* is a warning sign of a potentially problematic employee. Be on the lookout for individuals who have had high rates of job turnover with unclear or odd explanations for their lack of employment stability. Obviously, there are legitimate reasons for unusual employment patterns, such as taking time off to deal with a person or family illness or returning to school to further one's education. However, some applicants' histories may be indicative of problems that they have had in other work settings. Given the limitations on what many employers will tell you about a former problematic employee, do not assume that a past supervisor will provide you with information on your applicant's work history and job performance at that agency. Instead, be prepared to ask tough and specific questions about the applicant's job history during the interview. Furthermore, listen for statements that may be indicative of the applicant's attitude toward work and ability to get along well with others when talking about past jobs. Warning signs include the applicant reporting having left a previous job because he or she was smarter than the employer or saying negative things about a past employer during the interview (Feuer, 2000). At a minimum, such comments reflect bad judgment about how to conduct oneself in an interview and the ability to maintain a level of professionalism about a past employment situation.

*Psychodynamic orientation.* Finally, individuals with a *psychodynamic orientation* are not well suited for the field studies of today's research world. Most of the field research today is based on cognitive-behavioral or behavioral models that tend to be inconsistent with a psychodynamic perspective. If the individual has a degree in a mental health or human services field and has been working in the field for more than 15 years, he or she likely received at least some psychodynamic coursework or training. However, many of these individuals are well trained in other modalities of treatment and are not entrenched in a psychodynamic perspective. The interview process for candidates should include some discussion of the candidate's theoretical orientation and how this affects his or her view of the world.

### *Hiring from the Community*

Collaborative partnerships with those in the community are essential to recruitment (McCormick et al., 2000) as well as participant retention. Thus, in field research, some believe in hiring a person from the local community who already is connected with the community and understands the local population that is the focus of the research (Eke, Mezoff, Duncan, & Sogolow, 2006). This can be a good idea, but it also can be problematic. Any individual already connected to the community may not be connected by strong, positive bonds. In contrast, the individual's perceived connection to the community may be weak and damaged by the individual's past behavior. Simply knowing people in the community is not in and of itself an advantage when the individual has a negative reputation. For example, the individual may be known for a lack of follow-through on commitments or not being a team player on community projects. Perhaps even more problematic is a situation in which the person has been unprofessional or antagonistic with others. If this individual is hired, he or she may create havoc on the project. As Feuer (2000) notes, be cautious about people who claim to know everyone; instead, hire the best person who can get to know the important people. Those with strong interpersonal skills can build relationships with almost anyone. Furthermore, recognize that community norms of bonding are not sufficient to be an interviewer. Bonding is important, but so is accomplishing a set amount of work on a time schedule and delivering it with a large amount of administrative detail completed consistently.

### *Accountability and Management Information Systems*

All interviewers must be given targets to achieve on a weekly, monthly, and quarterly basis. Provide expectations regarding the number of participants to be contacted and recruited, interviews completed, and those followed longitudinally. A full-time interviewer can recruit and assess two new participants per day in a community setting, conduct 20 follow-up interviews per month on a consistent basis, have a follow-up rate greater than 80% at all times, and receive positive satisfaction ratings from participants on quality assurance evaluations randomly conducted by supervisors. Part-time employees are allowed 5 hours of time to conduct a 2-hour interview, including travel, preparation, and completion of the interview. All interviews are reviewed within 1 week to ensure that the interviews are complete, logged correctly, and that all tracking information is entered into a separate database. To ensure high-quality work and consistent productivity, each project has a tracking database for completion of each interview. We use a basic ACCESS database program tailored for each project. The ACCESS database generates a management information report biweekly or monthly for all supervisors and interviewers on each project. For the principal investigator (PI), this report indicates (1) the number of interviews completed on the project; (2) interviews by ethnicity, age, and gender; (3) any treatment conditions in the study; (4) interviews reflective of the percentage of the cumulative numbers needed; (5) interviews based on the targeted goals set at the beginning of the project; and (6) overall follow-up rates by ethnicity, age, gender, and treatment condition. The same management information system generates a per-interviewer report on the number of interviews per week, the follow-up rate per interviewer, the percentage of targeted interviews completed in the last month, and a cumulative chart per interviewer of productivity over time. These management information reports are critical for managing the team over time. These charts and reports are reviewed at monthly team meetings, where failures to achieve the team's goals are problem solved. Participants who are particularly hard to track or subgroups that are difficult to recruit are identified, and the team brainstorms strategies for improving the recruitment and retention of these participants. We have had interviewers who were specialists in subpopulations (e.g., Latino, HIV-positive women) or who were best at tracking hard-to-find participants. We provided special incentives for these interviewers.

### *Creating Access through Institutional Settings*

On every project, certain agencies have special relationships with the target population. For example, the juvenile justice courts likely have access to homeless youth. The Women, Infants, and Children (WIC) Program is a key resource for serving pregnant populations and those with young children. The Social Security Administration is excellent at tracking all Americans. For each population being studied, especially hard-to-reach populations, institutional agencies can facilitate recruitment and retention. Anticipating which agencies serve the population being studied is an important task. At the beginning of the study, get written permission, if possible, to follow the target participants through these agencies. These permissions can be obtained during a study, but it is often much harder once the study has begun. Credit reporting agencies are excellent sources for tracking adult populations. Interviewers' jobs can become much easier if the team anticipates these relationships and gets written, informed consent ahead of time.

### *Dealing with Human Resources*

The hiring personnel must work closely with the agency or institution's human resources department to ensure that the their staff understand the project needs, which may not be typical and could necessitate creative ways of staffing (Coen et al., 1996). Human resources professionals may not understand why the individual needs a certain type of experience to be a good fit for the position. For instance, if the candidate has experience working in a basic sciences laboratory, this does not mean that he or she will be the best person to conduct field research in homeless settings with individuals with severe mental health problems. However, on paper, the job titles may be the same, creating confusion when trying to fill the position. Furthermore, many interviewers do not work traditional hours and may have difficulty "clocking in" when they arrive and leave work. In fact, they may not even come into the office some days, which may necessitate creative ways of time keeping that may be frowned on by the institution.

Thus, it is important to communicate your needs for hiring early into the staff recruitment and hiring process. There may be institutional requirements about how searches for open positions are filled, such as requiring that internal candidates be given first priority. If the researcher and the HR representative are "on the same page" about the necessary qualifications and experience, it will help prevent wasted time interviewing candidates who are not appropriate for the position. When possible, drafting the actual job description or newspaper advertisement for the position can be very helpful in identifying appropriate candidates. Once an individual is hired, HR departments can be a resource for helping orient the individual to the institution and may also have training resources available at no or little cost to the researcher.

### *Training*

Adequate preparation to work on a field research study involves two phases: orientation and ongoing training. The new employee should receive a thorough orientation to the project and the agency. This initial orientation should be part of a process of ongoing training and learning that takes place throughout the employee's tenure on the project.

## Example: CHIPTS Training

At the UCLA Center for HIV Prevention, Treatment, and Services (CHIPTS), formal procedures are in place to orient new staff to the Center. In the first 2 weeks, the focus is on in-house training. As part of the welcoming process to the Center, new staff members and interviewers meet all current staff. This process is done through both individual introductions and an announcement at staff meetings. On arrival, the new employee has a desk already set up with a computer, office supplies, and other essentials. The next step in the orientation process involves becoming familiar with the specific project to which the individual is assigned.

### *Overview of Agency and Project*

As part of the project orientation, the new employee is given an overview of the office and a copy of the agency's employee handbook. In addition, the employee is told whom to ask about basic office questions, such as using the fax machine and receiving phone messages. Next, the employee is given a copy of all study-related project materials and publications to review. The supervisor then sits down with the individual and familiarizes him or her with the various staff members and their roles and duties.

### *Confidentiality Issues*

The employee's supervisor reviews the regulations regarding confidentiality, including a discussion of the Health Insurance Portability and Accountability Act (HIPPA) and the role of the IRB. Before having any contact with study participants, the new employee should first complete any training required by the funding agency or the institution. For NIH-funded research projects, specific requirements for training must be met.

### *Mock Interviews*

If the individual will be conducting interviews or be in any way involved with the interview process, he or she will receive intensive training in how to conduct study interviews. After reviewing the interview protocol or study assessment forms, the new hire observes an actual subject interview (if appropriate and the subject consents) or a mock interview. The next step is for the employee to participate in a mock interview with another staff member, with the supervisor observing. After successful completion of these tasks, the employee will conduct a real interview with a participant. This interview should be observed, audiotaped, and/or videotaped (with permission from the participant), depending on the project. Detailed, constructive feedback should be provided after each interview. This process should be done several times until the supervisor believes that the individual is ready to conduct interviews alone.

### *Field/Community*

A mandatory part of the orientation is scheduling time for the new employee to "hang out" with key players in the community. This is done for a minimum of 1 week. The time spent in the community should be scheduled with the agency, clinic, or organization, but the purpose of the time should not be rigidly defined. The individuals in the community should be told that the new staff member would like to meet them, understand what they do, and tour their facility (if appropriate depending on the setting). The meeting time should be set up either with the person in charge at the agency or with a staff member who is involved with the project. When the new employee visits each agency, he or she should meet all the key players. Ideally, the contact person for the meeting should handle introductions, but if not, the employee should be instructed to inquire about meeting the other employees. After the week (or more) spent being oriented to the community, the employee should meet with the supervisor to debrief and discuss the meetings.

## Retaining Staff

Many believe that the days of long-term employer–employee relationships driven by loyalty to the employer are gone (Koozer, 2003). A review of the research suggests that departure from jobs is related to job satisfaction, the meaning that the individual attaches to the job, and the individual's relationship with the employing agency (Pockett, 2003). In addition, with grant-funded projects, money may come and go, and employees may sometimes be frightened or apprehensive about working on studies funded by "soft money." Because of these issues, staff must be treated fairly, supported by the organization, and paid well. In addition, their safety should be taken into account when they are in the field.

### *Organizational Culture*

In field research, the creation of a strong, supportive organizational culture that promotes success among employees is vital. The PI and project manager should create an open environment where employees are encouraged to discuss concerns or problems instead of ignoring or burying them. Those in the field need to believe that their opinions and insights matter, particularly because they are on the front lines of the project. Thus interviewers may be an excellent source of information about participants and collaborators, and their opinions should be solicited and valued.

Furthermore, development of a team environment must take place. If you hire people who are committed to the project, you want them to feel invested in what they are doing. Regular team meetings or project meetings should be held in a setting where all staff can participate. Meetings should be an opportunity to review project updates as well as recognize accomplishments related to the study or staff successes. For example, instances commonly occur in field research in which a participant cannot be located and is perceived to be "lost to follow-up." By some creative networking, a staff member may locate the participant and conduct an interview that was never expected to take place. Recognizing this employee's efforts at a team meeting demonstrates the value of his or her efforts, illustrates that determination in tracking participants can pay off, and shows other staff members that such efforts are valued. Another option is an employee recognition program in which staff are given rewards for their contributions (Koozer, 2003). At CHIPTS, we instituted "Field Interviewer Appreciation Day" to recognize the hard work and contributions field interviewers make to research. Light refreshments are provided and the interviewers have the opportunity to interact with each other and the research staff.

On field research teams, difficult situations frequently arise that are upsetting to staff. An interviewer may have a participant get angry about something that is not the interviewer's fault. A participant may not show up for an interview even when the appointment was confirmed earlier in the day and the interviewer drove more than an hour to get to the meeting. When difficult events occur on a study, staff members should be encouraged and supported (Cotter et al., 2002). Without having the option of venting frustrations, being out in the field doing interviews can be quite stressful and lead to burnout.

Part of creating a supportive, team-oriented culture also involves setting clear boundaries and expectations for staff. Although expectations should be discussed as a part of the hiring process, these goals need to be continually reinforced and updated when appropriate, such as when project goals change or are modified. The PI or project manager should be clear about the expectations for personal behavior on the part of the employee. Clearly, personal relationships with study participants should not be tolerated, including any contact "off the clock." Staff also should behave professionally with other employees. Staff members who attempt to create chaos, such as spreading rumors or gossiping about other staff or the organization, should be told directly that such behavior will not be tolerated.

### *Safety*

In field research, one of the unique issues important to retaining staff is safety. Because staff members are going into the field, a potential for danger exists that is not present in an office setting. Staff members often find that working in dyads provides comfort in difficult situations (Simons et al., 1996). On most projects, staff have cell phones paid for by the agency to use when in the field. Staff also should be instructed to let someone on the team know where they are at all times when they are in the field. Staff members can canvass neighborhoods working at different homes, but within a block of each other and with a system for notifying each other in case of trouble. Not taking staff safety seriously will have a negative impact on staff retention and morale and increase liability issues. We have hired security services (e.g., Pinkerton Detective Services in New York City) to train staff to recognize danger, how to approach others to minimize their perception that you are dangerous, how to arrange a room for maximum safety (e.g., make sure you have clear exit access), and how to perform basic self-protective maneuvers.

### *Example: "Golden Rules" regarding Staff Hiring and Retention*

Investigators conducting field research must have some principles and guidelines for hiring and retaining staff. Following are the "golden rules" used at CHIPTS:

1. "When in doubt, do not hire." If you get a bad feeling about someone, or something about the applicant's past experience does not sound right, do not hire the person. At CHIPTS, if one person has doubts, then the candidate for the position is not hired. All that is required is one person saying "no."
2. "Fire quickly when a bad decision has been made." Occasionally the wrong individual is hired for a position. The person may have interviewed well, but later it became clear that the individual was not right for the position. The person may indicate willingness to do the tasks associated with the job but later fail to demonstrate motivation or a strong effort. On the other hand, the individual may have a great attitude but not be suited for the tasks that must be accomplished. Either way, keeping someone on board who is not right for the job is not fair to the individual or to the project. Waiting too long to let the person go can be a big mistake. The person's attitude or poor work performance may have a negative impact on other employees and must be addressed as soon as possible.
3. "When a good interviewer comes along, pay well and promote." Interviewers who are willing to work on a contract basis can be hard to find and retain. When you find someone who is a good fit for this type of position, be sure to pay them above what the current market offers and offer the individual room for advancement if he or she is looking to advance in the field. Some interviewers are content to stay in the interviewer role; give these individuals positive feedback and increases in salary as rewards. If you fail to do so, someone else likely will hire them away from you once their reputation is known in the community.
4. "Define employee goals and requirements early in the hiring process." One of the main ways to ensure that project needs are met is to make the expectations clear to the staff. Many new employees may be hesitant to ask about the office rules and make erroneous assumptions. For example, many interviewers work flexible or odd part-time hours. If a new project staff member sees one individual coming in at 10 a.m. every day, he or she may think coming in late is acceptable. The

reality, though, may be that other project staff members do not have the same flexibility given the responsibilities associated with their position. Simply explaining the office hours and that some individuals work different shifts will alleviate any misperceptions that can have a negative impact on the project. Furthermore, staff members who are new to research may not understand the importance of study-related forms, such as informed consent forms; the employer should clearly relate that some forms and tasks are critical to the implementation of the study. When flexibility in work hours is allowed, random, unscheduled checks in the field are necessary to ensure that the interviewer is where he or she is supposed to be when reporting to you. In addition, randomly follow up with approximately 5% of the interviews by verifying with the participant that the interview occurred as well as the circumstances of the interview.

## Conclusion

In today's competitive research environment, where fewer funds are available to go around, research should be thought of as a business. The interviewers and the research staff are the investigator's connection to the public and the customers—the participants. The participants in the research may never meet the investigator; thus the impressions they have will be based solely on the quality of the interactions with the study staff. If the staff members are unhappy or not committed to the study, the data may be compromised. Although prioritizing staff recruitment and retention is time consuming, failure to do so is a mistake that may be costly in the long run.

---

### POINTS TO REMEMBER

- Hiring staff
    - In-depth, multiperson interviews—small groups or one-on-one.
    - Use examples of situations to assess skills at handling challenges.
    - Clearly describe the role and expectations.
    - Check references and verify credentials.
- Selecting staffing patterns
    - Full-time staff members versus contract interviewers.

    - Graduate students can be good interviewers, especially if motivated, organized, and compensated per-interview.
- Finding good applicants
  - Positive factors: solid social skills, organization skills and attention to detail, positive attitude, committed to making a difference in the lives of others, fit with the goals of the project.
  - Potential negative factors: criminal record, poor personal or social boundaries, trying to save the world, unusual or erratic past employment, psychodynamic orientation.
- Accountability and management information systems
  - Interviewers provided with targets to achieve.
  - Tracking data base for completion of each interview.
  - Team brainstorms problems and strategies for improving recruitment and retention.
  - Special incentives for specialist interviewers.
- Creating access through institutional settings
  - Anticipate which agencies serve the population being studied.
  - Get written permission to follow the target participants through these agencies.
- Training
  - Overview of agency and project.
  - Confidentiality issues.
  - Stages of training:
    - Review the interview protocol or study assessment forms.
    - Observe an actual or mock interview.
    - Mock interview with another staff member with the supervisor observing.
    - "Real interview" with a subject observed, audiotaped, and/or videotaped.
    - Detailed constructive feedback until ready to conduct interviews alone.
- Organizational culture
  - Create a strong, supportive organizational culture that promotes success.
  - Regular team meetings or project meetings.
  - Recognize accomplishments related to the study or staff successes.
  - Set clear boundaries and expectations for staff.
  - Stop staff members who create chaos, spread rumors, or gossip.
- Safety
  - Work in dyads.
  - Cell phones paid for by the agency.
  - Let someone on the team know where they are at all times.

  - Hire security services to train staff to recognize danger, and learn self-protection.
- Think about research as a business

---

## REFERENCES

Bozell, J. (2001). Hiring hints: Be prepared before you begin the interview. *Patient Care Management, 17*(1), 12.

Coen, A. S., Patrick, D. C., & Shern, D. L. (1996). Minimizing attrition in longitudinal studies of special populations: An integrated management approach. *Evaluation and Program Planning, 19*(4), 309–319.

Cotter, R. B., Burke, J. D., Loeber, R., & Navratil, J. (2002). Innovative retention methods in longitudinal research: A case study of the developmental trends study. *Journal of Child and Family Studies, 11*, 485–498.

Eke, A. N., Mezoff, J. S., Duncan, T., & Sogolow, E. D. (2006). Reputationally strong HIV prevention programs: Lessons from the front line. *AIDS Education and Prevention, 18*(2), 163–175.

Feuer, L. (2000). The challenge: Hire the right people for the right reasons. *Case Manager, 11*(3), 24–26.

Jeffcoat, M. K. (2002). Hire for attitude, train for skill. *Journal of the American Dental Association, 133*(11), 1468–1471.

Kang, D. H., Davis, L., Habermann, B., Rice, M., & Broome, M. (2005). Hiring the right people and management of research staff. *Western Journal of Nursing Research, 27*, 1059–1066.

Koozer, G. K. (2003). *Work related problem report: Recruitment, development, and retention of employees at SCCI Hospital*. Mount Vernon, OH: Mount Vernon Nazarine University.

Levkoff, S., & Sanchez, H. (2003). Lessons learned about minority recruitment and retention from the centers on minority aging and health promotion. *The Gerontologist, 43*(1), 18–26.

Lewis, K., & Gardner, S. (2000). Looking for Dr. Jekyll but hiring Mr. Hyde: Preventing negligent hiring, supervision, retention, and training. *Hospital Topics, 78*(1), 14–22.

McCormick, A., McKay, M. M., Wilson, M., McKinney, L., Paikoff, R., Bell, C., et al. (2000). Involving families in an urban HIV preventive intervention: How community collaboration addresses barriers to participation. *AIDS Education and Prevention, 1*, 299–307.

Pockett, R. (2003). Staying in hospital social work. *Social Work in Health Care, 36*(3), 1–24.

Prinz, R. J., Smith, E. P., Dumas, J. E., Laughlin, J. E., White, D. W., & Barron, R. (2001). Recruitment and retention of participants in prevention trials involving family-based interventions. *American Journal of Preventive Medicine, 20,* 31–37.

Scott, C. K., & White, W. L. (2005). Ethical issues in the conduct of longitudinal studies of addiction treatment. *Journal of Substance Abuse Treatment, 28*(Suppl 1), S91–S101.

Simons, P. Z., Rietmeijer, C. A., Kane, M. S., Guenther-Grey, C., Higgins, D. L., & Cohn, D. L. (1996). Building a peer network for a community level HIV prevention program among injecting drug users in Denver. *Public Health Reports, 111,* 50–53.

# 4

# Managing the Data from Survey Development through Archiving

*Peter Doré and Arlene Rubin Stiffman*

> Civilization advances by extending the number of important operations which we can perform without thinking of them.
> —Alfred North Whitehead, 1911, p. 61

Many research texts address issues concerning reliability, validity, and data analysis. Some also discuss survey and questionnaire development (Dillman, 2007; Groves, 1989). However, we could find no publications that examined the continuum of data management from preparation for collection through analyses within the context of the real-world constraints of costs and time. All field researchers confront data management exigencies that must be dealt with to effectively and efficiently implement the data collection portions of the field research.

The goal for any researcher is to choose and use established, highly valid instruments and to use them to collect data without adding any errors due to mistypes, misunderstandings, omissions, or commissions. Yet at the same time, the field researcher must contain costs and time while dealing with the skill level constraints of responders, interviewers, and data entry personnel.

This chapter discusses the issues that researchers confront in meeting these constraints and illustrates each with specific examples of successes and failures. The issues we cover include adapting instruments and questions, selecting data collection venues (computer or paper and pencil), formatting instruments to maximize ease of managing the data through data entry, data cleaning, and archiving. We intentionally do not talk about content of the survey instrument nor give formal theory or advice on psychometric properties or item response theory. Our goal is to share our experiences in trying to keep costs contained while maximizing the smooth flow from data collection through cleaning and analyses to final data archiving. Clean, clear, accurate data mean everything to the success of research. Instrument selection, survey format, data entry, data cleaning, and data preservation cannot be overlooked without dire consequences.

Some literature addresses the issue of developing questionnaires and surveys. Chapter 2 of this book addresses many of these issues. Literature on formatting measures is also available. The key focus of this literature is to make it easier for the respondent and interviewer so as to reduce respondent error (i.e., Dillman, 2007; Fowler, 1995; Groves, 1989; Sudman & Bradburn, 1983). Most such publications emphasize devoting maximal effort during the development stage because later correction of problems is either costly or impossible to implement (DeMaio, 1983). Most published literature focuses on reduction of nonresponse and reduction or avoidance of measurement error. Thus those authors focus on design and graphics. This literature recognizes that respondents take clues from the layout and that poor layout will result in skipped words or misinterpreted questions (Dillman, 2007). Dillman emphasizes that layout must account for visual elements: location (spacing), shape, size, brightness, simplicity, regularity, and consistency of figure-ground formatting. The recent literature includes specific recommendations for placing answer boxes on the right, which is greatly facilitated by the sophistication of today's word processing programs. The rationale for this is to ease the respondent's finding and responding without ignoring skip instructions (Dillman, 2007). Further recommendations concern the use of narrow vertical lists rather than horizontal options. Many of the recommendations for keeping sections of questions visible within a narrower field of vision take account of research on the limitation of the focus of one's eyes (Kahneman, 1973). Recommendations for consistency of formatting assume the inevitable "training" of respondents as they get accustomed to your formatting and therefore might become confused if the formatting changes within the instrument.

This chapter goes beyond the literature by viewing the entire process of collecting and using data as one seamless process, with error reduction at each phase equally important, and with the recognition that when each phase anticipates the next, there are great economies to be gained.

## Adapting Instruments

Despite warnings to never change standardized instruments, it is often necessary to do so to capture your key concepts. Many scales are developed to measure one or more concepts thoroughly. Most likely you will be using many measures to assess many things. If a survey consists of too many long measures, subject burden becomes a major concern, and costs (interviewer time, entry time, subject reimbursement) increase. A related issue is finding a measure that collects only the data related to the study's aims. Although highly correlated with length, multiple subscale measures can lead you off topic if you do not purposefully select relevant parts and stick to your aims. Lastly, every measure included in a survey must be sensitive to the population under study. Can the average member of that population understand the questions, and are the questions relevant for the group based on culture or age, for example? Modifying instruments helps address the problems of length, relevance to aims, and population appropriateness.

Many standardized instruments cover vast topics, far more than you may need. One strategy to address this is to modify the standardized measure. Including irrelevant items increases the opportunity for error and cost at every phase. An excellent article talks about how unused data are dirty data (Orr, 1998). Regarding aims, perhaps the stickiest trap is balancing whether to collect additional data. After all, you already have the subject and asking more questions of each subject is relatively cheaper than creating another study. The lesson for avoiding unused data is to verify that you are choosing instruments that will answer your originally proposed aims (Fowler, 1995).

### *Using Selected Subsections*

At times only selected sections (and by sections, we mean entire subscales) of long-established surveys may be necessary. In our department, we are often interested in depression—not by itself, but how it relates to service use or length of hospital stay, for example. The Computerized Diagnostic Interview Schedule (CDIS) (epi.wustl.edu/dis/DIScdis.htm) provides *Diagnostic and statistical manual of mental disorders*, fourth edition (DSM-IV), diagnostic assessment of nearly every mental health disorder. Conducting the entire assessment takes hours, and we are interested in measuring many other constructs (school or neighborhood environment, insurance status, etc.). We have been able to successfully use only the parts needed in our studies by carefully including only the sections necessary to diagnose the conditions of interest. In the CDIS, the demographics section must be completed to calculate a diagnosis for any condition; because age is important, and occasionally gender or race may be needed for a diagnosis. However, the use of selective pieces of standardized instruments requires caution. The selection of entire subscales is easier, and if using only select items, scoring instructions should be checked to ensure everything is collected to calculate the scores. This is worth repeating: use scoring instructions for the scales as a guide to which items or sections are necessary to score the scale. You would also be wise to seek data from the originator of the scale to test if the items you wish to use are adequate. To use fewer items to save time or be more efficient does neither if your self-created subscale is not valid.

### *Shortening Instruments*

Shortening or modifying standardized instruments may be necessary to save time when an interview becoming too long, creating undue subject burden. Because many standardized measures have been subjected to factor analyses in their development, you can use the data, if published, to shorten them for your needs. If there are multiple factors, selecting only those factors that assess the constructs of interest may be a good idea. Alternatively, to shorten an instrument even further, you can select only the four or so items with the highest loading on the factors, as that will provide the best approximation of the original scale. We were once involved in a project with an extremely careful instrument construction. One full year of weekly meetings was allotted to compile, review, check, and double check the content, flow, and format of the survey. However, with pilot testing, the length revealed itself as a problem. Sections were removed (sections being asked only because we had a captive audience, not the ones related to the study aims), and in some cases individual questions were removed. Even after considerable caution two items from one subscale were removed to save time because it was forgotten that they were part of the algorithm for coding a standardized measure.

The lesson learned is to denote which items are your own original questions and which are necessary parts of standardized measures so you do not accidentally degrade an important variable.

As a data manager, the first author typically avoids these accidental errors from shortening or modifying instruments in two ways: first by listing the variable names used in the original in very small font under the question number in our surveys and, second, by including at the end of a standard measure, again in small font, the citation for the measure. These two techniques are effective in preventing the accidental removal of key items.

### *Crosschecking with the Original*

Remember that colleagues who are prior users of instruments may have adapted them, so after choosing an instrument, we highly recommend going directly to the original source. We had an incident on a pilot study in which a junior faculty member on a limited budget tried to save some time by cutting and pasting instruments in use by a senior faculty member. She assumed the scale was complete and even checked the first five items and the last five items against the original. Unfortunately, at analysis, when the subscales were being calculated, the subscale of primary interest was missing four of five items (items scattered through the middle of the measure). The senior faculty member had been studying a younger population for which those items did not apply, so they were left out (which made perfect sense for that study). The junior investigator was left without a major independent variable that was key to the study aims.

Another pitfall of adaptation when pulling items from other sources is dropped response choices. On one occasion we began to score a scale in which the middle value was somehow left off our version, making our scale 1:4 instead of 1:5, like the original. We adjusted by moving the top two choices up one value and having no middle value, so at least the range was the same—noting, of course, this adjustment in publications.

Ideally, before hitting the field, you will have developed the entry screens and scoring algorithms, and you will have collected, entered, and scored pilot data to ensure no items are omitted. Scoring algorithms leave no room for error. If an item is missing or contains an error, you will find it when you go to score the measure. The lesson learned is that it pays to be ahead of the game, there are benefits to having the entry screens done and the scoring algorithms completed before the first subject is interviewed.

### *Modifying for Cultural Appropriateness*

Another major reason for instrument modification is cultural appropriateness. One of our research projects examined the addictions problems and service use of American Indians (Stiffman et al., 2003). We wanted to use the Diagnostic Interview Schedule (Robins & Helzer, 1994). However, the instrument is based on the use of a variety of individual substances. If a subject uses a substance, then a series of frequency of use and addiction and dependence symptoms is generated. Some American Indians use hallucinogens during certain spiritual ceremonies, and the symptoms generated during these ceremonies should not be counted toward a diagnosis. Therefore, we modified the instrument with a few additional questions by asking the subjects if they used the substance only for spiritual ceremonies or at other times as well. The follow-up questions to arrive at a diagnosis were only generated if subjects use the substances outside spiritual ceremonies (Yu & Stiffman, 2007).

### *Modifying by Adding Items*

Standard scales often do not provide exactly what is needed for a study. One may come close, but perhaps a key component is missing. We recommend resisting the urge to change the items. A better answer may be to add some of your own additional items; you can still compare your findings with other published works by scoring only the original items. Then you can possibly show improved prediction with the addition of your items, or you can show that your items alone are better than the original. Changing the original, however, removes the ability to compare your sample to others because the data are now in the form of apples and oranges.

## Data Collection Venues

Once the instruments have been selected, a few questions must be answered:

1. Will the survey be on paper or computerized?
2. Will the instrument be self-administered (completed by the subject) or administered by a trained professional interviewer?
3. Will it be administered in person or over the phone?

Regardless of the answer to the first question, a paper version is needed to create the computerized version and serve as a backup if computer malfunction occurs in the field. The paper survey is formatted differently if it is the main mode of data collection versus if it is merely a backup to the computer. The importance of the instruments, specifically the physical appearance, is often overlooked in the grand scheme of a research project.

A major decision that field researchers must make is when to collect data with paper and pencil and when to enter data directly into the computer. This is as much art as science—or perhaps an educated judgment call. It all boils down to cost—all costs, including the cost of time and materials. Variables to consider include the number of interviews, length of the interview (number of questions), complexity of the interview, availability of technical skill, budget, and time to field.

In general, as the number of interviews goes up, computerization becomes more cost effective. Computerization is an up-front cost: the price of the computers, the cost to program the entry screens, and the cost to transfer the data to the analysis phase. The paper and pencil method seems inexpensive at first, but as you detail all aspects of the project, costs grow. An interview with more questions costs more to program the entry screens, but again, this is an up-front cost.

To give an example, one of our projects was interviewing 400 youths nine times (quarterly) using an approximately 80-page interview each time. The project used three interviewers to collect the data. Let us compare simply the cost of paper to the cost of three laptops for the interviewers.

Good laptops capable of handling the task of field interviewing can be found for \$1,200 each, for a total of \$3,600. If the price of a photocopied page is \$0.03 each, then $9 \times 400 \times 80 \times \$0.03 = \$8,640$. That makes the computers \$5,040 cheaper than the paper. Obviously, more costs should be considered. It is fairly standard practice to have an editor review and clarify paper surveys before handing them over to data entry. (The data entry person should have no decisions to make during entry. Their job is to enter data as quickly and accurately as possible.)

Editing usually takes about 15 minutes per 60- to 80-page survey, so that gives approximately (9 × 400 × 15)/60 = 900 hours of data editing time. Data entry requires a similar amount of time as editing, for a 60–80-page survey, you have 900 hours of data entry. If we assume editor level employees get about $18 per hour for editing and $14 per hour for data entry, the cost becomes (900 × $18) + (900 × $14) = $28,800. If a programmer for entry screens charges $40 per hour and can program one page an hour (a low estimate), assume 80 pages and nine waves at $40 per hour—dividing by four because most pages are the same from wave to wave and can be reused—plus approximately $5,000 yearly for the software license for entry screens (software is only needed through the programming phase). This results in [(80 × 9 × 40)/4] + $15,000 = $22,200 for programming and software for the entry screens. See websites such as http://www.spss.com/in2quest/ or http://www.sawtooth.com/ or a review such as http://chipts.ucla.edu/TEMPMAT/Broadsheet/Assessment.pdf to help you determine cost of the right package for your needs (we don't mention what we use, because you should evaluate the products based on your individual requirements).

This example shows the creation of the entry screens to be $6,600 cheaper than editing and entry of paper surveys. (Note that some version of entry screens is required to get paper surveys into electronic data, and that cost is not factored in.) So far, the computer is $6,600 + $5,040 = $11,640 cheaper than the paper survey. Of course we are not done, the programmer would need to be paid to convert the data to a statistical package, but this cost applies to both methods so we will call that a wash. Next comes storage of the data for any number of years; CD-R and DVD-R are inexpensive and do not require much storage space, certainly not as much space as 400 8-inch-thick paper files. Do not forget to calculate the storage cost of paper surveys; it can be a significant consideration.

The example mentioned seems to favor computerization dramatically. However, this example only considers the number of interviews and length of the interview. If fewer interviews and questions had been required, at some point the paper version would be less expensive. When other variables are included, this decision begins to involve more judgment. Complexity of interview, for example, approximates an inverted utility curve; as the complexity (number of skip patterns, text substitution of previous answers into current questions, etc.) increases, the cost of computerization increases. But at some point the interview can become too complex for a human being to accurately follow the skip patterns or recall previous answers. At a certain level of complexity the computer begins to become more cost effective relative to the time and errors made by human interviewers trying to follow a paper survey. (Also remember that more-complex paper interviews require more interviewer training and editing time, both of which are costly.) Availability of technical skill is also a cost issue. If reasonably priced programmers are not available, then the cost of computerization can change drastically, making it unfeasible. The budget, of course, plays a factor. In the example mentioned earlier, had the budget been small, the project could have been done without computerization. Cost restrictions require finding volunteer labor, or the principal investigator can serve as the editor and/or entry person, eliminating much of the cost but greatly increasing the time to finish the project. Timeline also can be an important factor. If funding begins on September 1 and you have to be in the field by October 1, computerization will be impossible; it can take months for a medium to large survey to be computerized.

Computerization also offers some intangible benefits, such as spot range checks and validations. For example, if a subject answers that the first time he ever gambled was at age 16, then 15 minutes later states he bought his first lottery ticket at age 12, the computer is much more likely to catch the discrepancy than a human interviewer is. The disparity can then be fixed with the subject immediately. On paper, this may not be caught until a year later, when the data have been entered and logic checks are being run during cleaning or analysis. Computerization offers a time-to-analysis advantage as well. Analysts should be able to review the first few cases minutes after the interviews are done. However, with paper, multiple interviews must be edited and entered before being able to see them electronically.

## Formatting the Paper Survey

So now you know if you are going with a paper survey or a computerized survey. Both require a paper version. The following section discusses how to format and lay out a paper and pencil survey designed to be conducted by trained interviewers. If computers are used, you can spend less time, and have less concern for these issues, because the paper version will rarely be used. In our research shop, we try to use a very standard format for every survey, which helps save time training interviewers and time designing entry screens.

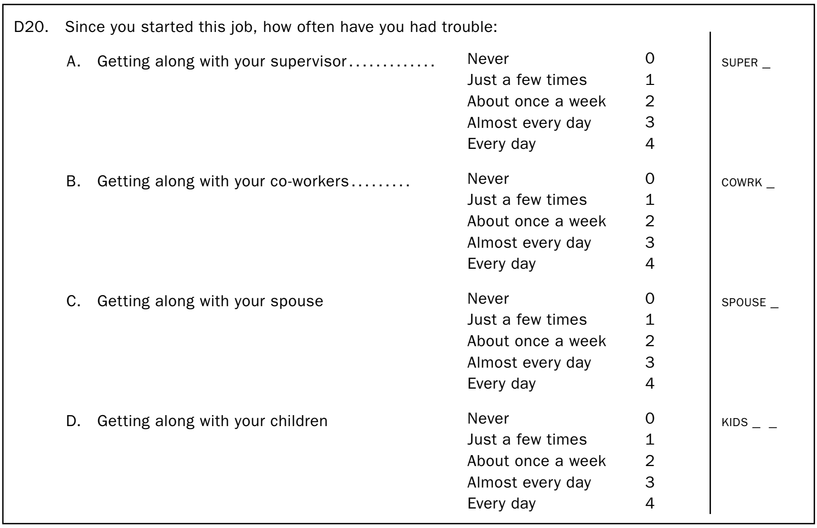
The main focus of much of the published survey format literature is on data collection error (Dillman, 2007; Fowler, 1995; Groves, 1989). A data collection error occurs when the interviewer or respondent mistakenly selects the wrong answer or omits an answer entirely (e.g., checks multiple boxes in a “select one” scenario, misses a row in an answer grid, or circles two responses with one circle). Thoughtful format and layout of the instrument questions, instructions, and answers greatly improve your chances of good, complete, and clean data (Babbie, 2002). The structure, spacing, layout, and consistency of the survey are critical in reducing interviewer and respondent errors. Formatting recommendations differ but are parallel for interviews and self-administered questionnaires. For interviews, recommendations cover choosing standardized conventions to differentiate instructions (for interviewers, for responses, and for skips) and the questions themselves. These differentiations often are made by placing optional wording in parentheses and using capital letters or italics for certain types of instructions. The published literature also includes one recommendation that is often ignored, even by well-known standardized interviews (e.g., the Diagnostic Interview Schedule; Robins & Helzer, 1994). As Sudman and Bradburn (1983) stated “If multiple filter questions are to be asked, try to ask all of them before asking the more detailed questions. Otherwise, respondents may learn how to avoid answering detailed questions.”

### *Layout*

For self-administered questionnaires, the usual focus is on making it self-explanatory, laying it out in a way that seems clear and uncluttered, keeping it short (De Vaus, 1986; Fowler, 1995), minimizing skip patterns (contingency questions), and using arrows and boxes to guide the skips (Babbie, 2002; De Vaus, 1986). For self-administered questionnaires, the format is especially important because it influences the rate of return and completion as well as data cleanliness (Fowler, 1995).

We present three main topics related to format and layout of the survey. They are proper use of white space, right-side coding, and a standardized look and feel for consistency.

Following are two examples. Take a pencil and try to circle “1” in answer to question D20.A, then circle “3” for D20.B in Figure 4.1. Your



D20. Since you started this job, how often have you had trouble:

| | | | |
|---|---|---|---|
| A. Getting along with your supervisor............ | Never | 0 | SUPER _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |
| B. Getting along with your co-workers......... | Never | 0 | COWRK _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |
| C. Getting along with your spouse | Never | 0 | SPOUSE _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |
| D. Getting along with your children | Never | 0 | KIDS _ _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |

**Fig. 4.1** Not a recommended format (not enough white space)

circle should completely surround the answer without touching any other answers.

Now do the same in Figure 4.2. Try to circle “1” in answer to question D20.A, then circle “3” for D20.B.

Circling one response should be easier and faster in Figure 4.2 because of the amount of white space used (Sudman & Bradburn, 1983). Common sense reveals that the interviewer will be able to progress more quickly and accurately with this layout, and the data entry person will find it easier to detect the circled answer and transfer that to the entry screen. This example only used two questions to demonstrate the speed and accuracy advantages gained by appropriate white space. Multiply this time savings and accuracy by the number of questions asked, and you now know how much time and money will be saved by giving interviewers “room” to record the respondents’ answers. Our basic advice is to use enough white space to clearly define questions from answers and for interviewers to be able to circle a response quickly.

Our next recommendation is for right-side coding. Right-side coding refers to having all responses to questions located on the right side of the page. However, our rationale, unlike that of Dillman (2007), goes beyond respondent ease to address the ease and accuracy of data entry after data collection. Please get out your trusty pencil again and this time circle “0” for A, “3” for B, “0” for C, and “2” for D. in Figure 4.3

| | | | |
|---|---|---|---|
| D20. Since you started this job, how often have you had trouble: | | | |
| A. Getting along with your supervisor............ | Never | 0 | SUPER _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |
| B. Getting along with your co-workers........... | Never | 0 | COWRK _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |
| C. Getting along with your spouse........... | Never | 0 | SPOUSE _ |
| | Just a few times | 1 | |
| | About once a week | 2 | |
| | Almost every day | 3 | |
| | Every day | 4 | |

**Fig. 4.2** This is a recommended format (showing proper white space)

(with a few questions added to better illustrate our next point about right-side coding).

It becomes apparent that your eyes and the pencil had to move all over the page to find the appropriate answers as compared to what had to be done while looking at Figure 4.2. The data entry staff would have to do the same thing: move their eyes all over the page to find the responses. Right-side coding, which is used in Figure 4.2, makes the task of recording and entering responses more efficient.

Right-side coding has many advantages: interviewers and respondents only have to look to one place to record answers, and entry personnel can scan down the right side of the page to enter responses without having to look at the questions. One speed advantage is that a data entry person can go from the last answer on page 1 to the first answer on page 2 before the page is completely turned, and so can keypunch with the right hand while turning with the left and be halfway down the second page while finishing the turn. The major speed advantage comes from not having to search the page or filter and process the information just to find the values to enter. Interpreting and processing are minimized by restricting where the eyes need to look.

Some may argue for the visual sense of scale, as in distance from one anchor on the scale to another that occurs in the horizontal presentation

D20. [Since you started this job/when you worked at this job], how often [have/did] you [had/have] trouble:

| | | | | |
|---|---|---|---|---|
| A. Getting along with your supervisor | | | | |
| Never | Just a few times | About once a week | Almost every day | Every day |
| 0 | 1 | 2 | 3 | 4 |
| B. Getting along with your co-workers | | | | |
| Never | Just a few times | About once a week | Almost every day | Every day |
| 0 | 1 | 2 | 3 | 4 |
| C. Getting along with your spouse | | | | |
| Never | Just a few times | About once a week | Almost every day | Every day |
| 0 | 1 | 2 | 3 | 4 |
| D. Getting along with your children | | | | |
| Never | Just a few times | About once a week | Almost every day | Every day |
| 0 | 1 | 2 | 3 | 4 |

**Fig. 4.3** Not a recommended format (not right-side coded)

in Figure 4.3. As a compromise, the horizontal version of the scale can be placed at the top of each page with that response pattern, or the interviewer can hand a card to the respondent while answering the questions (Sudman & Bradburn, 1983). You can then use the right-side coding, as in Figure 4.2 to record responses. This concern is irrelevant if the interview is administered by trained interviewers because the respondent will not see the orientation of the scale unless handed a card.

In our shop, we use a standard seven-column format for nearly all survey questions. As you can see from Figure 4.4, we use table format, and each column has a purpose. (In the figure, the table grid lines are visible for demonstration purposes and would not be left visible in the field version because the presence of too many lines creates visual clutter.) The consistent format has numerous advantages. First it allows us to cut and paste together "pieces" of the interview from a library of instruments (we save any standardized instruments, once they are electronically formatted, for future use) and have them all fit together in a consistent manner. This central repository of formatted measures saves time when collating a complete survey. The table format with seven columns makes it easy to cut and paste smaller pieces such as the actual wording of the question into other applications, such as a data management database. One can easily highlight an entire column to make font or color changes throughout the document. We usually break the table after each section so we can make our changes section by section. The two columns labeled as blank are for flexibility allowing for questions of different shape (i.e., a four-response grid pattern).

Another survey-formatting tip for variable names is simple but handy. Use the survey as a form of codebook and typically include the variable

| Question Number | Question | Responses | Blank | Numeric codes | Blank | Variable Names |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| K3 | Please tell us if you: own your home, pay rent or do you have some other arrangement? | | | | | K3 |
| | | Own home or buying it | | 1 | | |
| | | Paying rent | | 2 | | |
| | | Other (SPECIFY) ________________ | | 3 | | K3S |
| | | ________________ | | | | |

**Fig. 4.4** Recommended format, explaining what each column is used for

names for each field down the right side of the page near the responses. If the interview is self-administered, this would be undesirable; the variable names would only confuse the subject. Making a second copy and removing all the variable names would take a lot of work. One advantage of the seven-column table format is that all the variable names are in the seventh column; so we simply select the entire column and change the font color to white. When printed on white paper, the variable names are then invisible to the subject, and just a few clicks away from being visible for us to print a codebook. Hence, only one electronic copy of the survey needs to be maintained. One other tip before we refer you to Chapter 5 for more on codebook creation, we find it very efficient to store the numeric codes and their meaning that we create for "other specify" or short open text type responses directly in the codebook version of the survey. We place these codes immediately following the question to which they apply.

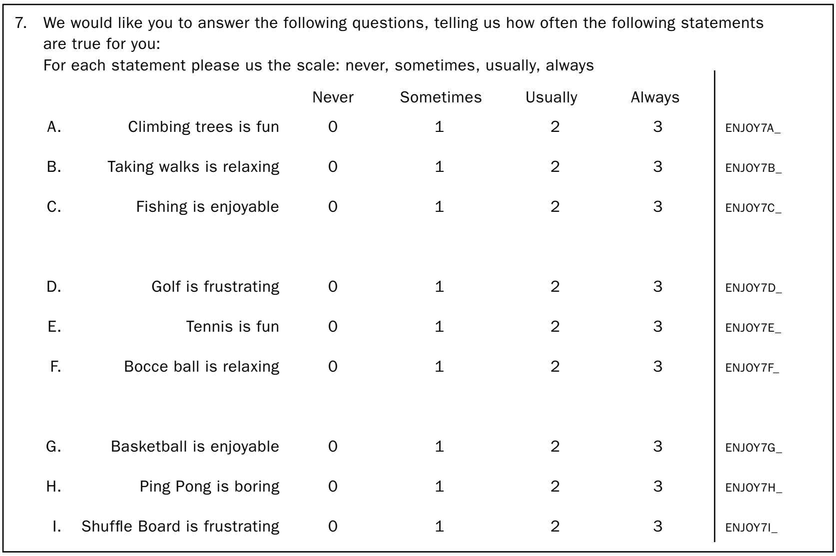
## Preparation for Data Entry

Although most of the literature focuses on formatting data collection tools to avoid errors associated with respondent answers (cf. Dillman, 2007), the issue of clean data is much broader than just the issue of data collection errors. Two other critical types of errors are transfer errors and data entry errors. Transfer error occurs when moving data from one place to another. An example is a data entry person entering the same answer twice; this usually happens because the response options are too close to each other so the answer to question 4 gets entered for question 4 and question 5. Data entry error occurs as a keystroke error, such as entering an 11 instead of a 1.

Transfer errors can be avoided by paying attention to instrument formatting. Instrument formatting alone does not prevent data transfer errors. However, these errors can be minimized in many ways by early planning. All the techniques discussed earlier cut down on transfer error. If you add white space to group responses in clusters of three, the error of duplicate entering can be greatly reduced. Data entry personnel can easily remember clusters of three numbers, so by simply adding extra white space after every third question, the data entry person will fall into a pattern of reading the three choices, remembering them, entering them, and then reading the next three. We have noted that data entry personnel were more accurate when entering 112 121 131 123 than when entering 112121131123. Try it for yourself, place your right hand on the keypad and see which is easier for you to enter. Figure 4.5 presents some hypothetical questions showing the preferred format.

## Computerized Entry

Proper paper survey format can reduce data transfer and keystroke errors, but format can only go so far. Filtering the data while entering it is a second method for finding and solving data transfer and data entry errors. Always check for syntactic and semantic errors with the data entry system (Van Bemmel & Musen, 1997). In short, if the data have a structure, such as a date, ZIP code, or time, for example, impose that structure on the data entry person by providing only the right "shape" field for entry of that value. This helps in many ways, not the least of which is giving the data entry person feedback that he or she is in the correct location on the entry screens. This can be referred to as setting range restrictions. Most survey questions have a predetermined minimum and maximum value. Setting these as "tight" as possible reduces keystroke errors. For example, if the data entry person is on a question with responses 1 to 4 with missing codes 96 to 99, make those the only allowable entries in that field. If the data entry person accidentally tries to enter an 11



7. We would like you to answer the following questions, telling us how often the following statements are true for you:
For each statement please us the scale: never, sometimes, usually, always

| | | Never | Sometimes | Usually | Always | |
|---|---|---|---|---|---|---|
| A. | Climbing trees is fun | 0 | 1 | 2 | 3 | ENJOY7A_ |
| B. | Taking walks is relaxing | 0 | 1 | 2 | 3 | ENJOY7B_ |
| C. | Fishing is enjoyable | 0 | 1 | 2 | 3 | ENJOY7C_ |
| D. | Golf is frustrating | 0 | 1 | 2 | 3 | ENJOY7D_ |
| E. | Tennis is fun | 0 | 1 | 2 | 3 | ENJOY7E_ |
| F. | Bocce ball is relaxing | 0 | 1 | 2 | 3 | ENJOY7F_ |
| G. | Basketball is enjoyable | 0 | 1 | 2 | 3 | ENJOY7G_ |
| H. | Ping Pong is boring | 0 | 1 | 2 | 3 | ENJOY7H_ |
| I. | Shuffle Board is frustrating | 0 | 1 | 2 | 3 | ENJOY7I_ |

**Fig. 4.5** Recommended format (clustering responses in groups of three; easy to remember)

in the field (a keystroke error), the entry system will prevent that error. Be sure to inform the data entry person of the legitimate values for that field so he or she knows how to proceed. The entry screens should display the valid responses when a non-legitimate value is attempted.

With right-side coding as described earlier, with the variable names on the survey to the right of the responses, designing entry screens can be relatively simple. We design the screens to look as much like the paper instrument as possible, using the same number of questions on the same pages as in the paper version. All we put on the screen for a question is the variable name and the space for the response. If each page has 10 questions, the same 10 fields will show on the entry screen. When the data entry person comes to the bottom of a page on the paper, and enters the last field on that page, the data entry application jumps to the next page and the appropriate field on that page. The data entry person should rarely have to leave the number pad. We try to avoid making the person reach for the mouse, which would slow the process down. If a skip instruction appears on the paper survey, the screens will skip to the appropriate field anywhere in the system so the entry person does not have to think at all about where to go to record the next response they find on the paper survey. We have found these techniques to reduce transfer error and some keystroke errors.

Computerized interviews (e.g., CAPI, Computer Assisted Personal Interviews; CATI, Computer Assisted Telephone Interviews; CASI, Computer Assisted Self Interviews) have sophisticated data entry checking programs to detect errors (e.g., valid ranges, filter checks, logical checks; De Vaus, 1986). Intentional redundancies can be built into the system so that the computer can alert you that a discrepancy has occurred for a particular question. This should be saved for the most critical variables, and used sparingly, because respondents do not like to be asked the same questions repeatedly. Cross-checks can be done through comparable responses to comparable items as the survey is being conducted so that the interviewer can be alerted and resolve any discrepancies. Logic checks can also be useful to detect errors. Many times the manner in which one question is answered determines how another question many pages later must be answered. This can also be detected in computerized interviews.

We would like to briefly mention Optical Mark Recognition (OMR), also commonly referred to as "bubblescan" or "teleforms." This mode of data collection involves software-specific formatting of your instrument or even special inks on special paper. This method is ideally suited for situations where all or nearly all responses can be coded into a relatively small number of choices such as 1 to 5. A drawback can be transfer error, this time more directly from the respondent to the page. Many people have difficulty accurately filling in the little bubbles or keeping their responses on the correct line of the answer sheet. Most modern OMR software has Optical Character Recognition (OCR) that is supposed to read hand writing and convert it to electronic text. Our experience has been that we have too many fields that do not fit easily into that format, and that the OCR is not accurate enough to process the data without a lot of additional manual processing (slowing the process). This technology is very promising in the right situations, so if you have all discrete response option questions, and relatively sophisticated respondents, OMR could be an excellent option.

### Variable Names, Algorithms, and Longitudinal Studies

Variables can be named in many ways, and the next chapter will detail many naming conventions. However, in this chapter we want to mention the importance and implications of naming choices in repeated surveys. Many, if not most, survey research projects are longitudinal, which implies repeated measures. This means that many, if not all, of the same variables are asked at regular intervals. Studies such as these should use longitudinal analysis techniques (random effects and mixed models, growth curve models using hierarchical linear modeling, etc.). To perform these types of analyses, datasets should have a "long" shape, not a "wide" shape. Most naming conventions tack on a wave number at the end of each variable to indicate the wave, and the datasets are merged together. This creates one row per subject, with many columns of variables, and is often referred to as "wide." To perform the proper analyses, these variables must be reshaped as "long," in which each subject appears in multiple rows (one row for each wave of data) but then each variable occurs only once (see Figure 4.6). A skilled data manager can change the shape of the dataset, but the most efficient choice is to collect the data in long format from the beginning.

If you collect your data as "wide," you have to create new entry screens for each wave to reflect the new variable names. Adding a wave or time variable to the dataset at entry allows you to keep the same entry screens for all waves because the variable names will not change, and only the value of the wave variable changes. This is, of course, easier if the survey is identical at all waves. But it can be done even with significant changes from wave to wave by naming the "long" variables with an underscore ("_") at the end of every variable except for variables such as ID and wave. This makes it easy to reshape the data to "wide." Collecting data

Long dataset: [Recommended data structure for longitudinal or repeated measures data] (Notice two rows for each subject, one for each wave.)

| Id | wave | marital_ | gender_ | hieduc_ |
|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 12 |
| 1 | 2 | 1 | 1 | 13 |
| 2 | 1 | 1 | 2 | 14 |
| 2 | 2 | 1 | 2 | 14 |
| 3 | 1 | 1 | 2 | 11 |
| 3 | 2 | 0 | 2 | 12 |

Wide dataset: [Not the recommended data structure for longitudinal data]

| ID | Marital1 | gender1 | hieduc1 | marital2 | gender2 | hieduc2 |
|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 12 | 1 | 1 | 13 |
| 2 | 1 | 2 | 14 | 1 | 2 | 14 |
| 3 | 1 | 2 | 11 | 4 | 2 | 12 |

**Fig. 4.6** Long versus wide dataset structures

"long" saves money and time because you only have to create one version of entry screens, not versions for each wave. Also, you can create constructed variables only one time; modifying the code to adjust the names for the new wave is not necessary—simply divide them by wave later. The time saved on entry screen development can be used by the data managers to construct variables for analysis while the data are still coming in, since they are no longer busy creating the second wave of data entry screens. Then use these same constructed variables wave after wave. Now data managers have more time to clean variables and datasets since they are no longer creating multiple versions of constructed variables for each wave.

There are a few more variable naming tips concerning longitudinal studies that every field researcher should know. (We learned them the hard way.) If data will be collected "wide," you have to ensure that you do not use the same variable name twice. For example, the first author started his career on a project that was well into the fourth wave of data collection. He discovered that a previous data manager had merged time 1 and time 2, saved only the final version of this merge, and apparently discarded the original data from both previous waves of data. However, approximately four to five variables in time 2 had the same name as in time 1 (the "2" was left off the end of the variable name). The merge, therefore, wrote the time 2 variables over the time 1 variables, and time 1 was lost. This could have been prevented in two ways. First, *always* keep the unmodified raw data safe and secure (on CD, DVD, or tape, physically off the system). Second, the variable names could have been imported into a database and checked for duplicates, which we recommend for all wide datasets.

Chapter 5 of this text deals in more detail with the process and steps of data management. Therefore, in the remainder of this chapter we briefly discuss only those aspects of data management that concern longitudinal issues.

## Data Archiving

Data archiving and survey format arguably do not have much to do with each other. However, certain strategies make archiving easier and more accurate. Preparing to archive your data should start from the first day of the survey process. Budget enough data management time in the grant-writing phase to track and record metadata (information about the data) as the project proceeds. Allow for at least 10% more data management time than you think necessary to complete the aim-related tasks. For example, we have not yet been part of a longitudinal study that did not change the number of questions in a scale, or worse, change the wording of items. Despite the scientific need to be consistent across time, with time comes information that shows a particular question is not working or a question is not understood by the subjects. An audit trail of all changes is a must for research, as is sharing the data through an archive. The project may have personnel changes over time, so it is important to set this type activity up as a process that is easily transferred from staff member to staff member. This can be done with a data management database that records any changes to an item, the date the change was effective, and the ID that started with the new item. Perhaps an easier way would be to maintain, in the codebook copy of the survey, all items, original to new, with indications of when an item was no longer in use, the first and last ID to use that item, and the same for the changed version of that item. Date is always important information to record about changes and what IDs were affected; first and last IDs usually are enough to mark where the change occurred (if assigned sequentially). Audit trails also are necessary for any changes to constructed variable algorithms, usually handled by dataset versioning. If any judgment is used to interpret the value of a vague response or the answers to a vague question, those decisions should be maintained in an audit trail or at least in the comments of the algorithm used to create the constructed variable.

Simply stated, comments cannot be detailed enough. Yet, as someone who has been doing this for 17 years, the first author has seen many project staff come and go across the life of projects, and even a simple line of code such as "If ID = 87345 then numfam = 3;", which seems clear, needs comments. A secondary analyst (a person using data after you have archived it) will want to know why you changed the value of numfam for that ID. At least add a comment ("*called id 87345 when numfam seemed odd at 33, subject confirmed only 3 family members, 6_17_2002, pmd *;"). Notice we included the why, when, and who made the change. If "pmd" left that project in August 2002, by October no one on the project would have any idea why 87345 was changed to a 3 if the comments were omitted. Imagine how hard it would be to re-create this type of metadata in 2005, when the project is over and someone decides to share the data with an archive—or even a colleague. In sum, begin preparation for data archiving at the beginning of the project and comment, track, and document everything, especially changes to the survey, its administration, and its questions.

## Conclusion

Most people think that the end result of a survey is obtaining and recording the data. But really, the end result is knowledge dissemination through papers and presentations and a shared dataset for archives. (Authors' aside: Data should be shared; it is the right thing to do. In addition to all the scientific advantages for the greater good, sharing data advances the ultimate goal of all research: knowledge development.) This requires a clean, clear, understandable, accessible, and usable dataset. This can be accomplished only by conceptualizing data management as a continuous process that begins with proposal writing and ends only after archiving. If all phases are kept in mind from inception, you can save money, time, and aggravation.

---

### POINTS TO REMEMBER

- For clean data, attend to the entire continuum of data management
- Adapting instruments
    - Modify measures to focus on aims.
    - Multiple subscales can lead off topic.
    - Every measure must be sensitive to the population under study.
    - To modify a standardized scale, take only what is necessary.
    - Cross-check your items against the original items.
    - Use entire subsections when possible.
    - Modify by adding items rather than changing original items.
- Data collection venues
    - A paper version of the survey is needed as a backup.
    - The more interviews, the more cost-effective is computerization.
    - Computerization combines gathering, editing, and entry into one step.
- Formatting the paper survey
    - Appropriate use of white space.
    - Right-side coding.

  - A standard look and feel for consistency.
- Preparation for data entry
  - Proper survey format reduces transfer error and data entry error.
- Computerized entry
  - Computerized entry cleans data by filtering invalid responses.
  - Redundancies and logic checks can also detect errors.
  - Optical Mark Recognition for data entry.
- Variable names, algorithms, and longitudinal studies
  - “Long” shape datasets save time and money.
- Data archiving
  - Start preparing data for archive at the beginning of the project.
  - Documentation is key.
  - Create an audit trail of any and all changes.

## REFERENCES

Babbie, E. (2002). *The basics of social research* (2nd ed.). Belmont, CA: Wadsworth Group.

DeMaio T. J. (Ed.). (1983). *Approaches to developing questionnaires: Statistical policy*. United States. Federal Committee on Statistical Methodology. Subcommittee on Questionnaire Design. Statistical Policy Office, Washington DC.

De Vaus, D. A. (1986). *Surveys in social research*. Crowes Nest, New South Wales, Australia: George Allen & Unwin Publishers Ltd.

Dillman, D. A. (2007). *Mail and internet surveys: The tailored design method* (2nd ed.). New York: John Wiley & Sons.

Fowler, F. J. Jr. (1995). *Improving survey questions: Vol. 38. Design and evaluation*. Thousand Oaks, CA: Sage Publications.

Groves, R. M. (1989). The interviewer as a source of survey measurement error. In *Survey errors and survey costs* New York: John Wiley and Sons.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Orr, K. (1998). Data quality and systems. C*ommunications of the ACH, 2*(41), 66–71.

Robins, L. N., & Helzer, J. E. (1994). The half-life of a structured interview: The NIMH Diagnostic Interview Schedule (DIS). Special issue: A ten-year retrospective on the NIMH Epidemiologic Catchment Area (ECA) program. *International Journal of Methods in Psychiatric Research, 4*(2), 95–102.

Stiffman, A. R., Striley, C. W., Brown, E., Limb, G., & Ostmann, E. (2003). American Indian youth: Southwestern urban and reservation youth's need for services and whom they turn to for help. *Journal of Child and Family Studies, 12*, 319–333.

Sudman, S., & Bradburn, N. M. (1983). *Asking questions* (p. 208). San Francisco, CA: Jossey-Bass Limited.

Van Bemmel, J. H., & Musen, M. A. (Eds.). (1997). *Handbook of medical informatics*. Heidelberg, Germany: Springer-Verlag.

Whitehead, A. N. (1911). *Introduction to Mathematics* Chapter 5. New York: Henry Holt.

Yu, M., & Stiffman, A. R. (2007). Culture and environment as predictors of alcohol abuse/dependence symptoms in American Indian youth. *Addictive Behaviors, 32*, 2253–2259.

# 5

# Data Preparation and Data Standards: The Devil Is in the Details

*Catherine M. Smith, Carolyn Breda, Tonya Simmons, Ana Regina Vides de Andrade, and Leonard Bickman*

> Form and function are a unity, two sides of one coin.
> In order to enhance function, appropriate form must exist or be created.
> —Ida P. Rolf (1996–1979), U.S. biochemist, physical therapist

You have an important and interesting research idea, a site to implement your research, and maybe even staff to help you. Your detailed proposal was approved by the institutional review board with flying colors. You are ready to go—right? Whether you are working alone or with others on the data collection, data entry, and/or data analysis for your important project, a few more details must be put (and kept) in order for the duration. Every research project that involves multiple staff requires a formalized communication system for talking about data—tracking and triggering research participants' participation (especially in cases of repeated waves of data collection), naming variables, entering data, and keeping track of all this information for the long run. And even if you are the only one who will be dealing with the data (e.g., a well-contained analysis of a well-documented secondary dataset), you will be surprised how useful a systematic approach, what we call *data standards*, can be—for those days a week from now, a year from now, or further into the future of your busy career when those perfectly transparent variable names that you made up in the middle of this project are no longer so perfectly transparent.

Every researcher has an approach to dealing with the nitty-gritty around data collection, entry, and analysis, and those approaches usually develop and change over the course of a research career. As with many aspects of life, much benefit can be gained by learning from processes utilized and decisions and mistakes made by those who have approached these types of tasks before. In this chapter, our goal is to provide an overview of methods we have found to be useful in the preparation and management of datasets.

The suggestions in this chapter arose from years of experience in a university research center. Our center conducts multiple projects of varying sizes at varying points of implementation at any given time. Some staff are project specific, but a core staff works across all projects, particularly on data analysis. As you can probably imagine, this requires a bit of juggling. Our projects also tend to be longitudinal (multiple waves) rather than cross-sectional; however, the data standards described here apply to both situations. We have found that factors such as number and types of respondents, and number of data points or waves influence the complexity of a project and its data. Data standards take complexities of projects into account, helping to make them manageable so that the dataset generated meets the immediate needs of the researcher and can continue to contribute to the field for years to come.

Chapter 4 considered several phases of dealing with data in the initial processes of data collection for research. In this chapter, we start with decisions that must be made and steps to be taken to get data into electronic format, including what you should consider when selecting software to use for data entry and additional details you will face when creating data entry programs. Pieces of this description will be more or less applicable to your project depending on how you intend to get your data; you may be collecting data directly from respondents by paper or Web-based surveys, or you may have a fully data-entered dataset for secondary analysis. Also discussed is what is required to check and ensure the quality of your electronic dataset and record-keeping tools to ensure those steps are done and the process runs smoothly. Finally, we describe elements we believe are critical for archival documentation of your dataset—work that will make the dataset as easy to use in the long run as it is when you are in the thick of the project and everything is fresh in your memory.

Let's turn to some of those data standards we believe can be especially useful to you as you make the critical decisions about data-related details. Overarching these details are two main themes: "be systematic" and "document, document, document."

## Getting Data into an Electronic Format: Considerations When Selecting Software

### *Budget Issues*

Squeezing every penny of value out of research funding is an important skill. Will you buy a canned program for data entry, pay a consultant to create one, or build one in-house? Can software you already use (e.g., Excel or Access) meet your needs? How many users require access to the system simultaneously? Will the system work over a server or on individual computers? All these factors play into cost. Remember to consider the personnel time required to get these systems up and running for your project team—either in the creation of the data entry file or learning how to use one created for you.

You may have a project that is so large or complex that it requires professional or outside help to set up the data entry program. Be sure that you think through and communicate with this professional about how the data get into the system and what the dataset should look like when it comes out of the system. As a researcher, you typically want to avoid multiple answers in one field (e.g., a list of instruments completed at a wave), but this is not an unusual setup for people with database experience but no research experience. In our experience, the data often are much easier to use if only one piece of information is included in a field. Be sure you have a clear plan and strong communication between the research team and the person(s) creating your system.

We conduct enough research that we generally create data entry programs in-house. To save time, money, and staff resources, we often start with pieces of a previous database when building something for a new project. This, and the fact that multiple people are setting up data entry systems, requires data standards that support consistency. We have formalized many of our rules about how we want variables named and data structured; pieces of this formalized system are what we share in this chapter.

### *Who Will Perform Data Entry? What Should the Data Entry Interface Look Like?*

Depending on the size of your project, you may have one graduate research assistant entering data, or you may be entering data yourself. On a larger project, several staff familiar with the project may perform this work, or you might hire workers from a temporary staffing service. Sometimes people write off data entry as a low-level skill and do not pay much attention (or money) to the people they hire for this work. Do not be fooled by this misperception; the people who perform data entry create the information you need to answer all the research questions you raised for this project. As data complexity and/or number of staff involved in data entry increases, clarity and robustness in your data entry program become even more important. For example, you may want to structure the data entry form to look like the questionnaire used to collect the data; however, this is not easily accomplished in a straightforward spreadsheet program. This usually has the added benefit of making an accidental slip from one participant's record into another record in the database more difficult. More complex data increases the need for simplicity in layouts for directly entered data as well; if you have several questions in your Web-based survey, respondents are likely to provide more complete data if the interface is clearly structured and limits the types of mistakes that can be made (e.g., not letting them leave a form until all items are completed). In one of our current projects, we are working with an organization whose staff are entering the data into a Web-based file. Data entry by individuals outside your direct purview raises special challenges; the loss of control over the process in one area (data entry) must be compensated in other ways if possible (e.g., simply structured questionnaires with data entry screens that show the actual questionnaire while people enter the data).

### *Considerations When Creating a Data Entry Program*

As a general rule, the less familiar the staff are with research or your project, the more safeguards you should build into your data entry database. For example, a common mistake is entering part of the data from one case in the row (or record) of another case, something that is entirely too easy in a simple spreadsheet. Another error is entering out-of-range values (e.g., valid response options are 1 and 2 and the data entry person tries to enter 3). Database programs such as Access or Excel can be set up to catch these mistakes before the data entry staff can continue data entry (or, at a minimum, so that such mistakes are easily identified after all data are entered). For complicated projects, such as those involving multiple respondent types filling out parallel forms, colored backgrounds on the screens that data entry staff see may be helpful. In fact, we use color to distinguish between the parallel versions of a questionnaire (e.g., youth and parent versions). We use the same screen color for data entry as on the hard copy form to ensure that people enter the data where they belong.

### *Will You Require Double-Entry Verification?*

Double-entry verification is a process in which all data are entered twice. The two datasets are compared variable by variable and case by case. Discrepancies are highlighted, and corrections are made in both data files—the first entry and the second (verification). When possible, real-time comparison as the data are being entered the second time is useful (we use Visual Basic behind our Access forms). That way the computer essentially says “wait; what you just entered for this variable does not agree with what you entered for this variable the first time you entered this case.” Reviewing the value with the protocol at hand rather than going back to the files and extracting it saves time. You will not need double-entry verification in situations in which participants directly respond into the database, as is the case with Web-based data collection. But double-entry verification, in cases for which data are being entered from paper protocols, can go a long way toward helping ensure data integrity.

### *Variable Naming*

As previously mentioned, we have several different people working together on datasets. We use some instruments fairly often across projects (e.g., measures of mental health symptoms) and others specifically generated to meet the needs of one project, with no expectation that they will be used again. Working together and adding new people to our teams over time has led us to formalize a system for variable naming to maximize efficiency within and across projects. In general, we still use eight-character variable names because some software programs still cannot use longer names, and others allow longer names but have been known to act erratically when they are used. We also have found that the odds of word processing mistakes when writing code to perform analyses increase with the length of the variable name. Our variable naming standards are shown in Table 5.1.

Consistent use of these standards allows users to tell quite a lot about the data stored in a particular variable just by looking at its name: the respondent, item content, item number, and any special feature of the item. In general, the standards shown in Table 5.1 are used in situations with a relatively large number of instruments being administered but

**Table 5.1** Variable Naming Standards

| Character(s) | Use |
|---|---|
| 1 | Respondent. For example, A for adult, Y for youth, C for counselor, or T for teacher. |
| 2 | Wave. This can be a number, a letter, or a marker. We often use “D” in longitudinal studies to indicate a specific date is in the dataset that must be considered when using the other data collected for that time point. This is especially useful in cases in which the actual amount of time (e.g., days, weeks) between administration for respondents varies and may affect an outcome of interest (e.g., time in treatment between data collection points). |
| 3–5 | Instrument code (e.g., “SFS” denotes the Symptoms and Functioning Severity Scale). |
| 6–7 | Question number *within a particular instrument rather than in the overall questionnaire or interview protocol that you have assembled*. Be sure to use leading zeroes so that your variables sort correctly when you look at them in another situation (e.g., “01” for the first question if you have 10 or more items). |
| 8 | We reserve the eighth character for special signifiers. Some that we use with regularity are “S” for a “specify” field (e.g., if ethnicity above was answered “other,” please specify) and “R” to identify an item that has been reverse coded for scoring purposes (the actual answers to the question are retained in the original variable name, e.g., AWSFS07, and the reversed answers are stored in a variable with an “R” appended to the name, in this example AWSFS07R). |

each being relatively short in length. When instruments have more than 99 items the standards are modified, with one less character used to identify the instrument (e.g., 3–4 instead of 3–5) and one more character used to identify the question number (e.g., 5–7 instead of 5–6).

To facilitate project and data management, we keep a central database that includes the instrument names and their acronyms (e.g., Symptoms and Functioning Severity Scale [SFS]) for measures that have either been used more than once or those we anticipate may become commonly used in the future. We also track variable names, variable labels, and value labels here so that we always use the same ones for a particular instrument. The database also includes the list of respondent indicators (the first character in a variable name) defined to date. When a new project starts, the researcher can access this database, borrow what is common to his or her project, and identify what remains "open" (unique) to them, suggesting new or different conventions if needed.

Our standards also include a review process in which core staff confer with researchers about plans for their data (e.g., variable names) to support them in using the conventions and ensure that divergences from the standards are appropriate. This helps limit the occurrence of foreseeable, avoidable problems with data structure. Many programs (e.g., analysis programs in SAS or SPSS) can be reused with minor tweaking when common conventions across projects are used. This creates a capacity for saving resources of staffing and time.

Of course, some projects are small and use measures not likely to be integrated into a later project. In these cases, more flexibility does not pose any serious threat to usually scarce project or center resources. Nevertheless, we have found comfort in knowing that if a variable name starts with Y, we know we are dealing with data from youth respondents. Life is complicated enough; such standards go a long way to simplify and reduce errors.

### *Formats and Other Fun Things to Think About*

When creating a set of variable names, the goal is to have a one-to-one relation between those names and the questions you ask participants. We have a rule that we never enter answers to more than one question into a single variable name. For example, consider parallel versions (A and B) of a form, each with three questions about an aspect of the person's experience; some respondents respond to version A and others to version B. At first, you may be inclined to use the same variable names for both versions and to remember that distinguishing between form version A and form version B is important. Giving these variables unique names across both form versions rather than just within a form version is safer. That way, after years pass or others pick up your now secondary dataset for analysis, no one forgets that these were six different questions split between two versions of the forms and not the same three questions asked on each.

We enter many of our variables as numbers and only a few as text or dates. This often makes for smaller file sizes, which can be helpful as the dataset grows. In addition, some data entry platforms only allow a certain number of characters to be entered for a given record, so if you do not need a variable to be text, code it as a number. Dates generally can be entered either as one variable (in date format) or three (month, day, year). Both ways have advantages and disadvantages. With one variable, coding the calculations for new variables that denote time intervals—for example, differences between two dates, such as date of interview and respondent's date of birth—typically is easier. On the other hand, other data manipulations (e.g., selecting records for which the month is six or greater) often are easier if the date is three variables. Be sure to think about your plans for the data (what new variables you may want to create, what analyses you have in mind), conduct some tests, or talk to people who have worked with the database and analysis platforms you are using to ensure that dates correctly transfer and can be used as you need.

When you select the program, evaluate whether it can prevent duplicate records from being entered or duplicate values from being entered into a particular variable. This is very useful if you have assigned unique participant IDs so you know only one record exists for a given ID (assuming the file is a single record per case file and not time series). Then, if data have already been entered for participant 1001 and the person performing data entry tries to enter another case for 1001, the case will be rejected and the data can be checked to ensure that the correct answers are entered for case 1001.

### *Creating a Codebook*

After you have made these decisions and created a clear and coherent variable naming and value labeling system, be sure to document it all for short- and long-term use. Document the variable names and values (codes) for each measure's response options (e.g., is "not at all" coded as 0 or 1?). Also identify what missing values are used (e.g., −9 for missing, −8 for "do not know"). Our codebooks start with an electronic copy of the instrument. We enter the variable names and values (often in a small font and always in a contrasting color) into the document. The values often are already there because they were provided as cues to the respondent when answering questions (e.g., when response options are shown as an ordinal scale: 1 = "strongly disagree" to 5 = "strongly agree").

This approach lets you search the document to find the precise variable name you are looking for. Storing this electronic file in close proximity to the data file(s) is a good idea. Doing so increases the likelihood that people wanting to analyze your data (even you!) will look at the codebook rather than try to guess what is what. Printing a hard copy of codebooks and binding them provides a reference for staff when electronic access to codebooks is limited.

### *Secondary Data Analysis*

If you are performing secondary data analysis, you do not have control over many data-related decisions. However, such situations make getting the type of information we suggest even more critical. What is that mystery variable that shows up in every other wave of the dataset? How is gender coded? What values are missing values, and have they been defined as such? If you have a good codebook, you are in much better shape to perform rational analyses and to develop findings that are clear and reliable.

## Cleaning Your Data

Have you ever been asked to respond to a questionnaire that used a 5-point scale and you really wanted to answer with 3½? We have, and some of your respondents will have similar inclinations. How do you know if a blank in your data entry file represents a question your respondent did not answer or an oversight made by the data entry person?

We have developed some data cleaning standards to apply before and after data are entered; both of these phases must occur before you begin real analysis of the data. As discussed in Chapter 4, cleaning before data entry often involves detailed reviews of forms completed by respondents following a detailed protocol for how to address or correct responses so that when data entry occurs, forms can be entered with as little interruption as possible. A cleaning process occurs after data entry as well—whether the data are entered in-house or already are electronically available through a Web-based survey.

### *Preanalysis Data Checklist*

We have identified several standards we believe are essential to ensure that your data are of the best quality possible before analysis begins. In our experience, these standards apply whether you have one or many datasets in your project. Depending on the size of your project and your organizational structure, you may want to have a process in which someone is responsible for determining whether the protocol has been correctly applied.

To standardize this phase of our review process and track our progress toward it, we complete a preanalysis data checklist for each dataset we construct. Table 5.2 shows the form we use, with some modifications. For example, it does not show the space at the top of the form to record the project's name and the name and location of the electronic dataset. This version includes some added comments to clarify the process. You can keep these forms in electronic format or work with paper versions. The most important thing is to have a standard procedure and supporting documentation that enable you to check the steps important for generating sound datasets.

### *Cleaning before Data Entry*

Completed forms (e.g., surveys) are reviewed according to a detailed protocol for checking aberrant situations, which include the following:

- Missing data
- Multiple responses
- Missed skip patterns
- Response patterns

Missing data are never fun but are a fact of a researcher's life. Rather than leave fields blank in data entry, we find it useful to distinguish the

**Table 5.2** The Preanalysis Data Checklist

| Initial and Date When Complete | Task |
|---|---|
| 1 | Create initial codebook. Have variable names reviewed by an appropriate signor before finalization.<br>Include codebook location here. |
| 2 | Create data entry system. Consult relevant center staff; be sure data entry person and date are included as variables; be sure to prevent (when possible) entry of out-of-range values and duplicate values (e.g., ID numbers). |
| 3 | Cleaning protocols. One answer per question (when appropriate); all skip patterns respected; all dates imputed where necessary. Use a red pen to record any changes to original data. Initial and date hard copy with the word “clean” or a checkmark. |
| 4 | Data entry. Initial and date the hard copy of the form with the word “enter” or “de” (data entry). |
| 5 | Data entry original vs. data entry verify. Are all the IDs listed in each table identical to IDs listed in every other table in that database (where appropriate)? Make corrections as necessary. |
| 6 | Data entry vs. tracking comparison. Are all cases and only cases in the data file (where the “real” data of interest are) expected based on the cases/information in the tracking database (including contact info, schedules for collecting the data, etc.)? Do all the cases that have been data entered have signed consent forms on file (as indicated in tracking database)? Do ID combinations (e.g., child/teacher/classroom/school) in the dataset match those in your tracking database? |
| 7 | Additional data entry and re-check (to rectify problems found in previous step). Document all corrections. Segregate any cases that need to be separated from final “good” dataset (and cannot be used for analysis) and put into a “bad cases” dataset. |
| 8 | Evaluate the data in table format. Do missing value patterns look correct? Do any strange patterns or blank cells need to be researched? Any out-of-range values? |

(*continued*)

**Table 5.2** Continued

| | |
|---|---|
| 9 | Additional data entry and re-check (to rectify problems found in previous step). Document all corrections. |
| 10 | Data "SAS-ified." Import data from all sources into SAS with proper labels and codes. Review univariate statistics for abnormalities. We use SAS for our datasets; however, the same applies to any other package used (e.g., SPSS). |
| 11 | Project manager reviews SAS frequencies. Are all variables from your database represented in the SAS file? Are the variable and value labels correct? Do the frequencies match your expectations? Are there out-of-range values? Are there duplicate IDs where there shouldn't be? Are all missing values marked? Document all corrections. |
| 12 | SAS data merging. SAS files are merged so that all data at a given level are brought into one file (e.g., child baseline information, repeated measurements over time). Merged file checked by lead analyst and project manager. |
| 13 | Project manager and principal investigator meet with data analysts to review the analytic plan and determine whether the collected data (size, structure, respondents, etc.) allow answers to the planned hypotheses. Do the research questions need reformulation? What is the appropriate methodology to use with the available data? |
| 14 | Plan how to organize reports of results. Typically, analysts first produce "result scraps," which are memos with charts, tables, and results language. Later these scraps can be assembled into results sections for articles and reports. Each project should have a tree of folders for results organized into major categories (e.g., descriptives, psychometrics, and the project's specific aims). |

various reasons as to why some items are missing or blank. In fact, some interesting analyses can be conducted on missing data, such as whether those with complete data differ from those with missing data in ways that may matter to your main research questions.

We use typical categories to describe missing data and assign the same set of codes across projects to define them (e.g., −6 = refused;

−7 = not applicable [such as when males can skip a question about the onset of menarche]; −8 = respondent indicated "don't know"; −9 = otherwise inappropriately missing). We use negative values to denote missing data because these values are easy to spot on frequency reports. Be sure they are labeled as missing and defined as missing as well so that their values do not get combined with "real" (valid) values in calculations. A downside of using negative values for missing data comes with date variables; storing a negative date electronically generally is not possible. Using 7/7/7777 or another implausible notation for a missing date (without negative sign) may work. The questionnaire form often already has the missing values and their text equivalents printed, but sometimes only the text is printed. In this case, cleaning the protocol requires the reviewer to write the appropriate code on the form so that staff can sail along when the time comes to enter the data.

Occasionally a respondent circles or marks more than one response option (answer), in essence giving multiple responses to a single question. Because data are valuable you do not want to define an item as missing if recourse is possible, which often it is when using ordinal scales (e.g., a question with response options that range from 1 [strongly disagree] to 5 [strongly agree]). When two answers are recorded that are contiguous (next to one another in the scale) and do not conflict with one another (i.e., not a situation in which one answer is negative and the other is positive), we flip a coin. For example, someone circles the 4 and 5, indicating he or she "agrees" or "strongly agrees" with the item. These values are contiguous and do not conflict. Keep a coin handy and flip it. If "heads" appears, code the answer with the higher value; if "tails" appears, code the lower value. If a respondent marks a 2 and a 4, sadly, you are left with missing data because salvaging reasonably reliable data is not possible. In this situation, "inappropriately missing" would be the best description to make and " −9" would be the appropriate value to mark on the form.

Skip patterns result when the answer to one question determines whether the respondent continues to the next question or skips to another question further down the page. This often is called branching. The answers to a set of such items should relate in a predicable way. For example, if a lead (trunk) question asks gender and a follow-up item (branch) asks for date of menarche onset, you would expect that all those who mark male on the trunk would skip the branch item. But life is not that clear-cut. Be sure to review and clean the data for any inconsistencies. Sometimes the set of questions clearly determines which item should be changed, the trunk or the branch. For example, a safe choice would be to change the male response to female because the person gives a date of menarche to the next question, and you also may have other information that tells you this is a female respondent. Or, you may determine that the answer to the branch should have been skipped after all and, therefore, should be coded as not applicable (e.g., −7). However, the right answer sometimes is not clear. Use whatever information you may have available to help make the best decision. When in doubt, we recommend coding any ambiguous data as missing (e.g., −9).

Response patterns sometimes are easy to see on the completed form if you know to look for them. Youth respondents sometimes find it "interesting" to circle answers to create a picture on the form; Christmas trees and zigzag patterns are popular. Or, they mark the same answer blindly for the entire questionnaire. Frankly, little can be done to salvage these data. You can try to re-administer a questionnaire or an interview, but this typically is logistically impossible and resource intensive and raises issues surrounding respondent resistance, with low odds that reliable data from any subsequent effort would be forthcoming. Get that red pen out and brace yourself for a series of −9 values or, perhaps better, −6 values if the results show this was the respondent's way of refusing to participate.

Decisions regarding missing data must be made when faced with the situation. Any corrections or notations you make on a form to reflect these decisions should be written next to the question, directly on the paper protocol, preferably in a high-contrast ink (e.g., red). For future reference, record the reviewer's initials and a justification for the coding decision made next to the question. No item should be blank; everything has some code that later gets entered, either a valid value or an indicator of a missing value. We also suggest that when this data entry cleaning is completed, you record some information on the cover of the protocol or form so you do not have to flip through each page to see whether the data have been cleaned or where you left off. We note which step was completed, the date the step was completed, and the initials of the person doing the cleaning so particular protocols can be reviewed if, for example, you later determine that two reviewers had different approaches to coding a particular question. Also keep track of the decisions you make; we recommend that you create a "decision rules" document that, along with codebooks, becomes part of the materials associated with a project. Such documentation will bring you great relief when the day comes (and it will) that you cannot quite recall why you did what you did when you did it!

### *Cleaning after Data Entry*

Before you jump headfirst into analyzing the data, a few more cleaning steps are strongly advised. First, run univariate statistics—at minimum, frequencies, minimum and maximum values, and means—for each variable. This will tell you fairly easily whether you have any common problems that arise in the data. A form to document the status of this review process can be helpful. Looking at frequencies and descriptive statistics:

1. Do the frequencies match your expectations? "Yes "or "No." If "No," record the missing categories or missing range.
2. Are there unexpected duplicate IDs? "Yes" or "No." If "Yes," identify the duplicate IDs.
3. Are missing values defined correctly (i.e., as missing and not as valid values)? "Yes" or "No." If "No," record how missing values should be identified.
4. Are there out-of-range values? "Yes" or "No." If "Yes," record the out-of-range values.

*Out-of-Range Values.* The variable is only supposed to range from 1 to 5, but you see from looking at the maximum value (or tail end of a frequency distribution) that you have a 6. If possible, you will have set up your data entry program to prevent entry of out-of-range values. However, errors are sometimes made, and out-of-range values do sometimes get through. As a result of this inspection, you can correct the data in the file before proceeding.

*Strange Distributions.* How can it be that 90% of the sample answered "strongly disagree" to a question about liking chocolate? On further review, you find that the programming that defines and describes this item (its data definition) was reversed; 10% do not like chocolate. Or, you see a conspicuously small number of cases defined as missing on items for which you would expect a larger number, such as an item about menarche. On review, you realize that data that should have been entered and defined as missing were not, so they are included among the valid values. Or, you know you interviewed 200 youths, but the univariate statistics show only 100. On review, you learn that a stack of forms in someone's drawer never got entered.

Second, print a "dictionary" of the file that, at minimum, shows the variable and value labels associated with each variable. (This function is available in most software packages such as SAS and SPSS.) You can create a form (as we have) for recording the status of this review.

1. Are all variables from your database represented in the SAS file? "Yes" or "No." If "No," record the missing variables in red ink.
2. Are other variables in the printout not included in the database (e.g., wrong variables out of place)? "Yes" or "No." If "Yes," record the additional out-of-place variables in the printout.
3. Are the variable and value labels correct and not missing? "Yes" or "No." If "No," specify or mark the ones that are not correct and write down the correct names or labels.

With this review process and documentation, you will be able to catch common problems before moving on to analysis. Be sure to run through this review process again after corrections are made to be sure that all the problems you noted on the documentation form were corrected.

Does this sound like a lot of work? It is, but it is manageable and, frankly, essential if you want to ensure the integrity of your datasets. Cleaning data on the front and back ends, variable by variable, gives you a chance to catch errors that you simply will not see once you delve more deeply into more sophisticated statistical analyses. Do as much cleaning as you can as early as you can. It is worth the investment!

### *Tracking All Project-Related Data: The Master Spreadsheet*

Some large projects may have different databases for different respondents or different waves of data collection. To facilitate management of large and complicated data structures, we developed a spreadsheet (the elements of which are shown in Table 5.3) to track the status of each database, including who has completed each step of the protocol. We store this spreadsheet on a shared computer drive; in our case the server, the same location we use to store all project-related information. The

**Table 5.3** MASTER: Data Table Checklist

| | |
|---|---|
| Common file name | The name you call the file in general, such as "Pediatrician Baseline" or "6mo Teacher." |
| Raw (access) data file name | The actual filename as it is stored on your computer; you could easily combine this with the row below and call it "file path." |
| Raw (access) file location | The path to finding your file on the computer, such as "N:\Projects\ADHD\Data\PedBackground." |
| Codebook created | The date that the codebook was created and the initials of the person who did the work. If you are working with a large team and want to show that this is in progress and not yet complete, you could mark here "*your initials* in progress." |
| Data entry completed | The date data entry was completed. You could use initials to identify the person who certifies that the data entry was completed or the person who did the data entry. This depends on dataset size and your processes, but some notation is useful. |
| Data entry vs. tracking comparison (step 6 of preanalysis checklist) | Ensuring that you have all the cases in your data entry database that you would expect based on your record of who completed what surveys or interviews is critical. For us, this involves a comparison of the IDs in the data entry database (and in multiple tables of that data entry database, if it is complicated in that way) with the IDs shown as completing that survey/interview in our participant tracking database, as discussed in the preanalysis data checklist. |
| Raw data cleaning (step 9 of preanalysis checklist) | At this point you have entered any cases that you discovered had not been entered, taken a look at the data file itself to see if any visual oddities presented themselves (e.g., blank cells) and corrected those problems. Again, note date completed and initials. |
| Expected completion of raw data cleaning | Whether this column is useful depends on the size of your team and how many projects you are trying to juggle. If you have many irons in the fire, it can be helpful if the person responsible for cleaning the data puts his or her initials and the expected date of completion here so that people who deal with the dataset afterward can plan their work. |

(*continued*)

**Table 5.3** Continued

| | |
|---|---|
| FINAL raw data on N:\ | The date the final raw dataset is completed. In our situation, the date released to the team of the analysts for what we call "SASification" because we often use SAS for this type of work. We refer to SASification as taking raw data, such as in Excel, Access, or FileMaker format, turning it into a file that can be directly accessed by the analytical software, getting the variable and value labels attached to the correct variables, and running descriptive statistics so that people familiar with the project can review them to ensure that all is well. Having the date, the data were released and the initials of the person who released the data can be handy for tracking an active workload and to help the analysts know whom to ask if they have questions. |
| Who writes SAS job | Another variable that is more useful for big teams or projects. This is good for divvying up workload. |
| DRAFT SAS | Date when the draft SAS file is completed and initials of the person who did the work. This also involves giving the appropriate project person a printout of the descriptive statistics for each variable for review. |
| Project manager approval | Date dataset has been approved by the project's responsible person (the project manager or data manager) and the initials of that person. |
| Comments | You never know when you need to record some information about special circumstances or events, or what was done and why, with a particular dataset. This is a great central location to store this information. |

sheet is updated by staff responsible for each step. The file is saved with a name such as "MASTER ProjectNameHere data table checklist." We even give guidance to users at the top of the document:

> This is the MASTER ProjectNameHere data table checklist. As you complete a task or edit a file, please date and initial the appropriate box. If changes/edits are made, please document the changes.

Table 5.3 shows the elements we review for each dataset. The pre-analysis data checklist gives more information about what needs to be done with a database. The spreadsheet described here tracks the landmarks. To be clear, the actual table we use has these items as the column rather than the row heads, and each individual dataset in the project is listed and tracked on a separate row of that table.

## Archiving Your Dataset

After completing all this hard work, you only have one more step. We promise, if you (or anyone else) ever look at these data again, you will be glad you pushed through. A big part of making your dataset useful for yourself or others at a later point is documenting what went on throughout the course of data handling. In subsequent sections we describe information that will be useful to include in your documentation. You may think of some other topics on your own and add them to your standard operating procedures for working with data files.

### *What to Include in Your Archive Documentation*

We currently store the documentation for each dataset on our shared electronic drive. We use a hierarchical folder structure to keep all related materials for a project in one general place (e.g., a project folder with a subfolder for data and a subfolder within data called "youth") and descriptive file naming to identify the specific content of a document or file. See Table 5.4 for a list of elements we have found to be important components of archive documentation.

With all these pieces of information about your dataset in one place, much of the difficulty of writing about these data is completed. Knowing that the information you need will be at hand long after it has faded from your memory is comforting.

## Conclusion

This chapter presents many details about dealing with data—devilish details that are especially relevant when more than one person handles the data or when one person handles the dataset at different points in

**Table 5.4** Components of Archive Documentation

| Element | Description and Reasoning |
|---|---|
| Description of project or study | For example, research topic, funding agency, grant number, dates of project. This information is useful when you are writing a paper based on these data 4 years after the grant ends. |
| Methodology | Primary or secondary data collection? Qualitative or quantitative data?<br>Cross-sectional or longitudinal?<br>Phone, mail, Web? |
| Types and sources of data; respondents | Where did you get the dataset? What types of information are included (a brief description can be handy)? Who were the respondents? |
| Sampling issues | Description of any sampling issues, considerations about representativeness of sample. |
| Norms for data | Are data norms available? If so, where can they be found (on computer, in journal articles)? What are they? List of references? |
| Access to data | What restrictions are there for accessing the data? Is approval necessary for access? If so, from whom? Contact information? |
| Variable naming conventions | What conventions did you use? |
| Data cleaning | What cleaning steps were taken? What decision rules were used? Locations of any relevant files. _________ . |
| Description of summary measures | If summary variables were calculated and included in the dataset, describe how they were scored and how missing data were handled (e.g., threshold for completeness of original items before computing; what imputations, if any, were used). |
| Descriptives | Values such as n, mean, min, max, standard deviation for each variable in the dataset. |
| References | Include references that should be included when writing about these data (e.g., references for any measures used). |

time. Although these are not the only issues that will arise as you manage your dataset, we hope that our review will help you avoid some of the problems that inevitably occur when conducting research.

We also acknowledge that some of the solutions we have developed for the research we do may not fit perfectly with the processes or data structures that you face. What we hope to communicate here is that being systematic in documenting processes and decisions is well worth your effort and provides a strong foundation for a successful research career.

---

## POINTS TO REMEMBER

- Use a systematic approach: data standards
- Getting data into an electronic format
  - Budget issues.
    - Choices: canned program for data entry, pay a consultant to create one, build one in-house.
    - Consider the personnel time for each system choice.
  - Who will perform data entry? What should the data entry interface look like?
    - As data complexity and/or number of staff involved in data entry increases, clarity and robustness in your data entry program becomes more important.
  - Considerations when creating a data entry program.
    - The less familiar your staff are with the research or your project, the more safeguards you should build into your data entry database.
  - Variable naming.
    - We use eight-character variable names.
    - The odds of typing mistakes increase with the length of the variable name.
    - Keep a central database that includes the instrument names, their acronyms, variable names, variable labels, and value labels.
  - Formats.
    - Never enter answers to more than one question into a single variable name.
    - Give variables unique names across parallel forms.
  - Creating a codebook.
    - Document variable names and values for each measure's response options.
    - Identify what missing values are used.
    - Store this electronic file in close proximity to the data file(s).
    - Print a hard copy of codebooks and bind them.
- Cleaning your data
  - Before data entry
    - Review completed forms for missing data, multiple responses, skip patterns, response patterns.

        - Write corrections or notations next to the question in high-contrast.
        - Record information on the cover of the protocol or form.
        - Create a "decision rules" document.
    - Cleaning after data entry
        - Run univariate statistics.
            - Do frequencies match your expectations?
            - Are there unexpected duplicate IDs?
            - Are missing values defined correctly?
            - Are there out-of-range values or strange distributions?
        - Print a "dictionary" of the file.
    - Tracking all project-related data: the master spreadsheet.
        - Store spreadsheet on a shared computer drive.
        - Sheet is updated by staff responsible for each step.
        - Save file with a name such as "MASTER ProjectNameHere data table checklist."
        - Give guidance to users at the top of the document.
- Include in your archive documentation
    - Description of project or study.
    - Methodology.
    - Types and sources of data; respondents.
    - Sampling issues.
    - Norms for data.
    - Access to data.
    - Variable naming conventions.
    - Data cleaning.
    - Description of summary measures.
    - Descriptives.
    - References.

## 6

# Cultural Sensitivity and Cultural Disparities: Ethical Dilemmas, Legal Issues, and IRB Requirements

*Arlene Rubin Stiffman*

> A principle is the expression of perfection, and as imperfect beings like us cannot practise perfection, we devise every moment limits of its compromise in practice.
> —Mohandas Karamchand Gandhi, www.brainyquote.com

When conducting a research project, the process of obtaining institutional review board (IRB) approval seems onerous enough. However, many of the ethical and cultural issues involved in research go far beyond what an IRB might consider. This chapter addresses those issues by using a study of American Indian youths to illustrate some of the conflicts, decisions, and solutions that researchers may confront in implementing human subject guidelines and regulations.

## History

It is hard to remember that IRBs are a relatively recent innovation. Before 1946 and the Nuremberg Code, there were virtually no institutionalized reviews of research ethics or protection of human subjects (Shuster, 1997). In fact, it was not until 1972 that all institutions were required to provide evidence that they subscribed to the Nuremberg Code. Only in the 1980s were a series of rules and regulations put in place to guide institutions in their review. The first major guiding document for these rules and regulations was the *Belmont Report* (1979). It was the first document to articulate the ethical principles behind the study of human subjects, which included respect for persons, confidentiality, beneficence, and justice in distributing the benefits and burdens of research (*Belmont Report,* 1979; National Institutes of Health [NIH], 1998).

The IRB is a group or committee that is given the responsibility for reviewing that institution's research projects involving human subjects. At universities, it is usually composed of staff, faculty, consultants, and nonaffiliated community members. Their primary purpose is to assure the protection, safety, and rights and welfare of the human subjects. The IRB is responsible for interpreting federal regulations concerning research protections. There may be considerable differences between institutions in the way they interpret their roles or their requirements for approval (wording, forms, and protections). In addition, some institutions have expanded the role to include education of faculty concerning ethics, assuring that researchers are qualified to do the studies proposed, and/or assuring that the institution is protected from potential lawsuits.

The entire culture and climate of research has changed dramatically in the past 15 years since review protocols and regulations have come into play. However, although IRB protocols are designed to protect human subjects, they often maintain the researchers' primary cultural bias, which is knowledge for knowledge's sake. Regardless, social service researchers in particular must consider more than just knowledge development in their research (Stiffman, Brown, Striley, Ostmann, & Chowa, 2005a; Stiffman, Freedenthal, Brown, Ostmann, & Hibbeler, 2005b). Social science research occurs within populations that have their own cultural views of the importance of knowledge, of what constitutes knowledge, and of cultural needs and priorities. In fact, in many cases these cultural perspectives may not value research. Little research addresses the need to maintain sensitivity to ethnicity and culture while preserving an ethical research program and ensuring the development of knowledge.

## Background

The ethical imperatives of confidentiality as specified in review board regulations may conflict with the broader ethical demand of not ignoring need in individuals (King & Churchill, 2000). Only recently have researchers confronted these additional pressures that lie outside the common human subjects concerns (Stiffman et al., 2005a; Stiffman et al., 2005b). A critical question for all researchers—particularly social science researchers—involves how to handle ethical issues related to particular cultures without destroying the integrity of the research (Hoagwood, Jensen, & Leshner, 1997). The IRB regulations may not fully address ethical imperatives in general or specific ethical issues relevant to particular cultural groups. Regardless of the culture of the community, ethical demands require that researchers maintain confidentiality, do no harm in the interview, protect the person from harm, and protect anyone else from being harmed. However, within a particular cultural group, an interview that may be innocuous in other cultures might be perceived as potentially harmful. For instance, some communities have a prohibition against revealing feelings and emotions that an interviewer may ask about, especially if that interview focuses on problems and risks. Furthermore, in some communally oriented cultures, protecting confidentiality may be exceptionally difficult because the community would immediately be able to identify both the interviewer and the interviewee.

Research that uses children as subjects further complicates the issue of the ethical adequacy of IRB regulations. The limited literature tends to focus on the difficulty of balancing human subjects concerns—particularly with child and adolescent research participants and their families—and the need to promote opportunities for the advancement of knowledge (Hoagwood, Jensen, & Fisher, 1996). The literature emphasizes the disparity of power between any adult researcher/interviewer and child subject. The child may potentially be unable to provide fully informed consent. This may be true even of adolescents; they may have a certain degree of family or peer pressure that interferes with their ability to make the choice on their own. Children and adolescents may be giving involuntary or nonvoluntary assent without the researcher being aware of it (Dorn, Susman, & Fletcher, 1995; Hoagwood et al., 1996; King & Churchill, 2000; Putnam, Liss, & Landsverk, 1996). Also, human subject regulations assume that the consenting guardian adult has the child's best interests at heart. This, however, may not be true, particularly in high-risk populations in which the guardian may be court appointed, another relative, or an institutional representative. Unfortunately, the history of child abuse by institutions and research programs is notorious

(Glantz, 1996), such as in the case of the Willowbrook State School conditions revealed by journalist Geraldo Rivera in 1972. Because childhood and adolescence are such critical developmental times for engendering productive adulthood, social science research on this age group must address these issues (Glantz, 1996; Koocher & Keith-Spiegel, 1990; Levine, 1995; Melton, 1992; Thompson, 1992).

Although researchers frequently are confronted with multiple cultural and ethical dilemmas, few articles provide guidelines for resolving them. The recommendations seem almost too simplistic. They often are limited to stating things such as "involve stakeholders and collaborators in research projects to avoid or resolve such issues" (Attkinson, Rosenblatt, & Hoagwood, 1996). Sometimes this stakeholder participation only means that a local IRB reviews the project or, if it is a multisite or multiagency study, several IRBs at different institutions may review a single project (Hoagwood et al., 1996). Research stakeholders often are families, and some of the literature recommends fully informing families of research decisions. The literature also states that a researcher should "strive for cultural competence and respect diversity and experiences of the subject and family" (Osher & Telesford, 1996). Nevertheless, leaving collaboration and involvement only to the development of the consent document and study design is inadequate. Simply making the language of the researcher accessible does not mean that the research is necessarily fully sensitive to conflicting or subtle ethical or cultural issues (Putney & Gruskin, 2002).

The major topic for ethics discussions in research has been child abuse protection. The Child Abuse Prevention and Treatment Act (CAPTA) obligates professionals to report suspected cases of child abuse or neglect (Kalichman, 1993; Meyers, 1992). However, the literature is full of arguments regarding whether that regulation applies to researchers or interviewers and whether certificates of confidentiality override the mandated reporting statute (Kotch, 2000; Steinberg, Pynoos, Goenjian, Sossanabadi, & Sherr, 1999). Certainly, it appears on the surface that the researchers' promise of confidentiality would directly conflict with the mandated reporting of abuse. Most researchers have solved this issue by including on the consent form clear language informing subjects of the potential necessity to breach confidentiality if they reveal child abuse (Attkinson et al., 1996). In fact, even children who are being researched *because* they are members of a protective service community may provide new disclosures, new identification of perpetrators, or additional information about their welfare that could fall under the CAPTA mandate regulations (Putnam et al., 1996). The assumption of most researchers is that reporting to authorities will protect the youth. However, such reporting may have some negative impacts. It may create emotional distress, provoke further abuse as a result of disclosure, or trigger intrusive investigations by law enforcement officials that provide more distress than solution. It may also separate the youth from parents by placement in a foster or group home situation that may be far less adequate than the original home. Little literature exists on alternative options open to researchers. One potential option is to encourage the youth or family to seek assistance and self-report any maltreatment, thereby bypassing the confidentiality regulations and removing the interviewer or researcher from having to make the report. Some literature states that doing this empowers the family to move forward in treatment and also assists the researcher (Putnam et al., 1996). However, we lack data on the impact of potential disclosure of maltreatment when done within a research setting in which confidentiality was promised (Singer, 1984). Furthermore, we have even less information about the impact on research itself. Does it reduce self-reporting of such problems? Does it reduce continuance in longitudinal studies?

The remainder of this chapter uses a research study as a case example to describe some of the issues cited. We will discuss how the issues were solved within the context of the research, whether these solutions were successful, what their impact was on the research and the community, and the research principles that we learned from our experiences.

### *Example*

The American Indian Multisector Help Inquiry (AIM-HI; a study funded by the National Institute on Drug Abuse) studied service use and drug information in two American Indian populations and confronted many of the conflicts involving IRB regulations, generic ethical imperatives, youthful subject issues, and culturally specific ethical issues. The AIM-HI study began in 2001 with a sample of 401 youths aged 14 to 19. Approximately 200 youths were from an urban area and 200 were from a reservation area. The design required that they be followed yearly until 2004 to obtain ongoing information about their service needs and use. This community is particularly appropriate to serve as an example for ethical and cultural issues because conflicts involving research, human subject guidelines, and ethics are particularly poignant among American Indians (Marshall, 2001; Stubben, 2001). American Indians have an unfortunate history of having been “cheated” by unscrupulous business and research arrangements. Promises have been made to American Indian communities to access them for research, while they received nothing in return (Doyle, 2001). Many American Indian communities are particularly wary of researchers. In fact, researchers must gain more than simple interpersonal trust. In reservation areas, research must first be vetted, reviewed, and approved by the Tribal Council. Gaining that trust involves having the researcher see the tribe’s and the tribal members’ best interests as primary while accepting the research interests as secondary (Beauvais, 1999; Norton & Manson, 1996; Weaver, 1997).

## Ethical Issues in Stakeholder Involvement

The AIM-HI research tried to address the ethical issue of involving stakeholders by initiating a multiyear planning process with the community prior to even submitting the first proposal to the NIH. The planning stage covered proposal redrafts for 4 years prior to actually obtaining funding. The principal investigator (PI) of the AIM-HI study (the author of this chapter) is not American Indian and would have had no contact or entry into the American-Indian community without the intervention of colleagues, including the co–principal investigator of this study, Dr. Ed Brown. As an American Indian who had been an active leader nationally as well as locally in the geographic area, Dr. Brown knew the needs of the tribe being studied. He approached the PI, saying that the community needed the type of research that the PI did and introduced her to the community. This allowed the respect that he had garnered to be transferred to the PI. This process involved multiple visits with the community members to discuss their needs, ideas, and the potentiality for research. The PI’s research career focus was basically reframed so that it would answer the community’s own research priority and needs. From the community point of view, this was the primary rationale and purpose for the research. And, from the researchers’ point of view, by agreement, this became the primary focus of the project. Pure knowledge development became an unspoken background issue that would never interfere or conflict with the research answering the community’s priorities or needs.

As part of the research planning process, the researchers initiated a "research implementation team" (RIT) to obtain the input of the stakeholders. This team was composed of American Indian elders, American Indian human service workers from the urban and reservation areas, tribal council members, parents, representative youth, and local university faculty. The RIT directed the types of questions they wanted the survey to address. As a whole, they were very aware of the problems in their community and anxious to obtain information that would address these problems. In many cases the information they requested was much more sensitive than what the researchers originally had conceptualized. For example, they requested the inclusion of items about child abuse, sexual orientation, and HIV risk behaviors. The original research ideas were to look solely at drug abuse and mental health issues and related services. Members of the RIT were particularly vocal in emphasizing the need to obtain information about the other problems. In some cases, they had personal agendas that developed from family members who were HIV positive, transgendered, or had a history of physical or sexual abuse. In fact, they were anxious to share these personal issues, and then generalize them to the information needs of the community as a whole. Over the course of approximately 1 year, the researchers and the RIT developed a mutually agreed on research plan and associated instruments. These formed the basis of the first research proposal sent to the NIH. The RIT had also arranged for anonymous piloting of the instruments with groups of American Indian parents and youths at a series of pizza parties. The parents and youths then participated in discussions about how they felt about the questions, and they suggested modifications, changes, additions, and deletions.

*Principle learned.* Involve the stakeholders from the first initiation of the research project through final reports, never letting contact be lost. The *Belmont Report*'s focus on respect for persons should be expanded beyond the research subjects to the community as a whole.

### Ethical Dilemmas at Implementation

The multiyear interaction with the RIT was supposed to address all the issues ahead of time. Unfortunately, a long time lag occurred between the initial establishment of the research implementation and funding of the project. It took approximately 3 years and three submissions to obtain funding. Through all this time, the researchers maintained contact with the RIT approximately twice a year. During that 3-year period, there were many personnel changes at the agencies that had sent representatives to participate in the RIT, and new members of the tribal council had been elected. Therefore the composition of the RIT changed over time. Also, even the original members who continued on the RIT lost track of their initial decisions concerning the research. Some individuals who had most vociferously demanded that the research address the more sensitive issues of sexual orientation, child abuse, and neglect had left the RIT; other members were no longer as concerned with those topics.

After funding was received, the new members of the team raised concerns about the intimacy of the questions, given their cultural prohibition against talking about intimate behaviors or feelings. They believed that even discussing parental problems, neighborhood violence, school violence, and life stressors could prove traumatic for the youths. They also expressed surprise and concern about the questions concerning sexual orientation, forgetting that those questions and topics had been initiated by earlier RIT members. The RIT was, however, excited about the potentiality of the research to serve as a service access point.

Despite a total of 4 years of contact with the community, the PI responded extremely foolishly to the new RIT's concerns. As a researcher and supposed expert on research, the PI repeatedly reassured the concerned RIT members that she had never had any youth experience stress reactions to the questions they were concerned with, and that youths usually enjoyed the process of talking about themselves. Relying on expertise and experience was a foolish mistake. The American Indians' historically based distrust overwhelmed any kind of reassurances or trust that had been built in the past. It took a while, but the PI began to see the validity of their concerns within the context of their own culture. Therefore, she moved from dismissing the concerns as a lack of knowledge of the research process to understanding their concerns as a valid cultural perspective. At this point, the PI and Co-PI began to work out a compromise that would enable the AIM-HI research project to accommodate all the concerns of the RIT while not removing or deleting data that were theoretically or potentially important for knowledge development about service access. Four areas of ethical concern necessitated compromises: (1) the impact of the potential stressfulness of the interview; (2) how the team responded to child abuse; (3) the ethical imperative to obtain services for youths whose needs were identified in the research; and (4) ownership of publication rights and data.

*Principle learned.* Accept the stakeholders' needs and concerns. Just because formal IRB issues are addressed does not mean that all ethical concerns are handled. Once again, the Belmont concept of respect should be expanded to the community as a whole, their culture and their perceptions.

### *Implementation Issue 1: Potential stressfulness of the interview*

The first ethical issue to be addressed was the conflict between research questions and the cultural prohibition against talking about potentially upsetting topics. The RIT was concerned that those questions would traumatize the youths because they would be placed in the position of violating their community norms. As a compromise, we agreed to allow the youths an easy "opt out" of those interview sections that the community believed were sensitive enough to cause the youth to become upset. Immediately before asking the questions in those potentially upsetting sections, we read a statement indicating that "some people may wish to not answer this set of questions, or may wish to answer them directly on the computer instead of having the interviewer read them out loud." Also, "Did they wish to skip the section? Answer it themselves on the computer? Or continue as before?" Their responses were then recorded in the computer. There were seven such potentially sensitive sections. Of course, this compromise involved a leap of faith on both parts. First, the interviewers would have to lead the youths through this section and not coerce them into responding to a section that they would rather opt out of, thus committing the ethical violation of nonvoluntary consent. Second, enough youths would have to see the importance of the questions and be willing to respond so that an overwhelming amount of missing data would not ruin the research. The second concern also involved the risk that the missing data would pertain to youths most likely to have high-risk responses and service needs.

*Consequences of real-world decisions regarding interview stressfulness.* Because we maintained data on youth responses, we know that this adaptation of the research project worked out quite well. Few youths elected to skip any sections, so the offer to skip did not compromise the integrity of the data. Ninety-four percent of the youths did not skip any sections. Three percent skipped only one section. And only 3% skipped two or more sections. Furthermore, the results from a reaction question at the end of the interview helped reassure the local stakeholders because more than 85% of the youths reported the interview to be interesting or helpful, and only one youth (0.3%) reported it to be upsetting (Stiffman et al., 2005a; Stiffman et al., 2005b).

*Principle learned.* The *Belmont report* mentions justice in the distribution of the benefits and burdens of research. Take risks concerning data completeness if it will assure responsiveness to what the community considers beneficence.—But document everything!

### *Implementation Issue 2: Responding to child abuse*

The second area of ethical concern focused on the community's concern about reporting child abuse. Responding to child abuse involves different degrees of ethical issues for close-knit communities than for urban communities. Especially in the reservation areas, all the families essentially knew one another. Our first concern in the case of child abuse was to protect the child. To do that, we wanted to be sure that we gave equal importance to protecting the child, the family, and the tribe. The RIT told us quite clearly that, because of the extended family atmosphere in the tribal area, breach of confidentiality could destroy the family, their reputation, their self-esteem, and their ability to live in physical proximity to the study. We, therefore, had to grapple with a number of questions that do not have clear legal guidelines: Would reporting cause more harm than good? What was the severity and currency of the youth's report of physical or sexual abuse? Was abuse likely to recur? Were these examples of cultural or ethnic variations of discipline that would only be interpreted as abuse outside the community? Were relevant service providers and protective services already involved? And finally, if we reported abuse, were services available?

We, therefore, stepped back from just the standard reporting requirement, and developed a detailed plan with the cooperation of the RIT and local protective services. We, along with protective services, provided extensive training in abuse to interviewers and field supervisors. The training acquainted interviewers with mandatory child abuse reporting laws and procedures for protecting themselves and the youths physically and legally. It dealt with the potential consequences of acting as a mandated reporter and with concerns that reporting abuse might escalate the abuse. The RIT and the interviewers were all aware that the foster care situation in the community was inadequate and that protective services had strained resources with few services and a high caseload. Therefore, the protective services group was wary about being overburdened with reports that would never become active cases or re-reports of existing cases. The community and protective services helped us develop a screening tool to eliminate obvious false-positives or pretreated situations.

Positive responses concerning child abuse triggered triage questions and dialogue at the end of the interview. The triage section reminded the youths of the consent form that they had signed, which described that confidentiality would be breached if they revealed something that indicated that they or someone else might be harmed in the interview period. The dialogue explained the interviewer's concern, and a series of stepped questions clarified the youth's risk level and told the youth about any concerns we had. Interviewers obtained youth permission so that confidentiality was not violated without the youth being aware.

*Consequences of real-world decisions regarding child abuse reporting.* The researchers did not anticipate much difficulty in the procedures for reporting child abuse because they had extensive experience conducting similar interviews in other high-risk populations in which few teens self-reported abuse (Auslander et al., 2002; Cunningham, Stiffman, Doré, & Earls, 1994; Stiffman, 1989a). However, the AIM-HI project immediately began confronting abuse reports. In fact, some of the first reports of abuse were about perpetrators who had worked with the RIT in developing the research project. More than one-fourth of the youths answered questions that triggered a flag for abuse. Luckily, we had the set of triage questions that eliminated one-fifth of those reports because they had already been involved with protective services for the latest incident of abuse. An additional 25%, when asked to clarify the incident, described an incident that was clearly not reportable abuse (e.g., peer bullying or being picked on at a playground). Seven percent of the youths agreed to immediately call protective services themselves (with the interviewer present). Noteworthy is the fact that the type of abuse reported by most of the youths would not be currently actionable by protective services. For approximately half the youths reporting abuse, the last occurrence was more than 1 year before, with approximately one-fourth of those occurring 10 or more years ago. One in five youths told us that they had already received help for their abuse situation, and many were in a new family situation with the abuser out of the home or in prison. Although protective services would have taken such a report, and monitored the family, in these cases they probably would not have done anything beyond monitoring. Given their situation at that time of work overload, it would have only created a stressful situation for the families and for protective services, with no positive outcome.

Unfortunately, one of the protective service providers breached confidentiality by revealing to a parent the presence and source of an incident report made because of an AIM-HI interview. The parent responded with a telephoned death threat (left on an answering machine) to the research supervisor, who was also a tribal member, and to the interviewer. Although nothing came of the threat, it was an unanticipated concern that would have been much less likely to occur in a larger, more anonymous and diverse community.

Because the PIs and their staff kept records on reports and on the answers to the triage questions, the records showed that longitudinal attrition rates for youths exposed to the triage questions were no higher than those for youths not so exposed (6% for both groups).

*Principle learned.* Abide by the legal regulations, but be aware of external issues such as privacy, changing circumstances, prior reporting, or treatment, and an already overwhelmed protective services system. Get agreements from all parties to triage as part of the research process, as reporting without triage may violate the Belmont concept of beneficence.

### *Implementation Issue 3: Obtaining services for needy youths*

The ethical issue of not ignoring youths who were identified as having needs varies in intensity depending on the critical urgency of the type of need. There was no ethical question about ignoring youths who were suicidal. However, the RIT also wanted us to address the generic need for services in youths participating in our survey. Our research aim was to examine their service use over a 4-year period. Thus, we were confronted with an issue: if our research became a service entry point for less critically needy youths, it would interfere with the natural history of services received by these youths. On the other hand, ignoring problems uncovered in the interview was not fair to the youths. The RIT quite clearly believed that their purpose in supporting the AIM-HI study was to get needy youths to services. After much consideration, we decided that we were not only interested in the natural history of services but also interested in service use, per se. Therefore, as long as we documented the youths who were sent to services by our project, we would still be able to talk about the services they received over the course of 4 years. Further, as social service researchers, we always felt somewhat guilty about learning that particular youths needed services while we simply moved on through the questionnaires and process of research without helping them. The American Indian community has resented research concerning their needs that never addressed services. Therefore, we agreed to use the research as an entry point to services. This is contrary to the historical research position, which has been to not react to any responses in a structured interview (Ventura, Liberman, Green, Shaner, & Mintz, 1998). In fact, it defied traditional research training, which taught that "the process of listening without reacting is helpful in and of itself" and to "never react." The PIs agreed with the RIT that not acting might place a child in future danger or leave the youth feeling hurt or unacknowledged. The team worked on identifying those areas of the interview that may indicate a need for services. These included mental health problems, suicidality, and alcohol or drug abuse. We also arranged for agencies with individual service providers to be on 24-hour call to respond to subjects participating in the AIM-HI project in both the urban and reservation areas.

Suicidality issues were particularly important for the RIT because American Indian suicide rates are higher than rates in other communities (Office of Technology Assessment, 1990). Although 8% of youths in the general community have been found to make suicide attempts (Centers for Disease Control and Prevention, 2002), rates of suicide among American Indian youths are thought to be as high as 23% (Manson, Beals, Dick, & Duclos, 1989). Suicide is, in fact, the second leading cause of death for American Indian youths. The RIT was concerned that, due to the cultural prohibitions about revealing feelings, if we asked about suicidality and if we overreacted to their responses, the youths would deny the feelings later. In the long run this would result in less help rather than more help. However, we had to prevent youths from killing or harming themselves. Also, despite any confidentiality issues, basic ethical principles indicate that no research contribution can outweigh the value of a human life. Therefore, another set of triage questions was developed for youths who had given positive responses to any of the questions about suicide. These triage questions helped the team determine the currency of the suicidal feelings, the access to lethal means, current involvement in treatment, and current involvement of a parent or mentor. Plans were made for an interviewer and the 24-hour caseworker to stay with any actively suicidal youth until a parent or provider was brought in. For all youths who were actively suicidal, the interviewer explained the necessity for breaking confidentiality based on concerns for the youth. As with the child abuse issue, interviewers obtained permission to break confidentiality. Also, as with the child abuse procedures, incident reports documenting the youth's responses and the interviewer's actions were developed for every youth who was suicidal to any degree.

*Consequences of real-world decisions regarding service provision to suicidal youth*. Suicidality reports were two to three times higher than in the researchers' previous research (Freedenthal & Stiffman, 2004; Stiffman, 1989b). One-third of the youths reported feelings of suicidality, and one in five had attempted it. The triage questions worked well because, under further questioning, only 14% of the suicidal youths reported being suicidal at that time. And only a few of those were actively contemplating suicide. However, the triage questions did not work as well in referring youths to services. Only 15% of those who were suicidal at the time agreed to call someone for help. None agreed to call a parent or physician, but some agreed to call a local agency or a hotline. Almost 90% of the youths who were suicidal but refused to make a call did verbally promise the interviewer that they would not attempt suicide before help was arranged. In these cases, an immediate report was given to the supervisor, who informed the 24-hour AIM-HI clinical backup at the local agency, and the interviewer stayed with the youth until a parent or provider or someone acceptable to the youth arrived.

The questions about suicide and the triage section of the interview became extremely important to the project and to the communities. In the year of the first interviews, four youth suicides occurred on the reservation. A rumor began that all the youths who committed suicide did so because of the stress of the interview. As you can imagine, the tribal council was concerned and called us in to address it. Fortunately, we had data on the suicidality of all the youths and on what actions we took. None of the suicides was an AIM-HI subject. We could also show that the suicidal subjects got help and that they were not stressed by the questions. (Of great importance to note, the youths who participated in the study were referred to services for suicidality. Because of the suicide attempts of youths who had not participated, the community immediately reacted by instituting a suicide prevention program.)

*Consequences of real-world decisions regarding service provision to youth with noncritical needs.* The plan that the PIs developed with the RIT for the youths who were less critically needy was designed to help provide services in such a way as to avoid influencing the interview's impact on future youth actions as much as possible. All youths were given lists of local services and hotline numbers for self-referral. In addition, those youths who, at the end of the interview, were flagged as having mild problems were taken through a set of triage questions that suggested that the youths consult a teacher, physician, social service provider, or parent. They were encouraged to get help from appropriate specific agencies on the referral list. For youths with serious mental health or addictions problems (meeting criteria for a diagnosis), the supervisor was notified and a clinical backup was initiated on behalf of the youth after the youth was told that such a procedure would be done.

Almost 90% of the youths had at least one problem that required directing them to services. This unexpected percentage put a time and financial burden on the interviewers. Each interviewer had to spend extra time discussing the issues with each youth, recommending services and, in some cases, getting the youth's agreement to break confidentiality. The biggest glitch was in the service backup. Neither the community nor the researchers had expected the high caseload uncovered by the research project. Although the 14-hour backup functioned effectively for youths in crisis, the long-term service capacity of the agencies was stressed by the number of youths referred by the reports. Also, implementation of services was crippled by individual provider actions. For example, although the agencies were prepared to respond to our referral of youths, providers had not been trained by the agency in how to respond to these types of referrals. Some providers believed that they could not offer services without parental permission, even if the youths were in danger. Thus, they sometimes made individual decisions to not follow through after receiving the names of the youths from the agency. They did not believe the research consent form was adequate, even though their agency representatives had shaped it. The researchers had no interaction with the practitioners and the agencies by which they were employed, so the researchers were powerless to monitor or encourage service provision.

The great demand on services, however, did trigger action within the reservation community as they became aware of the generic need for services for addictions, prevention, and mental health programs for youths. Also, within 1 year, multifaceted addiction prevention, family strengthening, and youth support policies and programs were instituted. After a 4-year period, the benefits were reflected by dramatically reduced behavior problems in the reservation youths. This same improvement did not occur with the urban youths, where no additional prevention or treatment programs were instituted in response to early waves of data (Stiffman et al., 2007).

*Principle learned.* Take research risks by being open about research results and assuring that a community's expected benefits concerning service referral balances the burden that they assume in supporting the research project.

### *Implementation Issue 4: Publication rights*

The final ethical conflict revolved around the concern of the RIT that any publications be acceptable to the American Indian community. This is an area totally unaddressed by IRB regulations but that involves its own set of researcher ethics and community issues. American Indians have had extensive experience with publications that speak about all their problems without ever acknowledging any strengths. In response to the community concern that this would not occur, we did a number of things. First, we refocused the interview to emphasize strengths so that every section that asked about problems or needs also asked about strengths in that particular area. Second, we had each community nominate key individuals who would review any papers and reports, with the agreement that they could veto or edit the paper prior to its presentation or publication. If multiple papers were presented on the same topic, the reviewers would not check each individual paper, but instead would review the exemplar one on the relevant topic. This guideline was instituted so that the reviewer would not be burdened with a lot of manuscripts and information. Third, the PIs agreed to not reveal the name of the participating reservation or urban area.

*Consequences of real-world decisions regarding publication*. The policy of inviting community representatives to review all publications had benefits and drawbacks. Unfortunately, some of the community representatives had no higher education experience. They found the documents extremely difficult to read. Although they had been part of obtaining certain aspects of information, such as information about the actions of traditional healers, they had second thoughts. For example, one reviewer became concerned that publishing anything about traditional healing would violate the traditional secrecy of this type of behavior. We had to explain that what we were publishing was not secret information about the conduct of traditional spiritual ceremonies, but information on the actions that traditional healers took on the part of the mental health of their young clients. Therefore, several papers triggered much discussion between the PIs and the reviewers. One positive aspect of this review system was that the community representatives identified a number of weaknesses in the manuscripts, which was extremely helpful.

The process of publication approval was interrupted by unrelated negative American Indian/researcher interactions occurring in an entirely different study with other reservations. A researcher who had been given access to clinic data by a tribe was discovered to have abstracted genetic history information that had not been part of the original agreement. Some years later, that researcher gave these abstracted data to another researcher for a totally different research project. This different research project was published in such a way that it could have redounded to discrimination and problems within the community itself. A member of that tribe was at a meeting where the information was presented with full tribal identification, and the member knew that the tribe had not agreed to that study nor worked with that researcher. The tribe was horrified at another dramatic violation of American Indian/researcher agreements. This incident verified many American Indian communities' historical distrust of researchers.

Unfortunately, at approximately the same time that this incident was reported by newspapers, the PIs submitted a paper for approval that was first-authored by a member of the faculty at a local university. That university had been part of the initial RIT but had not been part of the funded research project. The university participated in the RIT with the verbalized expectation that at some level they would participate in the research project itself. No one on the initial local team was experienced enough in academic research to actually conduct it. (They had been part of the initial proposal to the NIH, resulting in a critique about their high cost associated with little experience. This resulted in resubmissions that did not include them.) However, by the time the data were collected, new university faculty had been hired who did have that experience. Therefore, de-identified AIM-HI data were shared with local university faculty so that they could take lead authorship in papers that would advance their own research agenda and speed the publication of AIM-HI data. The PI of our team understood that this was part of the original open agreement. However, when a paper was sent to the American Indian representatives for review, and they saw the new local university names as lead authors on the document, they became upset. They made it clear that they did not want anyone else to have the data, even if it had no identifying information, or to use it for publication. They would only allow the PIs and their staff who conducted the research to use the data for publication. The community insisted on a narrow definition of who could publish from the data. Unfortunately, that particular paper had to be scrapped. Luckily, it was not a paper on a central theme or aim of the research itself.

At this point, the American Indian reviewers also decided that they would not let data on their youths be turned over to the NIH at the conclusion of the study. They insisted that they would keep the de-identified data and would personally vet any researchers wanting to use it. They explained that, culturally, taking information was akin to taking something tangible and personally owned. Respect for that information demanded that they not give up part of themselves to unknown others.

*Principle learned.* Recognize that publishing information about problems in a community is not a just distribution of the burdens and benefits of research and may violate the concept of community beneficence as well as confidentiality. Recognize also that researchers are entrusted with information that is a valuable commodity demanding respect and care.

## Conclusion

This chapter presents ethical, cultural, and interactive issues in conducting real-life research. We illustrated the issues with events and decisions confronted by a single research project. It is our hope that this chapter is a good representation about the general dilemmas that arise during many research projects. Hopefully, our solutions, our problems, and our data about the effects of our solutions will fill a needed gap in the research literature. Our desire is to let our experience serve as a thought-provoking guide for other researchers and to increase the repertoire of creative and innovative solutions for such issues.

The conflicts that we confronted in our research between a research culture and a community culture are not necessarily different from issues confronted by other researchers in other communities. However, the pressure within the American Indian community to provide services, the concern about the sensitivity of questions, and the historical feelings of distrust were different. We tried to balance research needs, ethical issues, and community viewpoints creatively without compromising research integrity. We believe that it is important to recognize that there are no easy answers to these conflicts. No matter what the researcher decides, there are potential pitfalls for the community, the youthful subject, and the research project. Stakeholder involvement, supervisor attention, expert consultation, and provider availability are all needed to effectuate workable compromises. As researchers we must accept that we have to live with competing pressures. Open communication among researchers will go far to advance our ability to handle these issues and will help build a backlog of exemplar solutions that can be used, avoided, or adapted by others.

## POINTS TO REMEMBER

- Many ethical and cultural issues go far beyond what an IRB considers.
  - Review board regulations may not fully address ethical imperatives or specific cultural issues.
  - Leaving collaboration and involvement only to the development of the consent document and study design is inadequate.
- Principles for applying cultural and ethical issues beyond IRB regulations.
  - Involve the stakeholders from the first initiation of the research project through final reports, never letting contact be lost.
    - Reframe research focus to fit within needs and priorities of community.
    - Develop a mutually agreed upon research plan and associated instruments.
    - Maintain ongoing contact throughout the process.

- Accept the stakeholders' needs and concerns.
  - Understand their view of the potential stressfulness of the interview.
  - Understand the benefits they want to get from the research.
- Take risks concerning data completeness if it will assure responsiveness to what the community considers beneficence—but document everything!
  - Find creative ways to allow subjects to choose to answer.
- Take research risks by being open about research results and assuring that a community's expected benefits concerning service referral balances the burden that they assume in supporting the research project.
  - Respond to child abuse by triaging.
  - Obtain immediate services for suicidal youth.
  - Obtain services for noncritical needs of youths through referrals.
- Recognize that publishing information about problems in a community is not a just distribution of the burdens and benefits of research, and it may violate the concept of community beneficence as well as confidentiality. Recognize also that researchers are entrusted with information that is a valuable commodity demanding respect and care.
  - Establish methods of review, editing, and vetoing publications.
  - Establish methods of authorship assignment.
  - Agree on methods to protect confidentiality.

## REFERENCES

Attkinson, C. C., Rosenblatt, A., & Hoagwood, K. (1996). Research ethics and human subjects' protection in child mental health services research and community studies. In K. Hoagwood, P. S. Jensen, & C. B. Fisher (Eds.), *Ethical issues in mental health research with children and adolescents* (pp. 43–57). Mahwah, NJ: Lawrence Erlbaum Associates.

Auslander, W. F., McMillen, J. C., Elze, D., Thompson, R., Jonson-Reid, M., & Stiffman A. R. (2002). Mental health problems and sexual abuse among adolescents in foster care: Relationship to HIV risk behaviors and intentions. *AIDS and Behavior,* 6(4), 351–359.

Note: Some of the information in this chapter has been published previously in other articles that are referenced in the references section.

Beauvais, F. (1999). Obtaining consent and other ethical issues in the conduct of research in American Indian communities. In M. De La Rosa, B. Segal, & R. Lopez (Eds.), *Conducting drug abuse research with minority populations* (pp. 167–184). New York: The Haworth Press, Inc.

*The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. (1979). Washington, DC: National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.

Centers for Disease Control and Prevention. (2002). Web-based injury statistics query and reporting system (WISQARS). Retrieved June 28, 2007 from www.cdc.gov/ncipc/wisqars.

Cunningham, R., Stiffman, A. R., Doré, P., & Earls, F. (1994). The association of physical and sexual abuse and HIV infection: Implications for public health. *Child Abuse and Neglect, 18*, 233–245.

Dorn, L. D., Susman, E. J., & Fletcher, J. C. (1995). Informed consent in children and adolescents: Age, maturation and psychological state. *Journal of Adolescent Health, 16*, 185–190.

Doyle, K. O. (2001). Opinion research in Indian country. *International Journal of Intercultural Relations, 25*, 511–530.

Freedenthal, S., & Stiffman, A. R. (2004). Suicidal behavior in urban American Indian adolescents: A comparison with reservation youth in a southwestern state. *Suicide and Life-Threatening Behavior, 34*, 160–171.

Glantz, L. H. (1996). Conducting research with children: Legal and ethical issues. *Journal of the American Academy of Child Adolescent Psychiatry, 34*, 1283–1291.

Hoagwood, K., Jensen, P. S., & Fisher, C. (1996). Toward a science of scientific ethics in research on child and adolescent mental disorders. In K. Hoagwood, P. S. Jensen, & C. B. Fisher (Eds.), *Ethical issues in mental health research with children and adolescents* (pp. 2–14). Mahwah, NJ: Lawrence Erlbaum Associates.

Hoagwood, K., Jensen, P. S., & Leshner, A. I. (1997). Ethical issues in research on child and adolescent mental disorders: Implications for a science of scientific ethics. In S. Henggeler, & A. B. Santos (Eds.), *Innovative approaches for difficult-to-treat populations* (pp. 459–476). Washington, DC: American Psychiatric Association.

Kalichman, S. (1993). *Mandated reporting of suspected child abuse: Ethics, law, and policy*. Washington, DC: American Psychological Association.

King, N. M. P., & Churchill, L. R. (2000). Ethical principles guiding research on child and adolescent subjects. *Journal of Interpersonal Violence, 15*, 710–724.

Koocher, G. P., & Keith-Spiegel, P. C. (1990). *Children, ethics, & the law: Professional issues and cases*. Lincoln, NE: University of Nebraska Press.

Kotch, J. B. (2000). Ethical issues in longitudinal child maltreatment research. *Journal of Interpersonal Violence, 15*, 696–709.

Levine, R. J. (1995). Children as research subjects, ethical and legal considerations. *Child and Adolescent Psychiatric Clinics of North America, 4*, 853–868.

Manson, S. M., Beals, J., Dick, R. W., & Duclos, C. (1989). Risk factor for suicide among Indian adolescents at a boarding school. (Rep. No. 104). *Public Health Reports*. U.S. Dept. of HHS, *104*(6), 609–614.

Marshall, C. A. (2001). Cultural factors in conducting research and ethical responsibilities in serving American Indians with disabilities. In C. A. Marshall (Ed.), *Rehabilitation and American Indians with disabilities: A handbook for administrators, practitioners, and researchers* (pp. 165–175). Athens, GA: Elliott & Fitzpatrick, Inc.

Melton, G. B. (1992). Respecting boundaries: Minors, privacy, and behavioral research. In B. L. Stanley & J. E. Sieber (Eds.), *Social research on children and adolescents: Ethical issues* (pp. 65–84). Newbury Park, CA: Sage Publications.

Myers, J. E. B. (1992). *Legal issues in child abuse and neglect*. Thousand Oaks, CA: Sage Publications.

National Institutes of Health. (1998). *NIH policy and guidelines on the inclusion of children as participants in research involving human subjects*. Bethesda, MD: National Institutes of Health.

Norton, I. L., & Manson, S. M. (1996). Research in American Indian and Alaska Native communities: Navigating the cultural universe of values and process. *Journal of Consulting and Clinical Psychology, 64*, 856–860.

Office of Technology Assessment. (1990). *Indian adolescent mental health*. (Rep. No. OTA-H-446). Washington, DC: U.S. Government Printing Office.

Osher, T. W., & Telesford, M. (1996). Involving families to improve research. In K. Hoagwood, P. S. Jensen, & C. B. Fisher (Eds.), *Ethical issues in mental health research with children and adolescents* (pp. 29–39). Mahwah, NJ: Lawrence Erlbaum Associates.

Putnam, F. W., Liss, M. B., & Landsverk, J. (1996). Ethical issues in maltreatment research with children and adolescents. In K. Hoagwood, P. S. Jensen, & C. B. Fisher (Eds.), *Ethical issues in mental health research with children and adolescents* (pp. 113–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Putney, S. B., & Gruskin, S. (2002). Time, place, and consciousness: Three dimensions of meaning for us institutional review boards. *American Journal of Public Health, 92*, 1067–1076.

Shuster, E. (1997). Fifty years later: The significance of the Nuremberg Code. *The New England Journal of Medicine, 337,* 1436–1440.

Singer, P. (1984). Ethics and sociobiology. *Zygon: Journal of Religion and Science, 19,* 141–158.

Steinberg, A. M., Pynoos, R. S., Goenjian, A. K., Sossanabadi, H., & Sherr, L. (1999). Are researchers bound by child abuse reporting laws? C*hild Abuse and Neglect, 23,* 771–777.

Stiffman, A. R. (1989a). Physical and sexual abuse in runaway youths. C*hild Abuse and Neglect, 13,* 417–426.

Stiffman, A. R. (1989b). Suicide attempts in runaway youths. *Suicide and Life-Threatening Behavior, 19*(2), 147–159.

Stiffman, A. R, Alexander-Eitzman, B., Silmere, H., Osborne, V., & Brown, E. (2007). From early to late adolescence: American Indian youths' behavioral trajectories and their major influences. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(7), 1–10.

Stiffman, A. R., Brown, E., Striley, C. W., Ostmann, E., & Chowa, G. (2005a). Research solutions for cultural and human subjects' issues concerning American Indian youth. *Ethics and Behavior, 15*(1), 1–14.

Stiffman, A. R., Freedenthal, S., Brown, E., Ostmann, E., & Hibbeler, P. (2005b). Field research with underserved minorities: The ideal and the real. *Journal of Urban Health, 82*(2), Supplement 3, iii56–iii66.

Stubben, J. D. (2001). Working with and conducting research among American Indian families. *American Behavioral Scientist Special Issue: Understanding American Indian Families, 44,* 1466–1481.

Thompson, R. A. (1992). Development changes in risk and benefit: A changing calculus of concerns. In B. L. Stanley & J. E. Sieber (Eds.), *Social research on children and adolescents: Ethical issues* (pp. 31–64). Newbury Park, CA: Sage Publications.

Ventura, J., Liberman, R. P., Green, M. F., Shaner, A., & Mintz, J. (1998). Training and quality assurance with structured clinical interview for DSM-IV. *Psychiatry Research, 79,* 163–173.

Weaver, H. N. (1997). The challenges of research in Native American communities: Incorporating principles of cultural competence. *Journal of Social Service Research, 23,* 1–15.

# 7

# Creating Interdisciplinary Research Teams and Using Consultants

*John Landsverk*

> Men who know the same things are not long the best company for each other.
> —*Emerson,* Representative Men: Uses of Great Men
> Good company and good discourse are the very sinews of virtue.
> —*Izaak Walton,* The Complete Angler. Pt. I, ch. 2.

Doctoral education, in general, and the final step of carrying out the dissertation project, in particular, have a focus on the doctoral student as a solitary scholar working alone. This focus stands in sharp contrast to the dominant mode of large-scale research projects created and implemented by interdisciplinary research teams. When working on a doctoral dissertation project as a solitary scholar, students gain little experience in developing interdisciplinary teams. This lack of training in setting up such teams in the formative years of a fledgling researcher is in marked contrast to the dominant model of most funded health research in the modern era, especially studies funded by the National Institutes of Health (NIH) in the United States. In fact, it may not be too much of a generalization to say that almost all federally funded research today is carried out not only by teams but also by teams composed of investigators from multiple disciplines. Few scientific questions do not require multiple investigators with multiple scientific perspectives and expertise, even if there remain disciplines where many excellent studies continue to be carried out by single investigators such as in anthropology. This may be a major reason why doctoral education often must be enhanced by postdoctoral work, an experience much closer in character to the real structure and processes of modern field research.

This chapter discusses lessons learned from the author's 20 years of experience in putting together interdisciplinary teams, with a special focus on the various roles that consultants may play. The chapter draws from a wide variety of field projects, almost all funded by federal agencies, including the NIH. In keeping with the intent of the book, the discussion draws from practical experience in conducting field research rather than the voluminous published literature about the use of expertise and consultation. The focus is on lessons learned rather than literature reviewed.

## Interdisciplinary Research Studies and the Need for Reengineering Research Teams

Almost without exception, modern biomedical and social science research is conceived and carried out by carefully designed investigative teams. These teams have a wide range of expertise that often cross discipline and subdiscipline lines among the investigators participating. The teams often include a group of consultants who have highly specialized skills seen as necessary for the project's aims but not available from the team of investigators.

The shift in modern research toward interdisciplinary research has been greatly accelerated by the recent NIH Roadmap initiative for reengineering the clinical research enterprise. This initiative includes a concerted effort to move research toward translational forms as well as focus sharply on the requirement to use interdisciplinary research teams (Culliton, 2006; Zerhouni, 2003, 2005; for a detailed overview of the Roadmap initiative and the programs that have come out of it, see the Web site http://nihroadmap.nih.gov.). This initiative is based on the premise that traditional divisions among scientific disciplines need to be bridged to speed the pace of scientific discovery. Planning for interdisciplinary research requires changes in many, if not all, aspects of scientific processes and support, including the training of investigators and the development of new research technologies.

The Roadmap initiative has identified three types of research leading to improvements in the public health of the nation: basic research that informs the development of clinical interventions (e.g., biochemistry, neurosciences), treatment development that crafts the interventions and tests them in carefully controlled efficacy trials, and what has come to be known as service system and implementation research in which treatments and interventions are brought to and tested in usual care settings. (For an interesting discussion of this interplay, see Westfall, Mold, & Fagnan, 2007.) Based on this tripartite division, the Roadmap further identifies two translation phases critical for moving from the findings of basic science to improvements in the quality of health care delivered in community clinical and other delivery settings. The first translation phase brings together interdisciplinary teams that cross and integrate the work being done in the basic sciences and treatment development science, or translating neuroscience and basic behavior research findings into new treatments. The focus of the second translation phase is to translate evidence-based treatments into service delivery settings and sectors in local communities. Although the second phase has received far less emphasis and support at the NIH, it requires interdisciplinary teams of treatment developers and researchers skilled in understanding service delivery systems, such as economists, sociologists, anthropologists, and clinical psychologists.

These translational teams are designed to break down the traditional silos that have arisen with the development of scientific disciplines (Zerhouni, 2003). In short, modern field research is an increasingly interdisciplinary game played with complex interdisciplinary teams brought together to speed the discovery of new methods for improving the quality of medical and mental health services provided to consumers in usual care settings. These complex teams need to be formed and sustained with great care and skill and with attention to what can go wrong. The issues related to putting such teams together and maintaining them as productive units during field research are the focus of this chapter.

## Forming Interdisciplinary Teams

Multiple types of research structures have been created through support under diverse funding mechanisms for carrying out scientific work. These include investigator-initiated single research projects, funded at the NIH under the R01 mechanism (technical terms for funding mechanisms at the NIH are used because of the need for specificity, but other counties and funding sources would use other technical terms for mechanism), as well as more complex structures, such as research program grants, research centers, and research networks. This chapter primarily focuses on individual research projects because they represent the least complex of the interdisciplinary structures. However, program grants, centers, and research networks all assume an enhanced interdisciplinary structure.

The creation of interdisciplinary research teams takes place in the proposal development stage of field research, when the research team drafts the key research questions and hypotheses as well as the methods to address these questions. These individual pieces of the proposal are packaged together with a tightly knit rationale supporting the innovation and rigor of the proposed research in an effort to secure funding to carry out the field research. A critical issue in the scientific review of the application for potential funding is whether the expertise and experience of the research team match the scope and goals of the proposed study. The application begins with a description of the specific aims to be addressed. The language used for these aims directly affects the nature of the interdisciplinary team to fit the proposed work expected by the reviewers. For example, if the aims contain a proposed economic analysis, the team will need to include an economist with specialized experience in the type of phenomenon related to cost data. Therefore, if the project is focused on mental health care, the economist must have experience dealing with mental health data and costs for mental health care. If the aims include the development and testing of a new treatment or intervention, the team should include an expert in the development of that type of intervention. An epidemiologist, for example, will not be seen as a good fit for that role on the investigative team. In short, the specific aims of the proposed study must be matched by the expertise of the investigative team.

Two sections follow the specific aims section, where the required expertise of the research team is projected and documented. First, the “background and significance section” lays out the scientific literature that directly relates to the aims of the proposed study. Here the task is to demonstrate in-depth knowledge of the research fields pertaining to the proposed work, with the critical task of showing how the proposed study will add an innovative piece to the existing scientific literature. Citations in this section by members of the investigative team can show how they have already contributed to the scientific field in prior studies.

An even more direct way to demonstrate the necessary expertise is provided in the preliminary studies section that follows the background and significance section. Here the application directly discusses the special and specific expertise of the selected investigative team members needed to carry out the proposed research project. This section documents their experience with the concepts and methods required in the study through detailing preliminary studies already completed to lay the groundwork for the proposed research.

The interdisciplinary team generally is divided into two types of research roles: a set of investigators and a set of consultants. Although both sets need to show specialized expertise, they differ markedly in the structure of scientific activities. In addition, within the set of investigators is the demarcation between the role and title of principal investigator (PI), who has responsibility for the overall scientific direction of the study as well as its administration (personnel and budget), and co-investigators, who partner in the scientific work with the PI. Another frequently used title is co-principal investigator, which has been used informally but has not been supported in any NIH program announcement as of this writing. National Institutes of Health has historically required a single scientist to carry out the role of PI. Recently, however, the NIH has promulgated a policy in which dual PIs can be named on an application for funding field research (NOT-OD-06-036 Establishment of Multiple Principal Investigator Awards for the Support of Team Science Projects). This development has been rationalized as a way to show multiple leaders in the structure of the team, in contrast to prior policies, which allowed only one PI. It also demonstrates the increasingly complicated structures seen as necessary to carry out complex field research. Although this development will likely be associated with decreased informal use of the co-principal investigator role, it is too early to tell if this will occur. This change also can be viewed as acknowledgment of the need on research projects for interdisciplinary leadership as these teams have become the norm. It also can be seen as a part of the trend of scientific studies increasingly being carried out in multiple sites by multiple interdisciplinary teams.

These investigator roles most often are filled by researchers who have some experience working together on other projects and have clearly identified and documented expertise specifically related to the proposed research study. Previous working experience is critical to document because reviewers of these applications often have considerable experience with problems that can arise in new research teams.

Although documentation of expertise certainly includes relevant advanced degrees, it also must demonstrate specific experience gained in prior research directly related to the methodological challenges of the proposed work. In addition, application reviewers always award more credibility to this experience if it was gained in competitively funded research and resulted in significant publications in highly regarded journals. In addition, reviewers will examine the application to determine whether good fit of team members exists without large overlaps and redundancies in background and experience.

In developing credible applications, balancing of the team must be delicate when the proposed PI is new to the role. Here an experienced researcher with considerable time assembling interdisciplinary teams can balance the inexperience of the new PI. The importance of this balance can be signified by naming the experienced researcher as a co-investigator with sufficient time allocated to the project so that reviewers will believe adequate guidance will be provided to assist the new PI in meeting the complex challenges of field research. In addition, reviewers will consider prior funded research by the new PI as important demonstration of his or her capability—not only to compete for funding but also to carry out the field work successfully. This prior experience is especially well demonstrated when the new investigator has already competed successfully for and carried out smaller-scale field research, which would be classified by separate funding mechanisms at the NIH as R03 (small grant), R21, or R34 (both the R21 and R34 would be small-scale developmental grants) funding mechanisms.

## Expertise in Specialized Areas: Methods, Economics, and Ethics

A distinction between substantive and methodological expertise is useful when considering interdisciplinary research teams. Substantive expertise is based in knowledge of a specific area, such as research on child welfare or adult public mental health. This expertise comes from both a discipline, such as clinical psychology or social work, and long experience in conducting research about a specialized phenomenon. Methodological expertise, on the other hand, is based in specialized experience around a set of methods or research tools. Two common types of specialized investigators are statisticians, who are experts in the analysis of quantitative data, and anthropologists, who are experts in the analysis of qualitative data, such as ethnographic narrative data or key informant or focus group data. These specialized methods of social science research have become required tools in research studies and are critical roles to fill. A recent example of such a team combining substantive (clinical psychology and implementation research) and methodological (anthropology) expertise is demonstrated in the article "Implementation of Evidence-Based Practice in Child Welfare: Service Provider Perspectives" (Aarons & Palinkas, 2007).

A common error made by fledging substantive researchers who submit applications to highly competitive sources of funding such as NIH is to put themselves forward as providing methodological expertise in the proposed study. This is especially true if they have gained their substantive expertise at highly qualified doctoral and postdoctoral research programs and have done their own statistical and/or qualitative analyses, as is required in these programs. Almost without exception, NIH review groups will not find an application acceptable unless the research team includes a statistician. Although anthropologists have only recently taken research team roles on interdisciplinary social science applications, they increasingly are seen as indispensable as statisticians. In fact, review groups in both the services and interventions areas at the National Institutes of Mental Health (NIMH) now include statisticians and anthropologists as standing members of the initial review groups. In addition, the bar has been raised regarding the experience of these experts, with the expectation that they have the expertise in highly specialized areas, such as longitudinal cohort analyses—for which exceptionally innovative work has been done over the past two decades with the development of modeling techniques such as random effects (Center for Health Statistics, University of Illinois at Chicago) or latent variable models (Mplus). My own experience in mental health services research suggests that counting data, such as that found in administrative databases, with episodes of outpatient and inpatient care can be highly skewed and requires considerable experience from well-trained statisticians.

Expertise in economics is another specialization that has become quite important in the field of health services research. Cost analysis and cost benefit analysis are methods often required for a research study to contribute to policy-relevant findings. Again, review groups will not accept substantive researchers as having credible expertise to carry out these types of analyses.

A third area of expertise with increasingly specialized knowledge is the field of ethics. Many research applications to the NIH now include an investigator with specialized training and experience in the ethics of research with human subjects. In my own research on the need for and use of mental health services by especially vulnerable children who are involved with the foster care system, I have turned to this type of specialized methods expertise for help with risk/benefit calculations and the determination of who can be approached in the consent process when a child has been removed from the biological parent and placed in foster care (Putnam, Liss, & Landsverk, 1996). Celia Fisher from Fordham University has been extremely helpful with these issues at our center; she has extensive training and experience in addressing these issues in field research (Hoagwood, Jensen, & Fisher, 1996).

## Identifying and Using Consultants

A clear distinction exists between the leadership and co-investigators of the interdisciplinary research team and consultants who are brought on board projects. Consultants are used for specialized expertise, much like that required of investigators, but that can be provided in short but strategic doses to the regular investigative team during critical phases of the research study. Often these consultants live and work at some considerable distance from where the field research will take place, constraining their ability to assume a believable co-investigator role. Sometimes the expert is placed in the role of consultant even when he or she lives in the same area as the research team but is unwilling to commit to the time required to be an investigator.

What constitutes such specialized expertise? I have taken the perspective of NIH-funded research because that is what I know best. From that vantage point, expertise involves not only content specialization but also experience working in competitive funded research projects with good track records of significant publication.

Following are three examples of types of consultants who were added to applications or funded studies in response to reviewer comments in the NIMH review process. These examples illustrate the perception of reviewers and instances in which even an experienced team did not anticipate that the application would be viewed as lacking specialized technical expertise.

The first application was for a project named “Integrating Evidence and Practice of Youth Psychotherapy” (R01 MH66070, A. Garland, PI; Garland, Plemmons, & Koontz, 2006). The application’s specific aims were to identify core common components of effective care for youth aged 4 to 13 years with conduct problems, based on efficacy research and clinicians’ judgment, and to examine the extent to which delivery of care in community-based outpatient clinics is consistent with these components. In addition, this project was designed to examine how delivery of care consistent and inconsistent with these principles is associated with changes in selected child and family outcomes. Although this project was funded by the NIMH after the first round of review, the reviewers raised a concern that the investigative team did not have sufficient expertise in the area of psychotherapy research. This had not occurred to the team when preparing the original application because the investigative team included multiple researchers who had doctorates in clinical psychology, considerable clinical experience in working with the type of child population to be used as a sample on the proposed field work, and extensive research experience in examining outpatient service delivery to this population. However, the reviewers argued that none of the investigative team had the important experience of conducting NIH-funded psychotherapy research. At the request of the NIMH branch involved in the funding decision, a consultant who met those requirements was added with sufficient resources to make one trip a year to the research site and consult with the team by e-mail and phone.

A second example comes from an application that requested funding to extend a large-scale mental health services research project titled “Patterns of Youth Mental Health Care in Public Service Systems” (U01 MH55282, R. Hough, PI; Burns et al., 2001; Garland, 2001; Garland, Hough, Landsverk, & Brown, 2001; Hough et al., 2002). This interdisciplinary research study examined children and adolescents with or at high risk for significant mental health problems who were involved in five child service systems: social services, mental health, special education programs in the public schools, juvenile justice, and drug and alcohol. The original study included a 2-year longitudinal study of 1,850 children and adolescents aged 6 to 17 years.

The application to extend the study proposed following the youth who had aged out of adolescence and passed their eighteenth birthday to examine the need for and use of mental health care during the transition to adult service systems, which often have quite different eligibility criteria and funding mechanisms. The reviewers of the new proposed study commented that the investigative team did not have sufficient expertise in the transitional-age population or the area of adolescent development. In response, the investigative team located two consultants who agreed to be on the revised application. One consultant had originated the important concept of "emerging adulthood," writing extensively on this specific developmental period. The other consultant had conducted a series of NIH-funded and highly productive longitudinal and developmental studies. Although the subsequent application was not funded, the review indicated that the addition of these two developmental experts well met the concerns from the prior review.

Finally, the research network-based study titled "Improving Care for Children in Child Welfare" (R24 MH67377, J. Landsverk, PI; Barth et al., 2005) had an overarching goal to establish a Child and Adolescent Interdisciplinary Research Network (CAIRN) with the focus on developing a heuristic model and practical strategies for the dissemination, implementation, and maintenance of evidence-based, parent-mediated interventions (E-BPMI) in child welfare settings for treatment of disruptive disorders and externalizing behavior problems in children and adolescents. Note that the sole focus was on externalizing behavior problems. One of the reviewers strongly commented that the investigative team was too narrowly focused on externalizing problems and was ignoring the entire area of trauma and resulting internalizing problems and disorders such as depression, anxiety, and posttraumatic stress disorder. Although the network R24 was funded by the NIMH, the review experience sensitized this researcher to consideration of expertise from other perspectives than the one on which the field research is based to consider a larger scope and also protect an application from the criticism of too-narrow a focus. This is a clear issue in the area of child welfare studies involving children who have experienced child abuse and neglect, an area in which a robust research tradition has formed around the development of trauma-focused mental health interventions.

## Consultants from the Perspective of the Investigative Team

I have argued earlier that consultants are specialized experts with experience and skills not available among the research investigators. Their status on a project typically differs in several ways from that of investigators. First, consultants generally are paid on a fee basis calculated in terms of number of days agreed upon by the consultant and the PI. In contrast, investigators generally are paid as a percent of time specified for each year of the full budget period. Investigators with university appointments often use salary resources allocated for them on the project to pay for a percent of their salary, sometimes buying back a part of their teaching responsibilities. This option is almost never used to pay for the services of a consultant.

Second, the two roles differ regarding intellectual capital related to the work of the project. Investigators share in the intellectual capital and should have clearly defined rights for publications based on the data and findings from the study. These publication rights are allocated either implicitly or explicitly on the basis of the investigators' intellectual contributions to the project. Consultants rarely are seen as having intellectual rights to use data and findings from the project, although they certainly can be invited to contribute to the project publications. However, contributions are by invitation rather than the social contract implicit in the investigator role.

Third, consultants serve "at the pleasure" of the project. Their work requires an invitation after the investigators have procured the research funding, and the time commitment often is negotiated across different funding years between the PI and the consultant. In many cases in which a research funding award is accompanied by a cut in the budget, consultant time is severely constricted or reduced entirely.

An implicit assumption in the above-mentioned discussion is that consultants all have advanced degrees and significant research experience. Although that is most often the case, some circumstances require the use of an experienced consultant who has neither an advanced degree nor research experience but instead has specialized expertise from nonresearch environments. For example, our research team has had projects in which case review or administrative data from a service delivery system have been used to address the specific aims of a project. Credible expertise for consultation on this type of project would include personnel who have managed these types of data sources for the service system. In this situation, consultation would be sought from data management personnel in the case of administrative data and from quality assurance personnel in the case of clinical chart review data.

An additional specialized role that would use the consulting format is an advisory group. Advisory groups could be scientific advisory boards, which would contain persons with advanced degrees and specialized academic research experience. Another form is a community advisory group, which would contain persons with a variety of academic backgrounds and specialized expertise in the local community. In mental health services research, developing an oversight panel to advise and monitor the work of the project from the perspective of stakeholder roles often is critical. One example is a project examining the use of mental health services in foster care, in which an advisory group was formed from stakeholder groups such as caseworkers, supervisors, program managers, and consumers (parents or adolescents who are or have been involved in the child welfare service system; Garland et al., 2006).

## Special Issues in Selecting and Approaching Potential Consultants

As previously noted, specialized consultants have advanced degrees and critical expertise well documented by publications in the professional literature as well as grant and contract funding obtained from competitive resources such as NIH. These potential consultants, however, may be members of the Initial Review Group (IRG) that would review the proposed application. Obviously, a person on such an IRG cannot review an application in which he or she has had a role in developing. Therefore, the PI, who typically makes the overture to the potential consultant, must carefully consider whether to create conflict for the consultant reviewing the application by securing his or her participation on the project, or not inviting the consultant so that he or she will not be in conflict. The PI certainly should discuss this issue with the potential consultant when first discussing possible participation. Early in the development of my research program addressing mental health care for children in foster care, I had a verbal agreement to mutually remain out of conflict with another expert on children in foster care who often reviewed NIMH applications. This allowed me and the other expert to be available to review the other's application, which was critical because very few experts in this specialized research area were invited to be part of IRGs. A fine point must be made regarding the distinction between discussing issues in a research area and discussing the development of a specific application. Two researchers certainly can discuss issues in their area and avoid conflict for the review process if the discussion does not include anything about the specific application under development.

However, once the discussion involves an application under development, the conflict of interest rule that governs the review process will apply and the potential consultant cannot be a reviewer, even if he or she does not join the specific application in a consulting role.

After the research team has selected potential consultants for the proposed application, members must craft an approach to each consultant to provide the highest probability of successful recruitment. Several elements in the approach have been important, in my experience. First, the person to initiate contact usually should be the investigator who best knows the potential consultant. In instances in which no one in the team has had any contact with the expert, but a member of the team knows someone who has had such contact or even a working relationship in the past, that third party could make an introduction. For example, an e-mail could be sent to the potential consultant regarding the importance of the research being developed, the expertise of the research team, and forthcoming contact by the PI regarding a possible role for the consultant.

Second, the research team needs to be clear about the specific expertise needed and the nature of the consultation being considered. Some or all of the proposal draft that has been developed should be shared—certainly the specific aims section at a minimum. The concrete details of the proposed consulting arrangement should also be stated, such as time and travel required and the fee being offered if the application is funded. In addition, discussion is important regarding whether the consultant is being asked to write and/or review any portion of the application and whether payment for that preparatory work is being offered. Finally, the approach should specify whether the consultant will be asked to participate as a co-author on papers that come out of the project.

A third element that may be critical in the negotiation is the travel burden. Many potential consultants have extraordinary workloads and many travel obligations. Additional travel to the site of the proposed project may be a powerful disincentive to the consultant agreeing to participate. I have found that this issue sometimes can be addressed by approaching travel as a bidirectional process. A recent successful approach to a critical expert for an application was to travel to the home city of the expert instead of asking the expert to travel to the author's city. This was especially important for gaining extensive consultation for the development of the grant application. Traveling to the consultant's home city twice in the development of the application gained almost two full days of rich consultation. In addition, because writing also is a time-consuming enterprise, I offered to prepare a draft of the application's section that needed input from the expert. This was received gladly and resulted in the consultant extensively editing the draft.

Finally, successful negotiation for a consultant's services, both before and after the award, depends on sensitivity to the potential *quid pro quo* of the relationship. Important reciprocities often may be offered or become available. For example, the consultant may wish to learn about the special expertise of the research team that could be useful in the consultant's own applications. Or the consultant may wish to use a member of the research team in professional events important to the consultant. For example, I have participated in workshops that an ethics consultant organized because of my special expertise in the ethical and human subject challenges experienced in research on foster care populations. Another example is responding positively to invited participation in a conference organized by a statistics consultant who believed that my experience in research with child welfare populations would be important to provide real-world examples for statistics experts at the conference.

After an application is awarded, the consultants should be contacted with the good news and further discussions should be held about their roles and any reciprocity that may have developed. Obviously, the full draft of the funded application should be provided to the consultant along with the summary statement of the review critique. If budget adjustments have been made in the award, discussion should also include any adjustments to the consultant's potential fees. The research team should almost always arrange a visit by the consultant to the research site so that he or she can be brought onto the project in an integral way; this way the team gains the benefit of the expert's experience right from the start.

Consultation from the perspective of the consultant also is important to discuss. The following section includes lessons learned from the author's experience as a consultant to many other research teams and projects.

## Consulting from the Perspective of the Consultant

The motivation to provide consultation to other research teams is based on several elements and is almost never driven by fees. I decided to consult with other research teams for potential reciprocities, research field development reasons, and intellectual stimulation. Taking a research consulting role almost always means that the consultant can call on members of that team to consult on his or her own applications and research projects. In addition, I work in the general area of mental health services research, a field that is quite new and growing. Therefore, I have agreed to consult in many instances because it is an opportunity to give back to the field and stimulate greater growth for the field of research. Finally, I have taken on consulting roles because they offer a great opportunity to learn about aspects of research phenomena for which I have little experience. For example, I have consulted extensively for a team of intervention developers at the Oregon Social Learning Center who continue to develop and test treatments appropriate for children involved in the child welfare system. This has provided great insight into intervention development and led to extensive collaborative research between my team of mental health services researchers and intervention developers, including a just completed Child and Adolescent Interdisciplinary Research Network funded by the NIMH under an R24 mechanism (Barth et al., 2005).

In the same way that the research team selects specific consultants, the consultant carefully evaluates the invited opportunity to participate in a specific project and with a specific research team. I have turned down a number of opportunities to consult after such an evaluation. A first-order reason not to join as a consultant is because the proposed project has little chance of getting funded or requires changes that the research team may not want to make. For example, the research team may have made a decision not to use a randomized design when the proposed consultant believed that the science would best be served by such a rigorous design and that the targeted review group would deem a nonrandomized design as inadequate. I have always further evaluated the proposed research in terms of personal intellectual stimulation and the opportunity to develop the field of research into the mental health care of high-risk children.

The performance of any type of consulting role involves a professional and often social relationship. This means that the consulting is always about the research project and the needs of the research team rather than about the advancement of the consultant or the demonstration of the consultant's greater expertise in specialized areas. The consultant is there to serve, with a special emphasis on providing what the research application and project need. In my experience, this means that pejorative language and arrogance never have a place in the consulting role. No matter how the consultant may view the quality of the design and plans for the proposed project, the advice given should be constructive by pointing out alternative approaches, what is known in the extant literature, and possible negative or less-than-optimal outcomes that might result from the design selected by the research team. Clear and strong opinions are expected from the consultant, but delivering those opinions with harsh or judgmental language will never advance the proposed research project.

## Conclusion

Modern social science and public health research increasingly is interdisciplinary in character and requires specialized expertise by both the members of the research team and consultants. Research of this type also requires the capacity to bridge these highly specialized areas in an integral way that breaks down their often disparate and strong boundaries. Consultant roles are especially important as a way to add these capacities to already talented and productive research teams.

---

### POINTS TO REMEMBER

- Increasing use of large-scale research projects by interdisciplinary research teams
  - Federally funded research by investigators from multiple disciplines.
  - Concerted effort to move research toward translational forms.
  - Divisions between scientific disciplines bridged to speed scientific discovery.
- Planning for interdisciplinary research
  - Three types of research leading to improvements in public health.
    - Basic research: informs the development of clinical interventions (e.g., biochemistry, neurosciences).
    - Treatment development: crafts interventions and tests controlled trials.
    - Service system and implementation research.

  - Two translation phases for moving from basic science to improvements in health care.
    - Bring together interdisciplinary teams which cross and integrate basic sciences and treatment development science.
    - Translate evidence-based treatments into service delivery settings.
- Complex teams need to be formed and sustained
  - Types of projects needing teams.
    - Investigator-initiated single research projects.
    - Research program grants.
    - Research centers.
    - Research networks.
- Create interdisciplinary research teams in the proposal development stage
  - Document in proposal.
    - Background and significance section.
    - Preliminary studies section.
  - Interdisciplinary team divided into two types of research roles.
    - Investigators.
    - Consultants.
- Expertise in specialized areas
  - Substantive expertise.
  - Methodological expertise.
  - Specialized investigators.
    - Statisticians.
    - Anthropologists.
    - Economists.
    - Ethicists.
- Identifying and using consultants
  - Expertise provided in short but strategic doses during critical phases.
    - Content specialization.
    - Experience working in competitive research with good track records.
- Consultants from the perspective of the investigative team
  - Consultants generally paid on a fee basis, while investigators generally paid as a percent of time across each year of the budget period.
    - Investigators share in the intellectual capital and rights for publications.
    - Consultants rarely have intellectual rights to use of data.
    - Consultants serve at the pleasure of the project.

        - Consulting format might be in the area of advisory groups.
            - Scientific advisory.
            - Community advisory group.
    - Special issues in selecting and approaching potential consultants.
        - May be members of the Initial Review Group (IRG). so must excuse self.
        - Whoever best knows the consultant should initiate the contact.
        - Research team clear about expertise needed and nature of consultation.
        - Specify whether the consultant may participate as a coauthor.
        - Travel burden.
        - Important reciprocities that may be offered or become available.
- Consulting from the perspective of the consultant
    - Consultant can call upon members of that team to provide consultation.
    - Opportunity to give back to the field and to stimulate greater growth.
    - Opportunity to learn about new aspects of research.
    - Reasons not to say yes.
        - Proposed project evaluated as having little chance of getting funded.
        - Requires changes that the research team may not wish to make.
    - Consulting is about the research project and the needs of the research team, not the advancement of the consultant.
    - Pejorative language and arrogance never have a place.

## REFERENCES

Aarons, G. A., & Palinkas, L. A. (2007). Implementation of evidence-based practice in child welfare: Service provider perspectives. *Administration and policy in mental health and mental health services research, 34*(4), 411–419.

Barth, R. P., Landsverk, J. A., Chamberlain, P., Reid, J., Rolls, J., Hurlburt, et al. (2005). Parent training in child welfare services: Planning for a more evidence based approach to serving biological parents. *Research on Social Work Practice, 15*, 353–371.

Burns, B. J., Landsverk, J., Kelleher, K., Faw, L., Hazen, A. L., & Keeler, G. (2001). Mental health, education, child welfare, and juvenile justice service use by very young offenders. In R. Loeber & D. P. Farrington (Eds.), *Child delinquents: Development, intervention,* *and service needs* (pp. 273–303). Thousand Oaks, CA: Sage Publications.

Culliton, B. (2006). Extracting knowledge from science: A conversation with Elias Zerhouni. *Health Affairs, 25*(3), 94–103.

Emerson, Ralph Waldo. (1899). *Representative Men*. Philadelphia: Henry Altemus.

Garland, A. F., Hough R. L., Landsverk, J. A., & Brown, S. A. (2001). Multi-sector complexity of systems of care for youth with mental health needs. *Children's Services: Social Policy, Research and Practice, 4,* 123–140.

Garland, A. F., Hough, R. L., McCabe, K. M., Yeh, M., Wood, P., & Aarons, G. (2001). Prevalence of psychiatric disorders in youth across five sectors of care. *Journal of the American Academy of Child & Adolescent Psychiatry, 40,* 409–418.

Garland, A. F., Plemmons, D., & Koontz, L. (2006). Research-practice partnership in mental health: Lessons from participants. *Administration and Policy in Mental Health and Mental Health Services Research, 33,* 517–528.

Hoagwood, K., Jensen, P., & Fisher, C. B. (Eds.) (1996). *Ethical issues in research with children and adolescents with mental disorders*. New York: Lawrence Erlbaum.

Hough, R. L., Hazen, A. L., Soriano, F. I., Wood, P., McCabe, K., & Yeh, M. (2002). Mental health services for Latino adolescents with psychiatric disorders. *Psychiatric Services, 53,* 1556–1562.

Putnam, F., Liss, M., & Landsverk, J. (1996). Ethical issues in child maltreatment research with children and adolescents. In K. Hoagwood, P. Jensen, & C. B. Fisher (Eds.), *Ethical issues in research with children and adolescents with mental disorders* (pp. 113–132). New York: Lawrence Erlbaum.

Walton, Izaak, and Charles Cotton. (1960). *The compleat angler*. London, New York: Oxford University Press,.

Westfall, J. M., Mold, J., & Fagnan, L. (2007). Practice-based research—"blue highways" on the NIH Roadmap. *Journal of the American Medical Association, 297*(4), 403–406.

Zerhouni, E. (2003). The NIH roadmap. *Science, 302*(3), 63–72.

Zerhouni, E. (2005). Translational and clinical science—time for a new vision. *New England Journal of Medicine, 353*(15), 1021–1623.

# 8

# "Indigenist" Collaborative Research Efforts in Native American Communities

*Karina L. Walters, Antony Stately, Teresa Evans-Campbell, Jane M. Simoni, Bonnie Duran, Katie Schultz, Erin C. Stanley, Chris Charles, and Deborah Guerrero*

> What a fascinating thing life is! I have survived the many, many stories of how I think, what I know, and who I am—all told by those who are well meaning, well dressed, and well ignorant of the deeper sides of my cultural epistemology.
> —Manulani Aluli Meyer (2001, p. 124)

It was nearly 2:00 AM and we were still cutting potatoes and browning buffalo meat. In 12 hours, we were hosting a kick-off feast for our new research project and were expecting 150 community members and tribal leaders. As we chopped and cooked, one of the team members wondered aloud how many faculty at non-Native projects expended this sort of effort to develop, nurture, and honor community partners. Given the state of research training, very few, we figured. In Native communities,

Acknowledgements: We thank the HONOR Project Native Leadership Circle, community partners, site coordinators, interviewers, and participants at each of the seven sites as well as faculty and affiliates of Nga Pae o Te Maramatanga, University of Auckland.

Note: View expressed in this chapter are those of the authors only and do not necessarily represent those of our funders or collaborators.

the aunties or uncles would be calling if Native researchers did not conduct themselves properly, as this type of personal involvement is expected to nurture meaningful partnerships and, ultimately, achieve the health and healing that our research efforts are all about.

Alas, despite good intentions, research in Native communities sometimes is not beneficial and may even have iatrogenic effects. Native communities and other communities of color have experienced research exploitation and, in some cases, cultural and economic devastation at the hands of even well-intentioned researchers. Developing and sustaining community-based partnerships with Native communities, who have endured colonization, historical trauma, genocide, and racism—as well as histories of exploitation by academics—is difficult, even for researchers from the very communities they are hoping to engage.

This chapter addresses some of the challenges in building community-based research partnerships with indigenous communities, illustrating potential conflicts and possible solutions. To inform these efforts, we introduce eight "indigenist" community-based participatory research (CBPR) principles. We refer to the HONOR Project, a national study of lesbian, gay, bisexual, and transgender Natives to illustrate the process of developing a working partnership. Although much of this chapter focuses on Native-specific experiences, the lessons translate with some specific tailoring to other racial/ethnic minority groups as well.

## Research in Indian Country

Indigenous communities are reclaiming rights to their own knowledge production and to science, which has been part of their communities for millennia (Castellano, 2004; Holkup, Tripp-Reimer, Salois, & Weinert, 2004; James, 2001; Meyer, 2001; Norton & Manson, 1996; Tuhiwai Smith, 2005). The advanced engineering projects of the Incans in South America and the sophisticated agricultural systems of the Haudenosaunee in North America suggest that indigenous peoples have long employed sophisticated techniques honed by research methods. For example, the ancient city of Cahokia in Illinois was an urban metropolis from the eleventh to twelfth centuries, with an estimated population of 20,000 to 50,000 (larger than London or Paris at that time). It contained sophisticated pyramidal structures, one of which (the Monks Mound) was the third largest structure in the Americas (Nader, 2001).

Part of the colonization process is to render invisible the successes of indigenous science and knowledge while simultaneously infusing public discourse with images of Indians as intellectually inferior. For example, the ancient mound structures were considered too sophisticated to have been produced by indigenous populations (Nader, 2001). Lay persons as well as anthropologists attributed the mounds to others, including the Vikings, Chinese, Lost Tribes of Israel, and lost civilization of Atlantis (www.answers.com/topic/mound-builders-2). Prevailing anthropological theories in the nineteenth century postulated that the mound builders had died off or had been annihilated by barbaric Indian tribes.

More recently, the media, in books and films such as *Chariots of the Gods*, have gone as far as to suggest that similarly complex architectural structures must have been constructed by aliens from other planets, presumably because Natives could not possibly have built them (Von Däniken, 1968). As Maori scholar Linda Tuhiwai Smith (2005) noted,

> The Western academy which claims theory as thoroughly Western…has constructed all the rules by which the indigenous world has been theorized…[as a result] indigenous voices have been overwhelmingly silenced. The act, let alone the art and science of theorizing our own existence and realities, is not something which many indigenous people assume is possible. (p. 29)

For indigenous peoples, therefore, decolonizing research methods include deconstructing and externalizing the myth of the intellectually inferior Indian, while simultaneously privileging and centering indigenous worldviews and knowledge to promote revitalization of indigenous epistemologies, research practices, and ultimately, indigenous wellness practices.

This call for revitalization and innovation in indigenous science follows the egregious mistreatment of indigenous peoples over the course of modern science. Indigenous peoples have endured generations of colonialism in the form of medical impropriety, abusive experimentation, and lack of protection of human subjects (Lawrence, 2000; Smith, 2006; Udel, 2001). Historically, indigenous peoples have been treated as scientific objects with scant regard to community needs or the potentially harmful implications of research processes and findings; research that communities have deemed as "helicopter" or "drive-by" research. Notably, medical impropriety and experimentation often have targeted the most vulnerable of our people, our children. For example, a report by the Truth Commission into Genocide in Canada (2001) documented the deaths of 50,000 Native children in Canadian boarding schools and detailed numerous instances of medical experimentation, including the removal of organs and radiation exposure.

Another area of exploitation causing concern for indigenous peoples is the seemingly relentless campaign to carry out genetic research on indigenous peoples. Genetic research all too often has been conducted without the approval of indigenous subjects. Researchers, for instance, have taken blood samples from earlier health studies and used them to carry out genetic research without tribal consent or consultation. An infamous example occurred recently among the Nuu-chuh-nulth people in British Columbia and is detailed at length by Schmidt (2001). As he reported, in the late 1980s a study was conducted to investigate the high incidence of arthritis in the Nuu-chuh-nulth community. The lead researcher collected 833 vials of blood from donors who signed consent forms allowing for the screening of biomarkers related to arthritis. The researcher was not able to find a gene related to arthritis but soon after used the samples to conduct further genetic research and shared them with other researchers without tribal knowledge (Schmidt, 2001). Such violations continue to occur, sometimes with federal government approval. For example, Schmidt also reported that the U.S. government filed patents on DNA cells taken from the Hagahai tribe in Papua New Guinea and the Solomon Islands. These samples were taken without informing participants or the tribe of the study. Alarmingly, the Hagahai cells are now available for purchase from a biomedical company. A more recently publicized case of harmful research involved the Havasupai Tribe who, in 2004, filed a lawsuit against researchers at Arizona State University (ASU) for misusing blood samples taken from tribal members. Specifically, tribal members were told that their blood samples would be used only for a study on the genetics of diabetes. However, their blood samples were also used for studies on schizophrenia, inbreeding, and migration studies of their ancestors. Tribal members reported that the published data from these studies were "humiliating and harmful to them" (Sahota, 2007). A major result of this particular case was tribal mobilization throughout Indian Country, including by the National Congress of American Indians (NCAI), to enact policies to protect tribes from research harm as well as increase tribal control over research in Indian Country (Sahota, 2007).

Being "researched to death" is both metaphor and reality for many indigenous communities (Castellano, 2004). Marlene Brant Castellano, while chairing a research session at the 1992 Royal Commission on Aboriginal Peoples, observed an elder who stated "If we have been researched to death, maybe it's time we started researching ourselves back to life" (p. 98). In response to this exploitation under the guise of scientific inquiry, indigenous communities are demanding accountability, in some places developing their own institutional review boards as well as guiding principles and protocols for all phases of the research process (Wallerstein & Duran, 2006). Research—by and for Natives—has prompted pipeline initiatives among indigenous communities (e.g., kaupapa Maori research) and universities to streamline indigenous scholars into research careers. As Linda Tuhiwai Smith (2005) noted

> Research, like schooling, once the tool of colonization and oppression, is very gradually coming to be seen as a potential means to reclaim languages, histories, and knowledge—to find solutions to the negative impacts of colonialism and to give voice to an alternative way of knowing and being. (p. 91)

The lack of indigenous individuals trained in conducting research has necessitated the development of partnerships with nonindigenous scientists. CBPR approaches as well as participatory action research approaches have guided some of these collaborations.

## Community-Based Participatory Research

> Research is not just a highly moral and civilized search for knowledge; it is a set of very human activities that reproduce particular social relations of power (p. 88).
> —Linda Tuhiwai Smith (2005)

CBPR has emerged over the last few decades as an alternative research paradigm that integrates education and social action to improve health and reduce health disparities. Wallerstein and Duran (2007, p. 312) noted that "CBPR is an orientation to research that focuses on relationships between academic and community partners with principles of co-learning, mutual benefit, and long-term commitment and incorporates community theories, participation, and practices into the research efforts."

CBPR is a way of approaching research that is consistent with social justice and, in the case of indigenous communities, tribal sovereignty. Action research, of which CBPR is one form, includes as well participatory action research, feminist participatory research, and cooperative inquiry (Holkup et al., 2004; Minkler, 2004). The approaches to CBPR are related to two traditions—the action research school promulgated by Kurt Lewin in the 1940s and the liberatory and consciousness-raising approaches of South and Central American scholar-activists such as Paulo Friere. Minkler (2004) noted that Friere, Fals-Borda, and other developing world scholars formulated "their revolutionary approaches to inquiry as a direct counter to the often 'colonizing' nature of research to which oppressed communities were subjected" (p. 686). Israel and colleagues (2001) defined CBPR as focusing on

> ...social, structural, and physical environmental inequities through active involvement of community members, organizational representatives, and researchers in all aspects of the research process. Partners contribute their expertise to enhance understanding of a given phenomenon and integrate the knowledge gained with action to benefit the community involved. (p. 182)

Moreover, CBPR, as an orientation to research, focuses on relationships among research partners with goals of societal and communal transformation (Minkler & Wallerstein, 2003) rather than a specified set of methods or techniques (Wallerstein & Duran, 2006).

Generally accepted CBPR principles recognize the community as a unit of identity and/or analysis; build on the strengths, resiliency, and resources of the community; facilitate co-learning, co-partnering, and community-capacity building throughout all phases of the research project, including dissemination; attempt to strike a balance between research and action; emphasize local relevance and ecological and historical contexts that contribute to multiple determinants; generate systems growth through cyclical and iterative processes; and involve long-term commitment to process and community (Israel et al., 2001; Wallerstein & Duran, 2006). Tribes have added to the list of CBPR principles some of the following mandates: (1) don't plan about us without us; (2) all tribal systems shall be respected and honored, emphasizing policy building and bridging, not a policy wall; (3) policies shall not bypass tribal government review and approval before implementation; and (4) tribally specific data shall not be published without tribal authorization (Turning Point Collaboration, Robert Wood Johnson Foundation, 2001). Explicit throughout CBPR implementation in indigenous communities is the recognition of the sovereignty of the tribe or indigenous community to be self-determining; that power and authority rest with the community or tribal entity; and that the process of knowledge exchange is reciprocal and always attentive to the best interests of the indigenous community, the ancestors, and future generations.

Authentic CBPR practice means expanding partnership-building dynamics to include the unpacking of power and privilege, specifically around the areas of racism, ethnic discrimination, and internalization of Western science as the only relevant form of scientific inquiry. For nonindigenous researchers partnering with indigenous communities, we caution that either partner's assumption of academic expertise may unintentionally hide or silence local voices, overriding traditional understanding of local phenomena. In addition, silence can be mistaken or coercively interpreted as agreement when, in fact, it represents resistance.

To facilitate partnership building and to reformat power dynamics, we promote a stance of cultural humility as opposed to cultural competence (Wallerstein & Duran, 2006). Cultural competence might never be acquired, and it may not even be an appropriate goal (e.g., in the case of spiritual protocols). Cultural humility, on the other hand, refers to a life-long commitment to critical self-evaluation regarding multiple, complex, and simultaneous positions of unearned privilege (e.g., being white and male) to redress power imbalances and nurture deeply respectful partnerships with communities.

## Decolonizing Research

> If we, as Indian people, are forced to reject our own indigenous knowledge and our ways of thought to participate in science, then we will be that much closer to cultural extinction.
> —Cornel Pewewardy (2001, p. 21)

As previously noted, the movement toward developing decolonizing methodologies among indigenous communities is a global struggle to counter hegemony. Indigenous researchers have been actively seeking protocols that disrupt and counter the history of exploitation, trauma, and discrimination in research inquiry and, instead, "privilege indigenous knowledges, voices, experiences, reflections, and analyses of their social, material, and spiritual conditions" (Rigney, 1999, p. 117).

As Tuhiwai Smith (2005) noted, this type of social movement to counter oppressive research structures is not new to research communities. Other populations, such as sexual minority communities, women, and other ethnic minorities, also have experienced dissatisfying and even exploitative research partnerships. They too are challenging the epistemological basis of scientific paradigms and the relevance to their communities.

For indigenous peoples, a layer is added to countering the hegemonic research imperatives; namely, to decolonize the research process while simultaneously indigenizing it. Tuhiwai Smith (2005) wrote that the decolonization research process involves multiple layers of struggle across multiple sites, including the unmasking and deconstruction of imperialism; manifestations of old and new formations of colonialism; as well as the simultaneous recognition of sovereignty for reclamation of indigenous knowledges, languages, and cultures and, ultimately, for the transformation of colonial relations between the colonizer and the colonized. She explained:

> Decolonizing research, then, is not simply about challenging or making refinements…to research. It is a much broader but still purposeful agenda for transforming the institution of research, the deep underlying structures and taken-for-granted ways of organizing, conducting, and disseminating research knowledge. (p. 88)

Decolonizing research practices promotes the ability of indigenous peoples and researchers to theorize their own lives; reconnect with past and future generations; acknowledge and prioritize indigenous ways of knowing and healing; respect and prioritize the community's role in defining problems, resiliencies, and strategies; and cultivate and build indigenous capacity to engage in both indigenous as well as Western research methodologies.

## Indigenist Research Principles

It is our conviction that it is not sufficient to decolonize research; we must go further and indigenize research. To guide the development of mutually beneficial research partnerships with indigenous communities, we suggest eight principles for decolonizing and indigenizing research: reflection, respect, relevance, resilience, reciprocity, responsibility, retraditionalization, and revolution. Specific strategies for incorporating each principle in research partnerships are discussed subsequently. Note that many of these principles incorporate Tuhiwai Smith's (2005) groundbreaking work on building indigenous research capacity. These guidelines are not exhaustive and should be appropriately tailored. They are a starting point, aimed to instigate further co-exploration of decolonizing and indigenizing approaches in research partnerships with indigenous communities (Figure 8.1).

Before engaging with Native communities, research partners could benefit from careful **reflection** upon their positionality vis-à-vis community members. Most university-based research partners have a

***Reflection***
True partnerships begin with reflection upon the privileged statuses from which most partners operate and the emotionally difficult task of acknowledging the pain of Native communities and developing empathy.

***Respect***
Research partners must value and prioritize indigenous epistemologies, knowledge, cultural protocols, and healing practices.

***Relevance***
The community should contribute to defining research problems and strategies, which should respond to their own self-identified needs and concerns.

***Resilience***
All aspects of the research must acknowledge the community's strengths and resilience.

***Reciprocity***
The partnership should be collaborative and mutually respectful with knowledge exchanged in both directions.

***Responsibility***
Research partners are obliged to enhance community capacity to conduct Indigenous and Western research, disseminate research findings in culturally meaningful ways, and anticipate the implications.

***Retraditionalization***
Traditional knowledge and methods must be actively integrated into the formulation of the research questions and the process of scientific inquiry.

***Revolution***
Research partners and community members must actively seek to decolonize and indigenize the research process to transform science as well as themselves, their communities, and the larger society for the betterment of all.

**Fig. 8.1** Guiding principles for decolonizing and indigenizing research.

privileged status in society, owing to educational and socioeconomic advantages. If they are members of the dominant racial group in the United States, they have additional advantages based on their White privilege as well. Acknowledging these privileges—not disingenuously denying them—can improve the partnership. Reflection involves an ongoing process of self-awareness of emotional reactions as well. The scale of human misery that Native peoples have endured since colonization and that many continue to confront on a daily basis is difficult to comprehend and accept among non-Natives. A common initial reaction is simply to reject it ("It isn't that bad") or to fight it with individual exceptions ("I know a rich Indian living very well"). Rejection of the experience precludes empathy. Devoid of context, Native calls for justice and inclusion might be misjudged as inappropriately angry or strident, even militant.

The principle of **respect** means that research partners must value and prioritize indigenous epistemologies, knowledge, cultural protocols, and healing practices. Indigenous "scientists" and expert knowledge already exist within indigenous communities and should be involved throughout the research partnership. As one medicine man once said to an academic when they were walking together in the woods on tribal lands, "Professor, out here you are my student. Welcome to *my* university."

For research partners to achieve **relevance**, they must actively engage the community from the earliest phases of the research endeavor in conceiving the aims of the project. This might involve meetings with key community members and tribal leaders, community forums and feasts, and extensive outreach to determine what the community itself defines as important to its health and well-being. For example, one young white student was strongly motivated to pursue research in eating disorders such as anorexia nervosa but had trouble gaining access to youth in the local tribal community. She was eventually forced to acknowledge that her own interests did not match the priorities of the tribe. Relevance also extends to the methodology of the research. For example, many tribes understandably balk at participating in trials with no-treatment control conditions. Given the extensive health needs of most communities, designs that involve waiting-list controls or comparison conditions of interventions with equivalent time and attention are more desirable. Of course, these should be developed with the needs of the indigenous community at the forefront.

Research with indigenous communities should acknowledge the community's strengths and its stalwart **resilience in** the face of multiple assaults on tribal autonomy and integrity. Much of the early work in Indian country focused exclusively on pathologic conditions, such as alcohol addiction and childhood abuse. Although these topics were serious concerns for tribes, the research was conducted without regard to contextual, structural, and historical factors that contributed to these problems (Walters & Simoni, 2002; Walters, Simoni, & Evans-Campbell, 2002) or to the large majority of tribal members who avoided these problems.

**Reciprocity** should characterize the research partnership, which should be collaborative and mutually respectful, with knowledge exchanged in both directions. Western and indigenous knowledge should be mutually understood, and respectfully exchanged. Often, communities have excellent ideas and possible solutions and just require some assistance in formulating these into research questions and translating them into fundable proposals.

Collaborators with indigenous communities assume a grave **responsibility**. They are obligated, first, to enhance community capacity to conduct research. Research endeavors should seek to incorporate youth and students into research activities to stimulate their interest in research and provide experiences to bolster their opportunities for future training. Researchers funded by the National Institutes of Health (NIH) have access to training mechanisms (e.g., F31, minority supplements) to build research capacity within indigenous communities. Creating an indigenous research workforce will help to replace the "Indian experts" with "expert Indians," observed Beverly Pigman, Navajo Nation Institutional Review Board Chair. These newly minted indigenous scholars will need research infrastructures that support and nurture their work; establishing indigenous research organizations is one way of ensuring they find what they need. Second, research partners have the responsibility of disseminating research findings in culturally meaningful ways. This can mean publishing in tribal newsletters as well as peer-reviewed journals, with community partners acknowledged as co-authors according to their contributions. The dissemination process should involve other forms as well, such as digital storytelling, documentaries, photography and other visual presentation, theater, or community events in which the findings are reported in an accessible fashion. Finally, research partners must be responsible for anticipating the implications of their findings. In one infamous example, the financial integrity of a tribe was threatened when questionable research findings were released to the press—without the prior knowledge or approval of the tribe. In addition, work that fails to adequately consider the contextual factors and the history of colonization among indigenous communities may lead to reports that inaccurately blame the victim, switching the focus away from the need for policy changes that address structural inequities.

**Retraditionalization** involves incorporating traditional and ancestral knowledge and methods into the formulation of research questions and the process of scientific inquiry. Building on the principles of respect and relevance, it involves the practice of co-embracing hybridized methodologies while maintaining an indigenous core. Whenever possible, partners and communities should co-develop mechanisms for developing innovative indigenous methodologies or hybrid methodologies that combine Western and indigenous approaches. Retraditionalization is a way to indigenize science. Reframing scientific processes from indigenous worldviews or, rather, reclaiming these processes involves consciously shifting from Western approaches to more holistic approaches in research methodologies. In this fashion, research methodologies become tools and the researcher is the vessel or the vehicle through which the tools come to life. Although some of these tools might contain their own spirit, how the researcher relates to the spirit of these research tools is what matters. In essence, the researcher can be witch or healer, practicing bad or good medicine depending on how he or she works with the spirit of these methods.

Truly indigenist research collaborations involve scientific **revolution**. Research partners and community members, by actively seeking to decolonize and indigenize the research process, can transform the structure and nature of knowledge production. This can be facilitated by challenging colonial or racist research practices within institutions (Mihesuah &Wilson, 2004). Decolonizing research practices include holding the researchers accountable as well as the institutions where research takes place. This process might involve promoting indigenous science and knowledge within universities, as well greater awareness of the need to recognize tribal sovereignty and treaty obligations. Indigenous peoples have established histories as astronomers, engineers, mathematicians, and physicians and, as a result, "science" and the processes of observing, developing, and testing hypotheses are not new to indigenous peoples. Through our research efforts we must continue to debunk, demystify, and deconstruct the intellectually inferior Native mythology. Indigenist science has implications for how indigenous communities are perceived and esteemed. This, in turn, has implications for the well-being of indigenous communities as well as the larger society. More importantly, indigenist science recognizes the contributions that indigenous knowledge has on the world. For example, many foods come from indigenous agricultural knowledge. Potatoes, chocolate, corn, beans, squash, and tomatoes, among other foods, originated in the Americas under sophisticated agronomist methods of planting, sowing, and harvesting (Mt. Pleasant, 2001). Yet, colonization erased knowledge of indigenous origins of these foods (Churchill, 1996). Potatoes became "Irish potatoes"; chocolate became "Dutch chocolate"; and vanilla became "French vanilla."

To paraphrase Aboriginal scholar Karen Martin, indigenist research must decolonize Western research practices via reclaiming the research endeavor (Martin, 2001). Specifically, she states,

> To reclaim research is to take control of our lives and our lands to benefit us in issues of importance for our self-determination. It is to liberate and emancipate by decolonisation and privileging the voices, experiences and lives of Aboriginal people and Aboriginal lands so that research frameworks are reflective of this. (2001, p. 2)

Moreover, she notes that indigenist research emphasizes the social, historical, and political realities that contemporarily shape indigenous lives and futures. Indigenist research, Martin notes,

> ...is undeniably political, emancipatory and confirming in its aim to control research on Aboriginal lands and regarding Aboriginal people and to regulate relations with governments, resource agencies, research institutions and visitors. (2001)

Indigenist research processes also involve redistributing and checking power. The *product* must be recognized as being always in the *process* of how and what we do. To honor ourselves and our partners we must "walk the talk" and be held accountable. Engaging in dialogue between indigenous and nonindigenous researchers and communities is essential to healing.

Accountability mechanisms include integrating elders, youth, and wise leaders into research efforts, for example, by creating leadership councils of community members (including all segments of the target population). Our efforts will not be measured by how many papers we have published, but by how much we have truly assisted in eradicating the suffering, injustices, and health disparities of our partnering communities.

## Implementing Indigenist Community-Based Participatory Research—The Honor Project

> In our every deliberation we must consider the impact of our decisions on the next seven generations.
> —From The Great Law of the Iroquois Confederacy

Most of the authors of this chapter were involved in a 5-year National Institute of Mental Health (NIMH) funded study called the HONOR (Honor Our Nations, Our Relations) Project. Discussed here are examples from our work on this study that illustrate the advantages—and the challenges—of indigenizing the research process.

After identifying a new NIMH funding priority in "Behavioral, Social, Mental Health and Substance Abuse Research with Diverse Populations," the first author contacted several indigenous health and HIV programs in major urban centers where HIV/AIDS surveillance data indicated disproportionately high HIV/AIDS rates, specifically among indigenous men who have sex with men. Although in most CBPR approaches the community identifies the research and researchers with whom to partner, the first author, who is American Indian and two-spirit, believed it was worth the effort to contact indigenous agencies about the announcement to determine interest in a collaboration.

Overwhelmingly, urban indigenous organizations were supportive and signed on. In fact, only budgetary restrictions limited our ability to partner with several rural, reservation, and Native Hawaiian partners who also expressed interest. Five agency directors were contacted and all five agreed to submission of the grant, eventually entitled "Health Survey of Two-Spirit Native Americans" in 2001. We were quite lucky and our first application was funded; thus began our journey in building more depth and breadth to our partnerships.

The HONOR Project had four major aims: to test an indigenist stress and coping model (Walters et al., 2002; Walters & Simoni, 2002); to establish preliminary baseline prevalence rates of trauma, HIV/AIDS, and health outcomes; to develop and evaluate an innovative sampling method; and to develop research infrastructure and capacity at the Native agencies. Our survey examined several areas, including traumatic stressors (e.g., historical trauma, microaggressions, traumatic life event); physical health (e.g., HIV risk behaviors, diabetes, human papilloma virus, cancer); mental health (e.g., posttraumatic stress disorder, depression, anxiety, substance use); spiritual health (e.g., traditional health and healing practices); and cultural protective factors (e.g., identity, community involvement, spirituality, enculturation). We imagined the findings could be used to inform service providers of critical health and wellness data so that they could better serve two-spirit (i.e., lesbian, gay, bisexual, or transgender; see Fieland, Walters, & Simoni, 2007, for a thorough explanation of the term) men and women at their agencies and to assist the local two-spirit communities in planning, implementing, and developing health policy. It was important to us that we incorporated research training opportunities for Native community members, students, and research professionals throughout the study. We were quite successful in the latter regard, hiring and training more than 30 American Indian and Alaska Native staff, investigators, and students on this project. Moreover, 60% of the investigators and professional staff self-identified as two-spirit, thus providing strong representation of the population of interest in directing the study. In addition, four American Indian investigators were key personnel (including the Principal Investigator [PI]), making this one of the few NIH-funded studies led by American Indian researchers. We had considerable support from community partners and non-Native research co-investigators and allies as well.

Participants in the HONOR Project were recruited from seven metropolitan areas in the U.S.: Seattle-Tacoma (Northwest Two-Spirit Society); San Francisco-Oakland (National Native American AIDS Prevention Center with additional support from Bay Area Two-Spirits); Los Angeles (United American Indian Involvement); Denver (Two-Spirit Society of Denver); Oklahoma City-Tulsa (Oklahoma City Native American AIDS Coalition, Indian Health and Community Resources Center, and Tulsa Two-Spirit Society); Minneapolis-St. Paul (Indigenous Peoples Task Force); and New York City (American Indian Community House and Northeast Two-Spirit Society). Eligibility criteria were (1) self-identifying as American Indian, Alaska Native, or First Nation; (2) being enrolled in a tribal nation or having at least 25% American

Indian blood; (3) self-identifying as gay, lesbian, bisexual, transgender, or two-spirit *or* engagement in same-sex sexual behaviors in the past 12 months; (4) being age 18 years or older; (4) English speaking; and (5) residing, working, socializing in the main city of the particular study site. Participants were recruited by two methods: targeted sampling and a modified respondent-driven sampling technique. A total of 452 participants were recruited in the survey in less than 2 years. In addition, 65 community-identified two-spirit leaders completed an extensive qualitative interview identifying wellness and resiliency themes among two-spirit persons. All participants received monetary remuneration for their participation.

### *Forming Team Partnerships across Differences*

Once the grant was awarded, we needed to nurture the relationships within our diverse interdisciplinary team of indigenous and nonindigenous scientists as well as our community partners and put into place systems of community accountability with the use of national and local community advisory boards. One of the first steps in decolonizing protocols previously noted is the focus on critical self-reflection on our own intentions, research capacities, and limitations—from internalized colonial processes to the positions of power and unearned privilege that indigenous and nonindigenous investigators, staff, and students held. We encountered and processed within our Seattle site our struggles with different worldviews, epistemologies, and methodological approaches. Not surprisingly, initial struggles emerged between indigenous and nonindigenous researchers in terms of comfort with different orientations to time, outcomes, and processes of engagement within the team itself. For example, for the nonindigenous researchers, the entrée into the field work phase was incredibly slow, involving high levels of staff time (i.e., money spent with little observable "outcome") as well as indigenous investigator time in community events, activities, and ceremonies that to nonindigenous researchers appeared to be only remotely related to the study aims. The PI and other indigenous researchers dialogued with the nonindigenous research personnel about the importance of community commitment, visibility, and genuine interest in all the communities' wellness as part of the development of trust in us. On the other hand, the indigenous researchers and staff struggled with nonindigenous researchers' push to move too quickly into the field. They also interpreted cultural differences in communication styles (e.g., being more verbally direct) as chafing and sometimes disrespectful of indigenous values and protocols. Through dialogue and sharing of differences in worldviews, expectations, and protocols, the research team began to work through their differences—good-natured teasing of nonindigenous researchers coupled with the researchers' demonstrated cultural humility quite often diffused tensions. Most importantly, however, indigenous and nonindigenous investigators became united in their intentions for the best interests of the community. One team member remarked at one point, "People are dying out there." His rallying call helped to refocus the team, who decided they needed to use all the available resources, on indigenous terms, to better the health of the Native communities. In addition, over time and through working together on the team, the focus on the research tools (epistemological, theoretical, and methodological) that each person brought to bear on the research partnership became less threatening once the relationship to those tools became was clear—that is, they were shown as being used only in the best interests of the community.

### *Tensions between the University and Community Emerge*

Similar to other studies, many tensions existed between university and community expectations, especially regarding the necessity of bureaucratic process around payment mechanisms. For example, many community members who worked directly with the project as consultants, transcriptionists, and artists had limited incomes and had to wait inordinately long periods to receive payment. This was particularly challenging because the Native staff understood the financial vulnerability of community members who were accustomed to being paid at the time services were rendered instead of waiting weeks and, in many cases, months to be paid. Moreover, many of our community members lacked sufficient savings to tide them over, thus increasing their economic hardship. Not only did this reinforce cultural mistrust among community members regarding the exploitation of Native peoples within university systems, but it also created unnecessary financial hardships on some of our most vulnerable community members. Thus, timing of payment schedules and reimbursements, and invoices versus payment at time of service all lead to some very challenging moments. We have not resolved this tension as of yet because, in large part, this involves shifting university financial systems, which takes tremendous resources and effort—but we are engaged and hope to shift policies to be more inclusive of community members who have much to offer in the ways of their services to research endeavors.

In addition, from the beginning of the study we had to attend to high levels of community cultural mistrust regarding research in Indian Country and university systems in general. Prior community experiences and perceptions of cultural appropriation; abuse of indigenous intellectual property rights; and general exploitation of indigenous staff, personnel, and professors also came to bear in the present project. To address these concerns, we simply realized that we had to "walk the talk" and conduct ourselves honorably and with transparency. Moreover, we relied upon our local and national leadership councils (i.e., community advisory boards) of key two-spirit community-based leaders to guide us and provide key insights into major thematic issues, topical areas, and cultural protocols. Importantly, we made sure that the leadership councils had representation from some of the most marginalized voices in our community (e.g., members who were homeless, transgender, or young people).

### *When the Community Calls, Come Running*

One of the most challenging issues has been the fact that the team largely consists of indigenous researchers, many of whom are two-spirit—thus we were "insiders" to the communities we were partnering with. Of course, this provided advantages (e.g., greater trust and authority delegated to the team by community members). However, being on the inside created other challenges. First, issues arose pertaining to boundaries. Indigenous researchers already carry many burdens and blessings in researching their own communities, some of which include familial or community obligations that come first but might be seen to interfere with research objectivity. For example, sometimes our indigenous staff is called on for ceremonial purposes (e.g., leading sweats, songs) within communities, including two-spirit communities; yet, these are the same communities with whom we are engaged in research. From an indigenous perspective, if someone is asked to give a blessing, sing a song, or pour a sweat and has the authority to do so, then it would be an insult to refuse without some strong spiritual or cultural justification. Another challenge is that sometimes, because of internalized colonization, our own community members devalued the indigenous researchers' knowledge and tended to hold nonindigenous counterparts' knowledge in higher esteem. Finally, to illustrate tensions in competing obligations (research and community), at one point during the study we received a call from an indigenous elder who had heard of our study and wanted to help the two-spirit community. He wanted to bring special gifts to the community but asked if our project could collect and deliver them. After discussion with the team, we decided that in the interest of helping the community, we would assist this elder with funds from our own pockets (ethically and fiscally, we could not use the study budget). A group of us traveled from Seattle to the Idaho border by car and transported the gifts back in the same day. The elder was pleased to have something to offer the community and we had an opportunity to hear about two-spirit people of that person's nation as well as sing together and have a meal together. Some of our nonindigenous colleagues thought we were "crazy" to travel 18 hours in a day to bring these gifts back and to meet with the elder. But the indigenous staff and researchers knew that not only was this culturally expected, but it was required. Now, upon reflection, we recognize that in doing so we learned many things that long day about two-spirit history and perspectives from an elder who wanted to help.

### *Almost Defunded*

In the second year of the study, just when we were about to hire a full-time project director and launch the survey part of our study, we received a disturbing phone call from our project officer. He called to tell the PI that within 2 hours a congressional amendment from Representative Patrick Toomey (H.AMDT.221 to H.R. 260; July 10, 2003, Congressional Inquiry) was about to be introduced that called for defunding several NIH-funded projects, including our study. In addressing his congressional colleagues, Toomey stated, "Who thinks this stuff up?...if they want to do this sort of research, we need to fund this privately and not with taxpayer dollars. I simply want to make the point that there are so many far more important, very real diseases that are affecting real people." Representative Chris Chocola went on to argue, "...I do not know that we can identify people who benefit from this taxpayer money being spent on these grants...we should be eradicating these horrible diseases that ruin families, ruin individual lives rather than grants that benefit no one that we can identify." Fortunately, the amendment failed, though narrowly, in a recorded vote of 210 to 212. Needless to say that we were relieved but very concerned as to how the community would respond to such public discussion and visibility of the study. In fact, the project was discussed on major media outlets *The O'Reilly Factor, The Scarborough Report*, and CNN *Headline News* during that week. We also quickly found out that once we were in the public eye, we were vulnerable to more attacks. Many of our investigators, including the PI, had community organizing experience and advocacy experience and, if this attack had been on a more individual level, would perhaps have taken the challenge more publicly. However, given that we were funding more than 30 indigenous people on this project, we had to think more broadly about the pros and cons of the response we would make. At that time, we decided to contact our community partners and all staff and colleagues associated with our project to let them know what had happened. One of our partners became so impassioned that he was able to give a speech and receive a declaration of support on the floor of the United Nations Indigenous Peoples Work Group meeting in Geneva in support of our study. Indigenous communities worldwide rallied in support of the study, noting that an attack on any segment of our community is an attack on all indigenous communities. Non-Native individuals and research organizations came out to support us as well. On the community partner level, we received a few phone calls from some community members who were concerned about the government accessing our records. We then applied for a Federal Certificate of Confidentiality to address this concern. Interestingly, our grant was requested through the Freedom of Information Act by a conservative journalist. We attempted to remove the names of our community partners to protect them but were not allowed to do so unless we wanted to invest our resources in a court battle that our attorneys assured us we would likely lose. Thus, once again, we notified our community partners listed on the application that this journalist would know who they were and we would support them however we could in addressing whatever might come of it given their fiscal vulnerabilities (most are small agencies operating on very tight budgets). Although this was a challenging experience, creating some community apprehension, for the most part our community partners and indigenous communities provided tremendous support. We may even have gained more community respect: reflecting the Native community's historic distrust of the U.S. government, one community member stated, "We must be doing something right if Congress wants to defund us."

### *The Power of Naming*

Researcher typically develop acronyms and pithy names for their studies, sometimes giving little attention to the process of naming. From an indigenous point of view, naming is sacred. In many communities, a name is seen as emanating from the spirit world. Along with a name come responsibilities as to how to conduct oneself and establish one's place in relation to the community. Naming creates relational accountability. At the time we initially discussed the HONOR project, we did not follow indigenous protocols and instead went the more Western route of devising an acronym. The community members discussed the importance of having a positive name, one that honored our two-spirit ancestors of the past and also would refocus on community strengths. Thus, the HONOR project was born—first as an acronym to reflect Honoring Our Nations, Our Relations, and later to simply honor the communities we are working with. From this naming experience, we learned the importance of following indigenous protocols, or indigenizing the research enterprise, and paying attention to the everyday details of research development and events, such as naming, which might be seen as insignificant in the scheme of the grand research project but may be critical to establishing accountability. Moreover, the community may choose to have an actual naming ceremony to provide a spiritual foundation for the study through the name. Of course, this should be done according to protocol that the indigenous community involved sets forth.

### *Feeding the Community*

Consistent with indigenous protocols, our team sought first the permission from the local indigenous tribe and their leader to conduct our study on their land. We offered the tribal leader gifts and tobacco and asked to make a speech at our opening feast to acknowledge the study and the leader's support for our efforts. In this way, we honored the indigenous peoples of that land and territory first and foremost, and then we sought permission to move forward with the study from tribal leaders, the urban indigenous community at large, and the two-spirit community more specifically. To accomplish this, we prepared a feast. Feeding the community is both metaphor and practice as we prepared to share the study with the community, solicit their feedback and guidance, and announce the development of the leadership council. Feeding the community also means feeding our community spirit. This event was well attended, and we were able to get critical support and guidance from the community during the presentation of the research study. In addition, the community partners had an opportunity to meet one another from across the country at this feast, allowing cross-fertilization of knowledge exchange among community members and partners.

Overall in the course of the HONOR Project, we learned several lessons. Specifically, we learned about the importance of self-care of indigenous staff, interviewers, and researchers as we delved into the interview phases of our study. At times, traumatic material invoked a secondary traumatic response among our own personnel; we needed to ground ourselves spiritually, in our own traditions, to deal with difficult stories. As elders have noted, stories have their own spirit, and in this way we as a team began to incorporate more ways to cleanse, center, or spiritually ground ourselves (e.g., smudging the room before team meetings; for those who wanted, team members going to sweat lodges to cleanse). In addition, we realized early on that a considerable investment of resources (financial, personnel, and personal) is required to develop research capacity among our community partners and among our own team members, including indigenous and nonindigenous researchers. Specifically, the learning curve for our own indigenous staff, many of whom had limited research experience, took time to nurture and strengthen. In many instances the work involved having our own team recognize their gifts and knowledge that they already possess as not only being valuable, but also being a necessary asset to the success of any research enterprise.

In general, indigenist CBPR takes many more resources and time to implement properly than would a more traditional research approach. Getting into the community took at least two to three times longer than even the indigenous researchers thought would be necessary. Seemingly little things such as creation of outreach materials and logo took tremendous time and energy but were well worth it in terms of nurturing and supporting indigenous artists and the community to be actively involved in co-creating the images and message for the community outreach materials. For example, the Native two-spirit artist who designed the posters for our outreach materials fortunately provided several designs. After several focus groups with different target populations (i.e., gay-identified

Native men, Native youth, elders, and middle-aged two-spirit activists), it became clear that three different posters needed to be developed to attract specific two-spirit populations to the study. This led to a higher cost than expected for outreach materials but at the same time gave us an opportunity to encourage hard-to-reach populations to enroll or find out more about the study (e.g., closeted gay Native youth) and simultaneously embraced generational cohort differences in values, worldviews, and issues (e.g., the middle-aged group liked the more "historic" poster, whereas the more "out" gay men preferred the image with two men nose to nose and dressed in traditional regalia).

## Conclusion

This overview of how to "indigenize" collaborative research efforts with Native American communities, along with the case example of the HONOR Project, highlight one key reality: research in Native communities demands years of personal and professional commitment. Maintaining consistent contact, fighting for indigenous rights and against injustices, and having an ongoing presence beyond the data and study are critical to nurturing mutually satisfying, liberating, and trusting partnerships with indigenous communities. The opportunity to work in indigenous communities is a great privilege, but it comes with great responsibility.

We are at a crossroads of Western and indigenous knowledge and science—a time for opportunity and growth and also a time to strengthen indigenous knowledge production and traditions through indigenous CBPR. With the recent passage of the United Nations Declaration on the Rights of Indigenous Peoples (United Nations General Assembly, 2007), the time is right to honor indigenous knowledge and intellectual traditions. As the Declaration states

> 1. Indigenous Peoples have the right to maintain, control, protect, and develop their cultural heritage, traditional knowledge, and traditional cultural expressions, as well as the manifestations of their sciences, technologies and cultures, including human and genetic resources, seeds, medicines, knowledge of the properties of fauna and flora, oral traditions, literatures, designs, sports and traditional games, and visual and performing arts. They also have the right to maintain, control, protect, and develop their intellectual property over such cultural heritage, traditional knowledge, and traditional expressions.

2. In conjunction with indigenous peoples, states shall take effective measures to recognize and protect the exercise of these rights. (Article 31, pages 14–15).

Reciprocity must be the defining feature of collaborative research efforts with Native American communities. Although non-Native partners might initially anticipate a sense of loss at the prospect of a truly collaborative partnership (indeed, they will be required to relinquish some control), their efforts eventually will be rewarded. As Oscar Kawageley (2001, p. 55) explained: "There is a need to broadly reconceptualize and revitalize Native knowledge and to integrate it thoroughly with mainstream science. The latter is an absolute must for our own people and for others; we have much to share with them" (p. 55).

---

## POINTS TO REMEMBER

- Indigenist community-based participatory research (ICBPR) principles and decolonizing partnership strategies
  - Research in Indian Country.
    - Demystify, externalize, and deconstruct the "intellectually inferior Indian."
    - Privilege and center indigenous knowledge, worldviews, and science.
    - Promote healthy revitalization of and creation of new indigenous epistemologies.
  - CBPR principles.
    - Recognize the community as a unit of identity and/or analysis.
    - Build on the strengths, resiliency, and resources of the community.
    - Facilitate co-learning, co-partnering, and community-capacity building.
    - Attempt to strike a balance between research and action.
    - Emphasize local relevance and ecological and historical contexts.
    - Generate systems growth through cyclical and iterative processes.
  - CBPR is Action research that includes the following:
    - Participatory action research.
    - Feminist participatory research.
    - Cooperative inquiry.
- CBPR includes the following:
  - Active involvement of community members, organizational representatives, and researchers.

  - Focuses on relationships between and among research partners with goals of societal and communal transformation.
  - Recognizes community as a unit of identity.
  - Builds on the strengths, resiliency, and resources of the community.
  - Facilitates co-learning, co-partnering, and community-capacity building.
  - Recognition of
    - sovereignty of the tribe or indigenous community to be self-determining, power, and authority rests with the community or tribal entity;
    - process of knowledge exchange is reciprocal and always responsible to the best interests of the indigenous community.
- Eight guiding indigenist principles for CBPR
  - Reflection
    - Acknowledge the privileged statuses from which most patners operate
  - Respect
    - Value and prioritize indigenous epistemologies, knowledge, cultural protocols, and healing practices.
  - Relevance
    - Actively engage community from earliest phases and aims.
    - Develop aims with needs of community at forefront.
  - Resilience
    - Acknowledge community's strengths.
  - Reciprocity
    - Collaborative and mutually respectful.
  - Responsibility
    - Enhance community capacity for research.
    - Disseminate results in culturally meaningful ways.
    - Anticipate the implications of findings.
  - Retraditionalization
    - Co-develop mechanisms for developing innovative indigenous methodologies or hybrid methodologies that combine Western and indigenous approaches.
  - Revolution
    - Challenging colonial or racist research practices within institutions.
    - Promoting indigenous science and knowledge within universities.
    - Recognizes the contributions that indigenous knowledge has on the world.

---

## REFERENCES

Castellano, M. B. (2004). Ethics of Aboriginal research. *Journal of Aboriginal Health*, (January), 98–114.

Churchill, W. L. (1996). *From a native son: Selected essays on indigenism, 1985–1995* (pp. 271–294). South End Press, Boston, MA.

Comptroller of the United States (1976). *Investigations of allegations concerning Indian health services*. Washington DC: Government Printing Office, 4 November, 4, 18–20, 25.

Fieland, K. C., Walters, K. L., & Simoni, J. M. (2007). Determinants of health among Two-Spirit American Indians and Alaska Natives. In I. H. Meyer & M. E. Northridge (Eds.), *The health of sexual minorities: Public health perspectives on lesbian, gay, bisexual and transgender populations* (pp. 268–300). New York: Springer.

Israel, B. A., Schulz, A. J., Parker, E. A., & Becker, A. B. (2001). Community-based participatory research: Policy recommendations for promoting a partnership approach in health research. *Education and Health*, *14*, 182–197.

James, K. (Ed.) (2001). *Science and Native American communities: Legacies of pain, visions of promise*. Lincoln, NE: University of Nebraska Press.

Kawageley, O. (2001). Tradition and education: The world made seamless again. In K. James (Ed.), *Science and Native American communities: Legacies of pain, visions of promise* (pp. 51–56). Lincoln, NE: University of Nebraska Press.

Lawrence, J. (2000). The Indian Health Service and the sterilization of Native American women. *The American Indian Quarterly*, *24*(3), 400–419.

Martin, K (2001). *Ways of knowing, ways of being, and ways of doing: Developing a theoretical framework and methods for Indigenous re-search and indigenist research*. Symposium B, Session 1 panel presentation made to the Australian Institute of Aboriginal and Torres Strait Islander Studies conference, 'The Power of Knowledge and the Resonance of Tradition', September, 2001 http://www.aiatsis.gov.au/__data/assets/pdf_file/5718/MARTIN.pdf

Meyer, M. A. (2001). Our own liberation: Reflections on Hawaiian epistemology. *Contemporary Pacific: A Journal of Island Affairs*, *13*(1), 124.

Mihesuah, D. A., & Wilson, A. C. (Eds.). (2004). *Indigenizing the academy: Transforming scholarship and empowering communities*. Lincoln, NE: University of Nebraska Press.

Minkler, M. (2004). Ethical challenges for the "outside" researcher in community-based participatory research. *Health Education & Behavior*, *31*(6), 684–697.

Minkler, M., & Wallerstein, N. (Eds.). (2003). *Community-based participatory research for health*, San Francisco: Jossey-Bass.

Mt. Pleasant, J. (2001). The three sisters: Care for the land and the people. In K. James (Ed.) (2001). *Science and Native American communities: Legacies of pain, visions of promise* (pp. 126–136). Lincoln, NE: University of Nebraska Press.

Nader, L. (2001). Anthropology distinguished lecture-2000. *American Anthropologist, 103*(3), 609–620.

Norton, I. M., & Manson, S. M. (1996). Research in American Indian and Alaska Native communities: Navigating the cultural universe of values and process. *Journal of Consulting and Clinical Psychology, 64*(5), 856–860.

Pewewardy, C. (2001). Indigenous consciousness, education and science: Issues of perception and language. In K. James (Ed.). *Science and Native American communities: Legacies of pain, visions of promise* (pp. 16–21). Lincoln, NE: University of Nebraska Press.

Rigney, L. (1999). Internalization of an indigenous anticolonial cultural critique of research methodologies: A guide to indigenist research methodology and its principles. *Wicazo SA: Journal of Native American Studies Review, 14*(2), 109–121.

Sahota, P. C. (2007). *Research Regulation in American Indian/Alaska Native Communities: Policy and Practice Considerations*. National Congress of American Indians Policy Research Center. December 12, 2007. http://www.ncaiprc.org/pdf/1196282550Research_Regulation_final_paper_110607.pdf

Schmidt, C. (2001). Indi-gene-ous conflicts. *Environmental Health Perspectives, 109*, A216–A219.

Smith, A. (2006). *Conquest*. Cambridge, MA: South End Press.

Tuhiwai Smith, L. (2005). On tricky ground: Researching the Native in the age of uncertainty. In Denzin, N. K., & Lincoln, Y. S. (Eds.). *Handbook of qualitative research* (pp. 85–107). Thousand Oaks: Sage Publications.

Truth Commission into Genocide in Canada. (2001). *Hidden from history: The Canadian holocaust: The untold story of the genocide of aboriginal peoples by Church and State in Canada*. The Truth Commission into Genocide in Canada, Nanaimo, B. C. http:// www.hiddenfromhistory.org/

United Nations General Assembly (2007), *The United Nations Declaration on the Rights of Indigenous Peoples* [electronic version]. Retrieved November 6, 2008 from http://www.un.org/esa/socdev/unpfii/documents/DRIPS_en.pdf

Udel, L. (2001). Revision and resistance: The politics of Native women's motherwork. *Frontiers: A Journal of Women's Studies, 22*(2), 43–62.

Von Däniken, E. (1968). *Chariots of the Gods? Unsolved Mysteries of the Past*. New York: Bantam Press publishers.

Wallerstein, N. B., & Duran, B. (2006). Using community-based participatory research to address health disparities. *Health Promotion Practice*, *7*(3), 312–323.

Walters, K. L., & Simoni, J. M. (2002). Reconceptualizing Native women’s health: An “indigenist” stress-coping model. *American Journal of Public Health*, *92*(4), 520–524.

Walters, K. L., Simoni, J. M., & Evans-Campbell, T. (2002). Substance use among American Indians and Alaska Natives: Incorporating culture in an “indigenist” stress-coping paradigm. *Public Health Reports*, *117*(suppl. 1), S104–S117.

# 9

# The Worst of all Possible Program Evaluation Outcomes

*Leonard Bickman and Michele Athay*

> I have missed more than 9,000 shots in my career. I have lost almost 300 games. On 26 occasions I have been entrusted to take the game winning shot, and missed. And I have failed over and over and over again in my life. And that is precisely why I succeed.
>
> —Michael Jordan

The history of innovation is replete with failures. Thomas Edison tried more than 6,000 different materials before finding a suitable filament for the light bulb; the Wright brothers experienced failure after failure in the tedious testing of kites and gliders before the successful flight of their famous flying machine; and Robert Hutchings Goddard, the father of modern rocketry, endured years of failed attempts and public ridicule before launching the first successful rocket. The fate of entrepreneurs is similar. According to the U.S. Small Business Administration (2007), approximately 50% of small businesses fail within the first five years.

Failure is not unique to inventors and entrepreneurs. Rather, it is a common occurrence in all fields including program evaluation in the behavioral and social sciences. Thus the first questions that could be asked are: How common is failure in social and educational experiments? What is an acceptable failure rate? How predictable is failure? We do not know of any social or behavioral data that directly speak to these questions, but other areas of study have compiled this type of information including the following:

- **Clinical drug trials:** Despite scientific advances, a drug starting human trials by the year 2000 was no more likely to reach the market than one entering trials in 1985 (roughly an 8% chance). Moreover, for pharmaceuticals, the product failure rate in phase III trials has increased to nearly 50% (U.S. Food and Drug Administration, 2006).
- **National missile defense:** According to Korb (2000), the vice president and director of studies at the council on foreign relations, attempts to destroy mock warheads have failed more than 70% of the time, since 1976, when research on hit-to-kill weapons began.
- **Information technology:** According to research done by The Standish Group (1995), 31.1% of all software projects are canceled before they are completed. In addition, the Organizational Aspects of IT Special Interest Group (OASIG) of the Operational Research Society (OASIG, 1996) concluded from its research that 7 out of 10 information technology projects "fail" in some respect.
- **Alcohol addiction:** As a testament to treatment failure for alcoholism, Sobell, Cunningham, and Sobell (1996) found that alcoholics who do not seek treatment are significantly more likely to recover and maintain sobriety than those who receive treatment. Alcoholics Anonymous is even cited as ineffective. In a 1996 analysis of membership data using the figure of five years of sobriety as the criterion of success, the Alcoholics Anonymous success rate was approximately 2.6% to 3.5% (Bufe, 1998).

It is safe to assume that the evaluation of social and educational programs also has its fair share of failure. However, the complexity and vagueness of social and educational programs make the cause of program failure difficult to determine. It is often not possible to answer such questions as: What critical elements were missing that must be present to succeed? How many of the components need to be well implemented if the program is to have an effect? And, probably the most difficult question, can the quality of the implementation be measured, and how good does it have to be to succeed?

In a quantitative evaluation, if the effect obtained is beyond what is expected by chance, then it usually is labeled as being statistically significant. This effect may be positive, that is, the treatment was better than the comparison, or negative, in which the comparison is better than the treatment. This chapter focuses on a third possibility: finding no statistically significant effect, or a null result. Null results refer to not finding any meaningful effect of the program, in contrast to negative results, which show the intervention may have been harmful. Some may consider a null effect to be beyond interpretation. A null effect can be attributable to any combination of a poorly implemented program, an incorrect program theory, or a poorly designed and conducted evaluation. The latter is the worst interpretation of a null effect for the evaluator, and the former is the worst for the persons delivering the program. Program developers must be most concerned about theory failure. The most useful interpretation of a null effect is when the theory can be concluded to be wrong. The other two interpretations may simply indicate incompetence or demonstrate that implementation or evaluation of the program is more difficult than anticipated (which is almost always true).

In this chapter we will discuss factors that contribute to finding null effects and some common reactions to reporting these findings. Also presented are some actions evaluators can take to prevent null effects and potential problems created for the evaluator who obtains null results.

## Finding Null Effects

> It is possible to fail in many ways…while to succeed is possible only in one way.
> —Aristotle

As noted earlier, the finding of null effects (disparagingly called "no results") is mainly attributable to flaws in three categories: (1) the program theory was wrong or too limited; (2) the implementation of the program was not of sufficient quality, strength, or fidelity; and/or (3) the evaluation's measurement, design, data collection, and analysis were inadequate. Each of these categories includes several factors that could result in null effects. This chapter argues that if the evaluation can demonstrate that factors 2 and 3 could not plausibly explain the null effects, then even a null effect may reveal important information about the effectiveness of the program and its underlying theory. The key is designing a study that provides first-class evidence on whether factors 2 and/or 3 are reasonable and plausible explanations for the null effects.

### *Limited or Wrong Theory*

One reason for finding null effects is that the original idea or theory underlying the program was not correct. Every program should have a conceptual framework or theory that describes how, why, and under what conditions the program should be effective (Bickman, 1987, 1990, 2000; Chen & Rossi, 1983). The theory is a description of how the intervention causes the outcome. For example, prisoners may be observed to have poorer penmanship than honest citizens. However, developing a penmanship course for prisoners will probably not be effective in reducing recidivism, simply because the theory that poor penmanship causes criminality is wrong.

A more pertinent example of program theory failure is a program evaluated by the first author that is discussed throughout this chapter. The program developers believed that if children with mental health problems were to have access to a full array of mental health services (a system of care), they would get better faster, and the services would be less expensive. This appears to be a reasonable basis for expecting services delivered in a system of care to produce better outcomes than those not in a system of care. However, the program model had many assumptions that the developers did not consider. For example, they assumed that the services themselves were effective. Ineffective services, regardless of how they are delivered, would still not produce an effect. They assumed that the clinicians would be able to match the children's needs with the appropriate services. They assumed that reducing expensive hospital stays would reduce costs without considering that long stays in less-expensive residential facilities would add up to be more expensive. The evaluation addressed many of these implicit components of the program theory and found little support for them. Moreover, data were presented that maintained that the program was implemented well and the evaluation had no important flaws. If the conclusions were accepted, then the field would have learned that the system of care approach is not a good investment unless there was evidence that showed that the underlying assumptions were correct.

### *Insufficient or Poor Program Implementation*

If a program is not implemented with fidelity to the underlying theory, then regardless of the outcome, the results of the evaluation cannot be used to judge the success or failure of the theory. The intervention needs to represent the theory underlying the program faithfully. Some of the program implementation factors that could be responsible for null effects are not providing sufficient intensity or duration of the intervention, not training staff sufficiently, not recruiting or keeping a sufficient number of staff, misidentifying the problem, including inappropriate clients, and allowing contamination between experimental conditions such as crossing from one condition to another. However, most interventions are so underspecified that determining which components of the program are the most important to implement with what level of fidelity is challenging. An evaluator also has difficulty determining the degree or intensity of participation needed to find an effect if the program developers do not provide plausible criteria. Is three hours of staff training sufficient, or is a week needed? Regardless of the length of training, how good a trainer is needed? Are the training materials understandable and appropriate? Are the program conditions needed to implement the training?

The study of implementation is in its infancy, so determining what is meant by proper or sufficient fidelity is often not possible. For example, a null outcome found in a large-scale evaluation of a system of care for children with mental health problems (Bickman et al., 1995) was criticized by one of the program developers, who believed that the parents were not sufficiently involved in the treatment to produce an effect (Friedman, 1996). In other words, an implementation failure occurred, so no conclusion could be drawn regarding the effectiveness of systems of care. However, the amount of parent involvement in the treatment and control groups was statistically significantly different. But was this sufficient? How much more parent involvement was needed to produce the desired effect? Unfortunately, the developer was unable to specify how much parent involvement was necessary. Friedman was not able to be that specific because the theory was weak and nonspecific, and the measurement of involvement was imprecise.

If a program is composed of several elements, then how many elements need to be present at what level of implementation for how many clients and for how long to find an effect? Currently few, if any, program developers can answer that question (Cook & Kilmer, 2004). Thus, it behooves the evaluator to include in an evaluation plan measures of implementation that program developers and implementers (potential critics) find acceptable. This recommendation is made realizing that the developers and implementers may have no special insight to provide to the evaluator. The recommendation is more to provide some protection for the evaluator than to provide useful knowledge.

### *Inadequate Evaluation*

The evaluation can also be a cause of null effects, if there is not enough statistical power to detect an effect if one exists. It may also be the case that the evaluator may have used insensitive instruments or instruments that were not relevant to the intervention. A good example of the latter is the use of IQ measures to judge the success of Head Start programs; IQ is a reliable measure, but the program was not supposed to affect IQ (Westinghouse Learning Corporation, 1969). In addition, the evaluator may contribute to a null effect by sloppiness in collecting and cleaning the data, or in the type of statistics used to analyze the data. In this case, the theory could be correct and the program well implemented, but a null result could still occur if the evaluation was critically flawed. When this happens the evaluator must be able to defend the soundness of the evaluation—or find a place to hide.

As you can see, there are many ways an evaluation can fail to find an effect even if one is present. Therefore, the fact that many evaluations do not find effects should not be surprising. The late Peter Rossi, one of the founders of the field of evaluation, called this Rossi's iron law: "the expected value of any net impact assessment of any social program is zero" (Rossi, 1987). Also relevant is his plutonium law: "program operators will explode when exposed to typical evaluation findings." We don't know of any empirical research that seems to support these laws, but they demonstrate the points being made in this chapter.

## Reactions to Null Findings

> I don't know the key to success, but the key to failure is trying to please everybody.
> —Bill Cosby

Given the real possibility of null effects, what determines the reaction when an evaluation indicates no evidence of effectiveness? First, it will depend on whether the evaluation is a hot topic. It may have been quite visible when it was planned, but in the several years required to complete the evaluation, people may lose interest, change jobs, discover other problems, and so forth. If the program is highly visible and expensive, then much attention likely will be paid to the evaluation outcomes. Second, the response to null findings in an evaluation is determined by the attitude and actions of the evaluator. An evaluator who is busy competing for the next contract may lack enthusiasm to publish or make the findings public. Moreover, if the individuals who commissioned the evaluation have already moved on to another position, little or no pressure may exist to make those findings visible.

Factors that influence the reaction and visibility of the evaluation findings include desirability of burying (or altering appearance of) results, the evaluator's role and goals, publication bias, and a change in program goals that negates the usefulness of the findings.

### *Burying Results*

If both the evaluator and the funding agency want to "bury" the findings, and the persons receiving funds like the program, then it is likely that nothing negative will happen. For example, the federal government has spent almost one billion dollars on The Comprehensive Community Mental Health Services Program for Children and Their Families over its 12 years of existence. Although Congress has re-funded this program several times, and several million dollars have been spent by the evaluation contractor, at the time of this writing no outcome studies of the program have been published in peer-reviewed journals. If the results are null, as suspected in this case, then it is in no one's interest to publicize or act on the findings. The program has nationwide support and thus its own constituency and the government would rather not publicize the continued funding of a program that lacked evidence of effectiveness. The only people who could gain from acting on the results are the clients, and they usually are satisfied if they receive services, regardless of the lack of evidence of effectiveness.

The movement from descriptive narratives of research findings in a field to meta-analysis also may have an effect on expectations about success and failure. For example, Multisystemic Therapy has been adopted worldwide as an effective program to help children with behavioral problems. It is recognized as a Model Program by the Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services, and the Office of Juvenile Justice and Delinquency Prevention. Model Programs are a step above simply being an effective program. It also has been labeled an Effective Program by the U.S. Surgeon General's Reports on Mental Health and Youth Violence and has received the Families Count Award from the Annie E. Casey Foundation. The SAMHSA Web site (www.modelprograms.samhsa.gov) describes the "proved results" being based on eight randomized experiments. However, a meta-analytic review (Littell, 2005) that has been vetted by both the Cochrane and Campbell Collaboratives reports no evidence of effectiveness and that narrative reviews have been wrong in the conclusions they have drawn. The findings of the meta-analysis have recently been published and have resulted in an acerbic exchange between the program developers and the author of the meta-analysis (Henggeler, Schoenwald, Borduin, & Swenson, 2006; Littell, 2006). The issue of effectiveness of the premier youth treatment program is far from over. Results of new studies can alter the conclusions about effectiveness. A recently published study conducted in Sweden reported no effects of the program (Sundell et al., 2008) but other studies in progress may find different results. What this should be telling us is that the simple listing of programs that work or don't work is not appropriate and is misleading. What we need to know are the conditions under which a program is effective since it is unlikely that all programs are effective regardless of how and where they are implemented. Moreover, we need to remind ourselves that in science all knowledge is provisional and dependent on future findings.

### *Evaluators' Role*

The evaluators' actions subsequent to a null effect are influenced by their values. If the evaluators see themselves as neutral parties working for the funding entity of the evaluation, then they may believe that when the final report is given to the sponsor the evaluators' activities are finished. Others may define their role as activists or advocates who champion the findings. This may especially be the case when the evaluation is of a publicly funded program and the evaluators believe that they have a responsibility to represent clients or society in general. Some evaluators take the stance that they owe it to their subjects to publish findings because their subjects provide their valuable time with the understanding and trust that they are helping others by taking part in the study.

### *Publication*

Although communication of research findings is one of the most important functions of researchers, as many as 50% of studies may not be published in a particular area of research (Scherer, Dickersin, & Langenberg, 1994). One explanation for this is publication bias against null results, often called the "file drawer effect." This has some potentially harmful effects, especially on the usefulness of meta-analyses. Publication bias can occur in two ways: by editorial bias preference and author bias.

- **Editorial bias:** This occurs when a journal decides which studies are suitable for publication on the basis of the statistical significance of a study's results. Editors have demonstrated a bias toward acceptance of positive outcome studies over those with null results (Coursol & Wagner, 1986; Scargle, 2000; Stern & Simes, 1997).
- **Author bias:** This bias occurs when investigators do not submit their research for publication. Deciding not to pursue publication may result from lack of interest, over-commitment in subsequent studies, a consequence of a drive to publish only statistically significant results, or the thought that null papers are given low publication priority. Several studies have shown author bias to be the leading cause of studies remaining unpublished (Begg & Berlin, 1989; Shields, 2000).

In some (typically nonacademic) organizations working under a contractual agreement, the contact may specify that the evaluator cannot submit for publication any document that has not been approved by the sponsor. These clauses are rarely allowed in contracts with academic institutions. Thus, many universities do not permit their faculty to engage in government research that has a secret or higher security level for that reason.

Does being a tenured professor and its guaranteed employment support more independent evaluations, especially when the results are not popular or are contrary to mainstream thought? We contend that evaluators in such protected positions have the added responsibility to act on unpopular outcomes. However, academics are also subject to wanting the next grant, publication, speaking invitation, consultation, and positive reputation that being associated with unpopular positions typically do not generate. As a result, null findings may not be pursued for publication as often as positive results. However, null results are most important when the program is well implemented and evaluation is conducted in a competent and fair manner.

### *The Changing Definition of Success*

When faced with null outcomes, evaluators or organizations may simply change their definition of success in a way that reframes the goals of the intervention or program and transforms them in a positive light. For example, the D.A.R.E. program is the most popular school-based drug abuse prevention program in the United States despite the fact that evaluations of the program have shown it to be ineffective in preventing drug abuse among adolescents (Birkeland, Murphy-Graham, & Weiss, 2005; Rosenbaum & Hanson, 1998; Weiss, Murphy-Graham, & Birkeland, 2005). Some of these studies even showed a negative effect in which students increased drug use after being exposed to the program. Faced with this information, the D.A.R.E. organization continued to expand the program by altering the goal of the program from reduction of drug use in adolescents to the development of better relationships between children and the police (Birkeland et al., 2005). In this case, the organization simply negated the impact of the evaluative findings by changing the goals of the program! Thus, D.A.R.E. remains a popular community-supported program in the majority of U.S. schools even though it does not do what it was originally marketed to do.

Because tying changes in policy to a specific positive evaluation typically is difficult, immediate direct consequences of identifying a program as effective when it is not usually do not occur. While there clearly are opportunity costs in incorrectly identifying a program as effective, positive findings usually support the status quo. Moreover, because almost everyone benefits from positive findings, the evaluation will not be as carefully scrutinized. On the other hand, identifying a program as ineffective may have more immediate negative consequences. The organization that funded the program and/or the evaluation expects the program to have positive outcomes and, naturally, the persons implementing the program do not want to think that the years they spent working in the program were a waste. Even the evaluator favors a positive outcome because association with something that is effective is much better than a program that is ineffective. Because most stakeholders want to see positive outcomes, these evaluations may not be critically examined so that everyone can remain pleased with his or her role in the program. However, different dynamics exist when the results obtained are not the expected positive ones. If outcomes are null, then a search for whose fault it is typically occurs, and the evaluator may be the first person in the hot seat. Cautioning not to kill or shoot the messenger, in this case the evaluator, has a long and distinguished tradition as far back as Sophocles in 442 BC and much later by Shakespeare in several of his plays (Wilson, 1934). The evaluator typically is the most vulnerable of the actors. Evaluators usually are not part of a government agency, do not have an advocacy group to support them, provide little or no employment opportunities for other stakeholders, and typically are the outsiders. We shall present some suggestions about how the evaluator can be shielded against some of the potential attacks, but be assured some damage will be done by critics of the evaluation. The following section discusses several ways to prevent these results.

## Prevention of False Null Results

> Integrity without knowledge is weak and useless, and knowledge without integrity is dangerous and dreadful.
> —Samuel Johnson

The first line of defense is prevention. False null outcomes can be prevented by conducting a flawless evaluation of a powerfully effective program implemented with high fidelity. Of course, this is a fantasy. Flawless evaluations may be planned, but they are never executed. Even programs that have been identified as “effective,” “model,” or “proven” by federal agencies have been found to have critical flaws in their evaluations that often are overlooked (Gandhi, Murray-Graham, Petrosino, Chrismer, & Weiss, 2007). Furthermore, null outcomes should not be avoided when they are true. However, we should recognize that false-positive outcomes do not have the problems associated with them that false null outcomes have. In any case, researchers can take many steps in the prevention of null results.

### *Integrity*

Evaluators with integrity do not try to minimize or soft pedal the bad outcomes. Simply said, the evaluator should "tell it like it is," but that implies that the data speak for themselves. Evaluators can slant the results by leaving out analyses that make the client uncomfortable, or torturing the data so much that it finally shows what people want it to say, or taking other subtle approaches that cannot be easily detected. These approaches are typically sins of omission and not commission, and may give the evaluator the feeling that he or she did not fudge the data by changing results. In the authors' opinion, these are not that different from each other. For the purpose of this chapter, let us assume a situation in which the evaluator does possess integrity but still wants to avoid, as much as possible the fallout from the null results.

### *Avoid Situations in Which the Evaluation Sponsor and Implementation Sites Are from Different Organizations*

In many situations the organization sponsoring the evaluation may not be the same one that is implementing the program. This most frequently occurs when the evaluator is hired by a government agency to evaluate a program. In this instance the evaluator is external to the implementing organization. The problem with this situation is that the evaluation may be seen as an additional burden that generates staff resistance that may express itself in several unpleasant ways. For example, if the program staff is responsible for some data collection, then a risk exists that the data will not be collected as carefully as the evaluator wants. This situation sometimes can be avoided if the evaluation budget contains sufficient funds for independent data collection.

### *Avoid Situations in Which the Evaluation Sponsor and Implementation Sites are from the Same Organization*

This is contradictory to the previous point but both situations can cause problems for the evaluator. When the evaluation sponsor and implementation site are from the same organization, the sponsor likely will have a larger stake in obtaining a positive outcome. One potential negative consequence is that the evaluator will be subject to pressure to report the findings in the most favorable light. Moreover, a greater tendency exists for the client to be dissatisfied and blame the evaluator for any negative outcome. The sponsor will also have more difficulty blaming the organization for poor implementation, thus again leaving the evaluator in the hot seat.

### *There Are Levels of Approval, and* Levels of Approval

Obtaining approval to conduct the evaluation from one level of the organization does not mean the other levels will be cooperative. For example, an educational evaluation may be supported by the director or superintendent of schools, but the evaluator may find that the principals and teachers were not consulted, and are not in favor of the study. Rarely will this result in a direct confrontation but more often as a lack of cooperation and passive-aggressive behavior that can result in a failed evaluation implementation. Even a sensitive superintendent may not have the time, given grant deadlines, to consult all concerned.

### *Statistical Power Issues*

Statistical power refers to the ability of the study to detect an effect if one is present. By now everyone who performs serious research should be sensitive to statistical power issues. Calculating statistical power is a significant planning tool, and with the advent of multilevel designs, statistical power has become even more of an issue. Power depends on several factors, including the expected effect size, the sensitivity of the statistical analysis, the characteristics of the measures, and of course the sample size. Power can contribute to obtaining a null result in the following ways:

- When the number of subjects to permit observation of a true effect is insufficient (low power to detect an effect).
- When the number of subjects is adequate but a meaningful effect does not exist (high power, no effect).
- When there is low power (small sample) and no real effect exists.

The evaluation planner typically tries to determine what sample size is needed to detect the desired level of an effect. Ideally, the research literature should be consulted to determine what effect size has been found for the type of intervention planned. Alternatively, if no reliable estimates are available, the evaluator should try to obtain the estimated minimum effect size that the stakeholders want to detect. However, effect size is a difficult concept to grasp. The evaluator will need to translate the metric of effect size into something meaningful to the stakeholders. For example, instead of saying an effect size of 0.50, which usually is called a medium effect size, the evaluator might say that this is equivalent to $x$ number of additional days of school for a reading intervention.

The planner typically will develop some post hoc reasons for predicting a minimum effect size based on the sample size available for the amount of time and money that the researcher has to conduct the research. This type of approach may provide the illusion needed to obtain funding, but it does not necessarily produce a well-designed study.

Detecting smaller effects requires more power (which translates to more time and money) than do larger effects. Stakeholders can be supportive in obtaining more resources if the evaluator is able to explain to them that the higher the power of the design, the more likely a statistically significant effect will be detected if one is present. If a realistic effect size is too small to be statistically significant, this may result in a change in design, altered evaluation questions, the addition of more resources, or the abandonment of the study. The latter rarely occurs if funding has already been obtained. Having stakeholders sign off on a minimum effect size may help prevent the evaluator from being blamed for an evaluation that found a small but not statistically significant effect size. However, the desired effect size should be able to be supported by either previous findings or what is needed for a policy decision. It is important to remember that statistical significance does not necessarily imply policy or clinical significance.

### *Measurement*

Another issue that the evaluator will want the stakeholders to sign off on is the measures or instruments used in the study. There is simply not enough time, money, or participant patience to measure everything. Choices have to be made between breadth and depth of measurement. Given a fixed amount of resources, should the measures cover many areas in the anticipation that to do otherwise would risk not finding an effect, or should measures be more focused with better psychometric properties to increase the probability of detecting an effect? This is a trade-off decision the evaluator should make in consultation with the stakeholders. Regardless, if null results are obtained, the evaluator wants to avoid as much as possible stakeholders accusing the evaluator of measuring the wrong things or using inappropriate measures.

### *Ensure Time to Implement Intervention Is Sufficient*

A rule of thumb in predicting the amount of time necessary to recruit participants, implement the intervention, and collect the data is to make a best estimate and then double or triple it. No matter how careful the planning has been, something will happen to disrupt it because the real world is not controlled by the evaluator. That is one of the important reasons some researchers like laboratory studies. As with royalty, they have great control over their subjects.

For example, in a study involving military dependents, the military member of the family was deployed to Iraq, thus disrupting the family and lengthening the time required for data collection (Bickman et al., 1995). When the family member returned, an additional disruption occurred. Then, in the same study, the investigator was assured by the sponsor that a dataset was available that could be used to locate youth starting treatment. What they did not tell the investigator was that the data would be available to use six to nine months after the child started treatment. In this study, the clients needed to be recruited within 30 days of entering treatment. What seemed like a simple recruitment plan involving sampling from a list developed into a time-consuming and expensive process. The investigators had to visit all the private service providers in the area each week to ask if they had any new eligible clients that week. They were visited because they could not be depended on to contact the investigators about new clients. The personal visits were time consuming but necessary because personal relationships with the staff were required for them to cooperate and help obtain participants.

In planning for recruitment, it is advisable to conduct what has been called a "pipeline study." Because most evaluations accrue subjects as they enter treatment, in contrast to a survey in which all subjects are recruited nearly at the same time, the rate of recruitment in an evaluation depends on the number of eligible clients willing to participate in the study. If the program is already in operation, learning how many clients enter treatment each month may appear simple. However, simplicity is deceptive. First, the program may not know the number. Do not trust an estimate; we have found that service providers overestimate how busy they are serving new clients. Several other factors that are difficult to predict affect recruitment. The senior author has been wrong most of the time in predicting the length of time required for recruitment. For example, he conducted a study in which youth were entering mental health treatment (Bickman, Summerfelt, & Noser, 1997), and the program was able to show how many new clients entered each month. But the program could not reveal the ages of the clients, so he had to guess at the number of youth aged 11 to 18 years. Similarly, because only one child from each family was eligible for the study, knowing how many siblings were present among the new clients would have been helpful. Both factors were underestimated, and a significantly longer time was required to obtain the necessary number of eligible subjects for the study.

Other factors affecting recruitment are even more difficult to predict. When program staff are in charge of the initial recruitment, any staff turnover affects their recruitment performance. Environmental conditions also affect recruitment. In the previously discussed study, some staff initially resisted the evaluation by claiming that few clients agreed to participate. However, when record keeping improved and supervision increased, the numbers went up. In another study involving physicians, the estimate of physicians volunteering in the study was much lower than expected. We suspect that this study may have had a lower rate than others involving physicians because the physicians were the subjects in the study, a role about which they were ambivalent. In recruiting clients from waiting rooms, the authors had to station a research assistant to recruit. Although the head of the department assured us that staff would be cooperative, in fact they were not. These examples illustrate the difficulty of accurately predicting recruitment.

### *Ensure Sufficient Time from Implementation to Follow-up*

Although any evidence concerning the length of time or the amount of exposure necessary for the program to show an effect is rare, discussing this with the stakeholders while designing the evaluation is a good idea. Two time periods are important: the amount of time a new program takes to become operational, and how long or how many sessions it takes for the treatment to affect the client. If null results are obtained, the argument may be made that the program was not fully mature or that more time is required for an effect to emerge. *A priori* estimates from stakeholders may be helpful in countering this criticism. When faced with that criticism, the first author has analyzed the data by number of sessions or some other indicator of exposure to see if the data were related to outcomes (Salzer, Bickman, & Lambert, 1999). The concern about whether the program was sufficiently mature can be tested by determining whether the program showed more effects with later clients compared with earlier clients. These analyses should be part of the analytic plan.

### *Ensure Sufficient Funds to Conduct the Study*

Insufficient funds jeopardize the integrity of the evaluation. It is the evaluator's responsibility to determine whether the client's expectations can be met by the amount of funding available. Sometimes a set percentage of the total program costs is allocated to the evaluation. This may seem reasonable, but little justification exists for these fixed percentages. In general, costs are related to the breadth of the evaluation and the degree of certainty of the correctness of the results wanted by the client. For example, not having a control group is less expensive, but the certainty that the program was the cause of the outcome is severely diminished. Hendrick, Bickman, and Rog (1993) have a chapter devoted to calculating evaluation costs. However, like time, the costs of an evaluation are usually underestimated.

### *Minimize Attrition*

As noted earlier, access to a sufficient number of clients is probably the most difficult aspect to control. Minimizing attrition is the second most difficult. Attrition can bias the dataset, result in a loss of statistical power, and reduce the internal and external validity of a study's findings. Two main statistical concerns exist. First is differential attrition from the treatment or comparison groups. A randomized experiment is no longer randomized when a large differential attrition is present. Selection artifacts can occur during the evaluation as well as during the recruitment phase. Attrition can also be from the program or from the evaluation. For example, biases typically occur when data are collected only from clients in treatment. This type of attrition from the program is not under the control of the evaluator. It is better to collect data independently of treatment; however, this is more expensive and not always feasible. Prevention of attrition from the evaluation has not been sufficiently studied, although several sources provide hints on how to control attrition.

Techniques for retention of study participants include incentives for participants, continual mail or telephone contacts by project staff, providing postage-paid postcards for participants to inform the researcher of a change in address or telephone number, and attempts at tracking participants if contact is lost (Drotar, 2000; Senturia et al., 1998).

### *Obtain an Externally Valid Sample*

Having just enough clients to meet statistical power issues is not sufficient. The clients in the evaluation must represent those in the program; otherwise the results may be ignored due to lack of representativeness. Drawing representative samples is not usually done in evaluations because of the difficulty in obtaining a sufficient number of clients. Despite its impact on an evaluation, few studies have experimentally examined the recruitment process of acquiring subjects (Brown & Liao, 1999). Having some data from all clients, as with demographics, at least allows the evaluator to make some comparisons between those who participated in the evaluation and all clients. Although far from proving representativeness, this analysis may provide some indication of bias in the evaluation sample.

### *Build in Research Questions*

Although null results may not be seen as useful by the sponsor or implementers, including sound research questions increases the value of the evaluation to the research community. Evaluations typically have more resources than are available to researchers conducting basic research. Often the addition of important research questions in the design and analysis of the evaluation adds little or no cost. The senior author's Fort Bragg study (Bickman et al., 1995) produced more than 50 publications on information collected in the evaluation that were not evaluative. However, the inclusion of these questions requires that the evaluator or someone on the evaluation team be both interested and knowledgeable about the substantive area of the evaluation.

### *Who Controls What?*

The issue of control over the evaluation is critical. Many decisions must be made within the constraints of the budget. Decisions are made about research design, instrumentation, sample size, analysis, and written reports. Trade-offs will have to be made because funds and time to do everything at the highest standard are never sufficient. There are also major differences in opinion among the evaluation community regarding the sharing of these types of decisions. For example, the empowerment orientation (Fetterman, 1994) involves almost all stakeholders in these types of decisions. The authors' approach is to recognize the strengths that each stakeholder brings to the evaluation; the evaluators should know more about sampling, psychometrics, design, measurement, and analysis, whereas the program personnel should know more about the context, feasibility, and which questions are important to the service community.

After explaining all the trade-offs, the evaluator should negotiate with the client the best approach to take. Avoid situations in which the evaluator must follow the golden rule, which some define as "those who have the gold make the rules." At this juncture, the evaluator's professional identity and integrity are tested. If a client's expectations are that only a positive outcome is acceptable, the evaluator should drop further involvement. If the design, method, measures, analysis, and program are such that the evaluator is confident that the conclusions the client wants to draw from the evaluation could not occur and no amount of discussion appears to have an effect, the evaluator should not conduct the evaluation. As a preventive method the evaluator must be certain that the contract with the client allows the evaluator to end the relationship. Drawing a contract in two phases may be preferable, in which the first phase is the acceptance of the evaluation plan, and the second is execution of the plan.

### *Transparent Statistics*

As designs get more complex and the availability of sophisticated computer programs increases, understanding how the results were obtained becomes more difficult. This implies that the evaluator has to make an extra effort to clarify complex results for the reader. In the case of null results, the issue is even more complex. As all basic-level statistics classes teach, hypothesis testing attempts to reject the null hypothesis and conclude that an alternative hypothesis (that a significant difference exists) is true. Students have been taught for generations that the null hypothesis cannot be proven. For example, proving unicorns do not exist is not possible because that would require omniscience. No one can "prove" that the null results were caused by an ineffective program because, as previously noted, several reasons exist as to why null results would occur. As an alternative to traditional null hypothesis testing, an interest in equivalence evaluation has occurred (Rogers, Howard, & Vessey, 1993; Seaman & Serlin, 1998; Stegner, Bostrom, & Greenfield, 1996). In equivalence testing, the goal is not to demonstrate that the difference between two groups is zero, but rather to show that the difference is smaller than what would be considered meaningful.

The emergence of computer programs to analyze multilevel designs (e.g., students nested in classes and classes nested in schools) has provided an approach that recognizes the dependency of observations. However, it was simpler to consider a classroom level intervention as simply having an effect on students that is not affected by the fact that they are in the same class and school. Although the use of sophisticated computer programs can help analyze complex data, the researcher's familiarity and fluency in the correct use of such programs becomes more critical. The evaluator wants to avoid demonstrating that "to err is human, but to really foul things up requires a computer."

### *Plan for Moderators and Mediators*

An intervention typically does not work the same way for all people in all situations; rather, it has different effects on different subpopulations (Brown & Liao, 1999). The overall result of the study may not show an effect, but by measuring mediators and moderators the program may show an effect under certain circumstances. For example, if a program to teach people how to repair a car is being evaluated, randomly assigning students to the teaching or no teaching conditions and recording their ability to make the repairs is an appropriate evaluation design. However, if a mediator such as knowledge is theoretically predicted and measured, and findings showed that the teaching group did not learn anything additional, then this provides a hint about where to look for the failure. Was it the teaching? Did other teachers also show no effect? Was it the curriculum? Did experts review the curriculum to determine if it was clear and understandable? The mediator approach helps determine if the problem was in transferring knowledge into skills because no difference in learning occurred. However, these subgroup analyses do not maintain the internal validity of the original design. A study that is labeled a randomized experiment does not mean that all the statistical comparisons the evaluator makes have the same advantages (e.g., selection) that the primary analysis has.

The approach mentioned earlier relates to the theory that underlies the program. This has been termed program theory (Bickman, 2000), or theory-driven evaluation (Chen, 1998). This orientation has developed into a sophisticated approach to planning called logic modeling. Several good sources detail this approach (Hernandez & Hodges, 2003; Millar, Simeone, & Carnevale, 2001).

The use of moderators can sometimes uncover an effect that was present, but only for some of the participants. Imagine that the car repair scenario used a coed class (adopting traditional stereotyped attitudes), and the boys thought the class was great and learned a great deal, which was demonstrated by their high repair scores. On the other hand, the girls were turned off by the class and now have a much more negative attitude toward repairing a car and thus did not try that hard. Examining the moderator of gender can help better explain the effects of the car repair class.

### *Review and Approval of Manuscripts*

A sticky point that often arises in contractual work is who controls publication. In a project the senior author negotiated, the government wanted to approve anything submitted for publication. Those terms were refused. The government then countered that if the authors did submit something, then the government would have the right to include whatever it wanted in the manuscript. The authors replied that if the government wanted to publish an article, their representatives should write their own manuscript. The government next wanted to ensure that the companion piece would be published. The authors stated that decision was the editor's. The authors finally settled on providing a copy of the manuscript, for information only, 90 days before it was submitted for review for publication. Although attempts at control are not necessarily as heavy handed as this, they do occur. In a different federal agency, the authors had to agree to not publish evaluation findings until the government had published their own report. The major problem is that the government report will not be published until after the funding is concluded, possibly several years later. In another situation the group being evaluated wanted to publish with the authors as coauthors. The authors found that acceptable as long as they did not feel pressured into changing the text. Publishing the results of an evaluation with the program personnel is a risky proposition.

### *Study Implementation*

Another way to avoid blame for null results is to study the implementation of the project carefully. Such data can serve as formative feedback in helping the staff improve the implementation of the project, or it can serve as summative feedback concerning the fidelity of the implementation. The formative process is difficult to accomplish because it has to be done in real time. Telling a staff person that he or she was doing the wrong thing six months after it happened usually is not helpful. One way to accomplish this is to build in the implementation data collection as a part of the intervention. If null effects occur, the implementation data are critical in answering the question why the program failed. It also takes the pressure off the evaluator if the data show the program was not well implemented. However, as previously noted, specifying how well the program needed to be implemented to produce an effect is difficult.

### *Can the Sponsor and Program Deal with a Bad Outcome?*

We believe that similar to subjects in an experiment, the sponsor should be fully informed of the possibility for null or negative outcomes of an evaluation. Whether the sponsors are the funding provider of the program, the developer, or the implementer, they will have positive expectations about the evaluation's outcome. After all, why invest time and money in a program if it will not do any good? The evaluator must prepare the client for the possibility that the outcome may not support all or even any of the expectations of the client. Some role-playing may be useful here by asking the client to imagine how he or she would respond to null results. How would the client deal with funders and clients served by the program and staff? This is an important step for ethical reasons and because the evaluator wants to arm the client with responses other than what may unfortunately be a typical response—it was the evaluator's fault.

## Evaluation Aftermath

> What we call the beginning is often the end. And to make an end is to make a beginning. The end is where we start from.
> —T.S. Eliot

### *Criticism Will Occur*

If the program and evaluation are important, the evaluation will receive criticism. No evaluation is flawless, but investigators must be prepared to deal with flaws that could affect the conclusions about the program. The prevention tips presented earlier are useful at this stage.

### *"The Evaluation Was Perfect, but We Don't Do that Anymore"*

One of the most difficult tactics to deal with is when advocates for a failed program publicly state that the evaluation was nearly perfect and the evaluator was fair, but the program has now changed to take into account the evaluation findings. This may be an attempt to stifle all discussions about the evaluation findings because they are presented by the program advocates as no longer relevant. Instead of focusing on the findings, the evaluator must backtrack and question the advocates on how the program has changed to take into account specific findings. This is a much less interesting debate than discussing the evaluation and its findings.

### *Differentiating Evaluation Failure, Implementation Failure, and Theory Failure*

The triad of causes of null outcomes is not obvious. Facing a null result, the evaluator should be prepared to address the potential flaws in all three causes. The strengths and weaknesses of the program theory, program implementation, and evaluation must be discussed to arrive at a better understanding of the null results.

### *Invite Debate*

If the evaluators have confidence in their methods and conclusions, they should invite critics to appear with them in symposia and journals.

Airing the differences may not change the minds of the advocates from either side but may influence bystanders to learn more about the controversy. The senior author participated in discussions in journals that had special issues or sections devoted to an evaluation and its critics (Bickman, 1996b, 1997).

### *What Was Learned?*

We contend that just as much can be learned from an evaluation that had null effects as from one that demonstrated the effectiveness of a program. As previously discussed, several plausible explanations exist for null effects, but fewer alternatives exist for a positive outcome in a well-designed and implemented evaluation. A principle of the scientific method is that a researcher can only provisionally disconfirm, never confirm, a hypothesis. However, null effects can inform policy makers what direction they should not take; but the value of this information is tempered by the availability of other interventions that can be used to ameliorate the problem.

Examination of unintended results of the evaluation is also beneficial. At times the results of a failed experiment become a success as an answer to an entirely different research question. For example, in the late 1980s Pfizer scientists began testing a drug for angina but soon began to see it showed little benefit in human clinical trials. Instead of throwing in the towel and moving on to another drug, the scientists noticed an interesting side effect that they decided to pursue. That led to the development of the drug marketed under the brand name of Viagra. In the same haphazard way, the famous child's toy Silly Putty (Binney & Smith, Easton, Pa.) was created during experiments initially aiming to create a new type of rubber for the United States' use in World War II.

### *Plan Another Study to Resolve Questions*

Because null effects seem to generate more questions than answers, they can be a fertile source of additional studies. Unfortunately, evaluators rarely have the opportunity to conduct follow-up studies. The first author was fortunate to be able to secure funds to replicate an evaluation that produced null effects and design it so that the new evaluation answered some of the important questions raised by critics. In this way, evaluation may be similar to baseball statistics; on a baseball team, the player who holds the records for the most home runs usually holds the record for another statistic—strikeouts. Baseball legend Babe Ruth put it this way, "Every strike brings me closer to the next home run."

### *Change Areas*

If the evaluator is blamed for the null effect and his or her reputation is severely in question, then it may be time to move to a different field of study. Evaluators are flexible in what they evaluate. We know of several people who simply changed fields and have discovered that their negative reputation did not follow them. As Douglas MacArthur once said "We are not retreating—we are advancing in another direction!"

### *Sometimes Null Effects Do Matter*

We briefly mentioned the Ft. Bragg study earlier. This was a large-scale evaluation of an innovative system of delivering mental health services to children and adolescents. It was based on a major movement in that field to provide a system of care that provided a continuum of coordinated mental health services as well as nontraditional mental health services. This was expected to provide more effective treatment and improved mental health outcomes at lower costs (Bickman, 1996a; Bickman & Mulvaney, 2005). Results showed that the comparison condition, with no integrated services, had similar clinical outcomes but the costs per child were 59% higher in the demonstration site. Given that systems of care were the centerpiece of many human services reform efforts, it should be easy to imagine the controversy these results produced. It is informative to read what one evaluation scholar wrote about the evaluation about nine years later in an article dealing with influential evaluations.

Henry (2003) noted

> When an evaluation concludes that a program produced positive effects on important outcomes, it is easy to see that the findings could influence more positive attitudes about the intervention and that the findings could justify expanding the intervention.... Negative results could produce opposite reactions to avoid harming those that the program was supposed to benefit. But it seems more difficult to see how null findings could be influential. Implementation flaws or methodological problems

> could be the source of null findings, making them more difficult to act upon or to change pre-existing beliefs. (p. 520)

Henry's conclusion was that the results of the evaluation were

> ...to stall the enthusiasm for the "continuum of care" approach to managing mental health services. Instead, attention turned to managing costs...and evaluating the effectiveness of individual mental health treatments regimes, rather than management systems.... The Fort Bragg Demonstration Evaluation indicated that social betterment, in this case in the form of improved mental health outcomes for children and adolescents, was not to be found by moving toward a continuum of care system for mental health services, thus, forestalling costly changes in service delivery that would have left children no better off. If debunking bad ideas is the cognitive contribution that is a defining characteristic of democracies...then null findings, when supported by rigorous methods and careful assessments of program fidelity, can contribute. (p. 520)

Henry used the phrase "stall enthusiasm." In the subsequent years the uncritical enthusiasm decreased and there developed a significant movement toward developing evidence-based treatments. We do not think that evaluation and a subsequent study was the only activity responsible for these changes. However, the SAMHSA has continued to invest in a system of care program that has cost close to a billion dollars in the last decade. A clear outcome is unlikely when research confronts the strongly held beliefs of politically savvy advocates.

## Final Thoughts

> What is important is to keep learning, to enjoy challenge, and to tolerate ambiguity. In the end there are no certain answers.
> —Martina Horner

This chapter focuses on the evaluation's worst outcome, null results: what they are, how they are responded to, how to avoid them, and what to do when they occur. Although it would be desirable if all social and educational programs were built on strong theory and implemented with high fidelity so that an evaluation held to rigorous standards would always produce unambiguous results, this is simply not the world in which we live. Null results happen, and determining why is often difficult. These often dreaded results can occur for numerous reasons and are far more common than many researchers would like to believe. Discussion of null results receives little energy and is swept under the rug in lieu of the heart-warming topic of positive and significant results. However, ignoring the topic of null results does not make the reality of obtaining them disappear and does not help researchers prepare for, protect against, and handle them when they do occur. It is our hope that the information contained in this chapter, much of which was learned the hard way by the first author, can arm researchers and evaluators with the necessary insights, knowledge, and tools to help navigate their way through the muddy waters of null results and emerge relatively unscathed.

---

## POINTS TO REMEMBER

- Steps toward launching scientifically valid field research
    - Openly acknowledge the differing purpose between the goals of your research and the goals of the practice setting, and then identify the shared aims.
    - Create an advisory board to guide the research project.
    - Be clear about roles and boundaries. Set up organizational structures to respect those roles and manage the project but remain flexible and be prepared to change these structures if new needs arise.
    - Structure the research project to reflect the practicing setting's need. Be flexible.
    - Be prepared for the funding priorities of your grantor to shift. If they do, remain flexible and responsive, but do not compromise the integrity of your study.
    - Make the budget, the staffing, and the research goals explicit and transparent.
    - Share the credit and do not surprise your colleagues.
- The Community Intervention Development Model (CID)
    - Theoretically and clinically informed construction, refinement, and manualizing of the protocol within the context of the practice setting where it is ultimately to be delivered.
    - Initial efficacy trial under controlled conditions to establish potential for benefit.

- Single-case applications in practice setting with progressive adaptations to the protocol.
- Initial effectiveness test, modest in scope and cost.
- Full test of the effectiveness under everyday practice conditions, including cost effectiveness.
- Effectiveness of treatment variations, effective ingredients, core potencies, moderators, mediators, and costs.
- Assessment of goodness-of-fit within the host organization, practice setting, or community.
- Dissemination, quality, and long-term sustainability within new organizations, practice settings.

## REFERENCES

Begg, C. B., & Berlin, J. A. (1989). Publication bias and dissemination of clinical research. *Journal of the National Cancer Institute, 81*(2), 107–115.

Bickman, L. (Ed.). (1987). *Using program theory in evaluation. New directions for program evaluation, No. 33*. San Francisco: Jossey-Bass.

Bickman, L. (Ed.). (1990). Advances in program theory. *New directions for program evaluation, No. 47.* San Francisco: Jossey-Bass.

Bickman, L. (Ed.). (1996b). Special Issue: The Fort Bragg experiment. *Journal of Mental Health Administration, 23*(1), 7–15.

Bickman, L. (1997). Resolving issues raised by the Ft. Bragg findings: New directions for mental health services research. *American Psychologist, 52,* 562–565.

Bickman, L. (2000). Summing up program theory. In P. Rogers, T. Hacsi, A. Petrosino, & T. Huebner (Eds.), *New Directions for evaluation, 87,* 103–112.

Bickman, L., Guthrie, P., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C., *et al.* (1995). *Evaluating managed mental health care: The Fort Bragg experiment.* New York: Plenum Publishing.

Bickman, L., & Mulvaney, S. (2005). Large scale evaluations of children's mental health services: The Ft. Bragg and Stark County studies. In R. Steele & M. Roberts (Eds.), *Handbook of mental health services for children, adolescents, and families* (pp. 371-386). New York: Springer, doi: 10.1007/0-387-23864-6_24.

Bickman, L., Summerfelt, W. T., & Noser, K. (1997). Comparative outcomes of emotionally disturbed children and adolescents in a system of services and usual care. *Psychiatric Services, 48*(12), 1543–1548.

Birkeland, S., Murphy-Graham, E., & Weiss, C. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (D.A.R.E.) program. *Evaluation and Program Planning, 28,* 247–256. doi:10.1016/j.evalprogplan.2005.04.001

Brown, C. H., & Liao, J. (1999). Principles for designing randomized preventive trials in mental health: An emerging developmental epidemiology paradigm. *American Journal of Community Psychology, 27*(5), 673–710.

Bufe, C. (1998). *Alcoholics Anonymous: Cult or cure* (p. 90). Tuscon, AZ: See Sharp Press.

Chen, H. (1998). Theory-driven evaluations. In A. J. Reynolds & H. J. Walberg (Eds.), *Advances in educational productivity, vol. 7: Evaluation research for educational productivity* (pp. 15–34). Greenwich, CT: Elsevier Science/JAI Press.

Chen, H., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7*(3), 283–302.

Cook, J. R., & Kilmer, R. P. (2004). Evaluating systems of care: Missing links in children's mental health research. *Journal of Community Psychology, 32*(6), 655–674. doi: 10.1002/jcop.20024

Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology, Research and Practice, 17*(2), 136–137.

Drotar, D. (Ed.). (2000). *Handbook of research in pediatric and clinical child psychology: Practical strategies and methods* (pp. 77–96). Dordrecht, Netherlands: Kluwer Academic Publishers.

Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice, 15*(1), 1–15.

Friedman, R. M. (1996). The Fort Bragg study: What can we conclude? *Journal of Child and Family Studies, 5*(2), 161–168.

Gandhi, A. G., Murray-Graham, E., Petrosino, A., Chrismer, S. S., & Weiss, C. H. (2007). The devil is in the details: Examining the evidence for "proven" school-based drug abuse prevention programs. *Evaluation Review, 31,* 43–74, doi: 10.1177/0193841X06287188.

Hendrick, T. E., Bickman, L., & Rog, D. J. (1993). *Applied research design: A practical guide.* Newbury Park, CA: Sage Publications.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., & Swenson, C. C. (2006). Methodological critique and meta-analysis as Trojan horse. *Children and Youth Services Review, 28,* 447–457, doi:10.1016/j.childyouth.2005.07.001.

Henry, G. T. (2003). Influential evaluations. *American Journal of Evaluation, 24,* 515–524.

Hernandez, M., & Hodges, S. (2003). *Crafting logic models for systems of care: Ideas into action. Making children's mental health services successful,* *Vol. 1.* Tampa, FL: University of South Florida, The Louis de la Parte Florida Mental Health Institute, Department of Child and Family Studies. Retrieved October 13, 2008 from http://cfs.fmhi.usf.edu/TREAD/CMHseries/IdeasIntoAction.html.

Korb, L. J. (September 8, 2000). *Testimony. National missile defense: Test failure and technology development.* Council on Foreign Relations. Retrieved January 9, 2007, from www.cfr.org/publication/3800/national_missile_defense.html.

Littell, J. H. (2005). Lessons from a systematic review of effects of multisystemic therapy. *Children and Youth Services Review, 27,* 445–463, doi:10.1016/j.childyouth.2004.11.009.

Littlell, J. H. (2006). The case of multisystemic therapy: Evidence or orthodoxy? *Children and Youth Services Review, 28,* 458–472, doi:10.1016/j.childyouth.2005.07.002.

Millar, A., Simeone, R. S., & Carnevale, J. T. (2001). Logic models: A systems tool for performance management. *Evaluation and Program Planning, 24,* 73–81, doi: 10.1016/S0149-7189(00)00048-3.

OASIG (1996). Why do IT projects so often fail? *OR Newsletter, 309,* 12–16.

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113*(3), 553–565.

Rosenbaum, D. P., & Hanson, G. S. (1998). Assessing the effects of school-based drug education: A six-year multilevel analysis of project D.A.R.E. *Journal of Research in Crime and Delinquency, 35*(4), 381–412.

Rossi, P. (1987). The iron law of evaluation and other metallic rules. *Research in Social Problems and Public Policy, 4,* 3–20.

Salzer, M. S., Bickman, L., & Lambert, E. W. (1999). Dose-effect relationship in children's psychotherapy. *Journal of Consulting and Clinical Psychology, 67*(2), 228–238.

Scargle, J. D. (2000). Publication bias: The "file-drawer" problem in scientific inference. *Journal of Scientific Exploration, 14*(1), 91–106.

Scherer, R. W., Dickersin, K., & Langenberg, P. (1994). Full publication of results initially presented in abstracts: A meta-analysis. *Journal of the American Medical Association, 272,* 158–162.

Seaman, M. A., & Serlin, R. C. (1998). Equivalence confidence intervals for two-group comparisons of means. *Psychological Methods, 3*(4), 403–411.

Senturia, Y. D., McNiff-Mortimer, K., Baker, D., Gergen, P., Mitchell, H., Joseph, C., et al. (1998). Successful techniques for retention of study participants in an inner-city population. *Controlled Clinical Trials, 19,* 544–554.

Shields, P. G. (2000). Publication bias is a scientific problem with adverse ethical outcomes: The case for a section for null results. C*ancer Epidemiology, Biomarkers and Prevention, 9*, 771–772.

Sobell, L. C., Cunningham, J. A., & Sobell, M. B. (1996). Recovery from alcohol problems with and without treatment: Prevalence in two population surveys. *American Journal of Public Health, 86*(7), 966–972.

The Standish Group. (1995). *The chaos report*. Retrieved January 5, 2007, from www.standishgroup.com/sample_research/chaos_1994_1.php.

Stegner, B. L., Bostrom, A. G., & Greenfield, T. K. (1996). Equivalence testing for use in psychosocial and service research: An introduction with examples. *Evaluation and Program Planning, 19*(3), 193–198.

Stern, J. M., & Simes, R. J. (1997). Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal, 315*, 640–645.

Sundell, K., Hansson, K., Lofholm, C. A., Olsson, T., Gustle, L., & Kadesjo, K. (2008). The transportability of multisystemic therapy to Sweden: Short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology, 22*, 550–560, doi: 10.1037/a0012790.

U.S. Food and Drug Administration. Office of Management. (2006). B*udget formulation and presentation: FY 2007 budget*. Retrieved January 5, 2007, from www.fda.gov/oc/oms/ofm/budget/2007/TOC.htm.

U.S. Small Business Administration (2007). *Get ready: Is entrepreneurship for you?* Retrieved January 9, 2007, from www.sba.gov/smallbusinessplanner/plan/getready/SERV_SBPLANNER_ISENTFORU.html.

Weiss, C. H., Murhpy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation, 26*(1), 12–30, doi: 10.1177/1098214004273337.

Westinghouse Learning Corporation (1969). *The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development*. Athens, OH: Ohio University.

Wilson, J. D. (1934). *The manuscript of Shakespeare's Hamlet and the problems of its transmission*. New York: Macmillan.

# 10

# The Influence of Research on Policy and Practice: Lessons from Studies of Asset Building and Low-Income Families

*Michael Sherraden and Trina Williams Shanks*

> We should do more to help all working families save and accumulate wealth. That's the idea behind the individual development accounts, the IDAs. I ask you to take that idea to a new level, with new retirement savings accounts that enable every low- and moderate-income family in America to save for retirement, a first home, a medical emergency, or a college education.
> —William Jefferson Clinton, 2000

A long-standing debate in the social sciences asks whether scholarship should be basic to build knowledge or applied in current issues and challenges in human interaction. In large measure, this debate derives from an overdrawn dichotomy. The differences between basic and applied social sciences are not very clear, and the debate, although frequently impassioned, is in the end rather unproductive (see an excellent and still timely statement on this topic by Rossi, 1980).

To be effective in the real world, applied social research requires purposeful development of an idea along two fronts—the academic and the

---

This chapter is based on part on several previous publications, including Sherraden (2000, 2007a), Sherraden, Slosar, & Sherraden (2002), and Williams Shanks (2005a, 2005b, 2005c).

applied. At times the activities along these two fronts overlap; at other times they diverge sharply. The skills and activities required for operating on the two fronts are sometimes similar but often are very different. Success on one front can sometimes contribute to success on the other front, but often it does not. To be successful in the long run, the applied social researcher must operate simultaneously, or at least intermittently, on both fronts. The process is not neat and linear. The involved scholar may be working on any one of several steps at a particular time. The challenge is to cover all of the bases all the time, though this ideal can never be fully achieved.

Success in influencing policy and practice in applied social research follows directly from how the research is conceived and carried out. It cannot be an afterthought. All the practical strategies for influencing policy and practice follow from asking a timely and important question, specifying theory in a form that has applied relevance, gathering and analyzing data with impeccable care, and reporting in multiple formats that can be used by different audiences. If research is well conceived, then news articles, speeches, testimony, advising, op-eds, legislation, and other practical tools emerge almost as a matter of course. At that point, the greatest challenge is not how to have applied influence, but instead how to build partnerships so application does not take up all of the researcher's time. The researcher's time should be devoted primarily to scholarship.

This chapter is based upon our research in a particular area—asset building—as an illustration of building knowledge and at the same time informing and influencing policy and practice. We hope this detailed example provides lessons that can inform other social research and application projects. We use a detailed example because this may be more engaging and more informative than a dry list of prescriptions detached from an illustrative context.

## Background: Assets and Income

How does a body of applied social research begin? Of course, there are many origins and pathways, but quite often a key publication and/or a particular scholar will initiate thinking and research. In the present example, in 1991 a book entitled *Assets and the Poor* proposed asset building as a strategy for economic and social development of impoverished families and communities and suggested individual development accounts (IDAs) as universal and progressive savings accounts. IDAs were proposed (1) as savings accounts for everyone; (2) to be started as early as birth; (3) with savings matched for the poor, up to a cap; (4) with multiple sources of matching deposits; (5) to be accompanied by financial education; and (6) to include savings to be used for investments in homes, education, business capitalization, or other development purposes (Sherraden, 1991). What has become of this idea? How did it happen? What was the role of research, and what might we learn from this example?

First, let us consider a brief summary of the context and approach. A shift to individual asset accounts in social policy is occurring in many countries around the world, and policy discussion is emerging (OECD, 2003; USAID, 2004; World Economic Forum, 2003). Examples in the United States are the expansion of 401(k) retirement plans, IRAs and Roth IRAs, 529 College Savings Plans, Medical Savings Accounts, and many others. As a global phenomenon, this is a near revolution in public policy and is especially pronounced in retirement policy. These trends appear likely to continue.

Unfortunately, in the United States and most other countries, growing public expenditures for asset building (mostly through the tax system) are highly regressive. More than $300 billion is spent annually in the United States for asset building in homes, investments, and retirement accounts, and more than 90% of this goes to households with incomes greater than $50,000 per year (Corporation for Enterprise Development, 2004; Howard, 1997; Seidman, 2001; Sherraden, 1991).

At the same time, the poor do not have the same opportunities and subsidies for asset accumulation. The reasons are threefold. First, the poor are less likely to own homes, have investments, or have retirement accounts, where most asset-based policies are targeted. Second, the poor have little or no tax incentives, or other incentives, for asset accumulation. Third, asset limits in means-tested transfer policies discourage saving by the “welfare poor,” and probably also the working poor (Nam, 2008; Powers, 1998; Ziliak, 1999). In effect, the United States and many other countries have a dual policy consisting of asset-building subsidies for the nonpoor and asset-building disincentives for the poor. This system is both unfair and counterproductive. If asset building is how individuals, families, and communities develop, then a sensible public policy would promote asset building for all because this would have the greatest payoff in social and economic development. Given these conditions, we may ask: Why not asset accumulation by the poor? And thus the rationale for this research.

The use of income as sole definition of poverty and well-being is now being questioned in many quarters. Sen (1993, 1999) and others are looking toward capabilities. Sen uses the term capabilities to refer to what people are able to be and do. Asset-based policy can be seen as part of this larger discussion. Asset holding is one pathway and measure of long-term capabilities. As public policy, asset building is a form of social investment (Midgley, 1999; Sherraden, 1991). Asset-based policy would shift social policy from an almost exclusive focus on maintenance of households toward development of individuals, families, and communities. In this sense, asset-based policy is an explicit complement to income-based policy (Sherraden, 1991). The goal of asset-based policy should be *inclusion*. Inclusion in this sense means that policy should (1) bring everyone into asset-based policy; (2) make asset-based policy lifelong and flexible; (3) provide at least equal public subsidies for the poor in dollar terms; and (4) achieve adequate levels of asset accumulation, given the purposes of the policy. The reader may note here that the research program in asset building has a strong motivational foundation based on evidence and reasoning as well as a clear image of an ideal outcome. Although this ideal is unlikely to be reached, the context is essential in framing research questions and design.

## Policy and Practice Innovations

It is not possible to consider policy and program applications without a deep understanding of the applied context. Applied social research almost by definition cannot be an "ivory tower" activity. Even to ask a useful and productive question, engagement is required.

When IDAs were first proposed (Sherraden, 1988, 1991), asset building was considered by many social policy experts to be an odd idea and perhaps misguided. Many argued that poor people could not—and perhaps should not—save. Others were concerned that funding for asset building might compete with funding for income support. Others have been against individual asset accumulation in favor of more community-oriented approaches. These are all well-meaning responses, and we have taken them all seriously. To greatly oversimplify, our position has always been that savings and asset building should be voluntary, and that the poor, like the nonpoor and as a matter of fairness, should have subsidized opportunities to do so. Also, we have sometimes had to point out that the nonpoor are not accumulating their 401(k) or home ownership assets in communal forms, and ask why this should be imposed on the poor? If people choose to be communal, they can always do so, but should communal ownership be imposed by public policy? These positions seem reasonable to us, and over time we find ideological opposition declining as evidence accumulates that the poor can and do accumulate assets and use them to improve their lives. What is the lesson in all this? Probably it is that no new policy or practice direction will be without opposition from current policy and practice patterns and established interests. Resistance can be expected. The applied researcher should keep her eye on the core reasoning, data, results, and implications. Over time, we have found that actual information slowly wins out over entrenched patterns and ideologies.

Despite initial resistance, the idea of asset building has taken hold in many quarters across the political spectrum. The Center for Social Development (CSD)[1] at Washington University in St. Louis was advising asset building in many applied settings. Overall, since asset building and IDAs were proposed, considerable progress has been made in policy and practice in the United States. In all cases this has required partnerships with many state legislatures and policy organizations in Washington. A research center in a university does not have the capacity—and in our view should not even try to have the capacity—to carry out policy change without key partners.

For example, an important, though largely unheralded, policy change has been the easing—and sometimes elimination—of welfare asset limits in nearly all states since 1991. Before the early 1990s, there had been almost no focus on easing or eliminating asset limits in means-tested programs. This was not an agenda on either the "left" or the "right," and little research on the topic existed. Due to numerous state-level efforts, this has now changed, and progress has also been made at the federal level. Corporation for Enterprise Development (CFED)[2] and the New America Foundation have led regulatory efforts to exempt federally supported IDAs from asset limits in any means-tested program. Altogether, this quiet policy change is profound in its implications. If greater assets can be accumulated without penalty, then asset building by the poor is much more possible.

Individual development accounts were included as a state option in the 1996 Welfare Reform Act. The federal Assets for Independence Act, the first public IDA demonstration, became law in 1998. Other legislation to extend IDAs has been debated in Congress over the past decade. Many other types of asset-building policies also have been introduced to promote home ownership, retirement saving, and other key policy objectives (New America Foundation, 2007). With CFED, CSD has co-planned seven national conferences on IDAs and asset building. In these large meetings, research is delivered and discussed directly with practitioners. These planned interactions enable practitioners to understand the value and importance of research, and also point to new research questions that arise from practice. It would be hard to imagine a successful applied research agenda in the absence of such interactions.

Center for Social Development has been particularly active in networking among state policy leaders in IDAs and asset building, organizing and hosting five national conferences that have focused on state policy. Nearly all 50 states have adopted some type of IDA policy (Edwards & Mason, 2003) though most programs started as a demonstration and are not universal (no state has a statewide IDA program). Initially, IDAs were greeted with skepticism by many state legislators, who were convinced that poor people could not save, or that it would take too long for the poor to accumulate the funds necessary to invest in assets such as homes, businesses, and post-secondary education. Moreover, in the early years of IDA policy development, some policy makers perceived IDAs as a threat to income maintenance programs, believing that IDAs would compete for the same resources. Research on IDAs suggests that these concerns are largely unfounded. Research has shown that states have not diverted cash assistance dollars to IDAs (Edwards, 2005). Research on state IDA policy suggests that this strategy has helped institute a new policy focus on building assets and underscored the importance of helping families accumulate wealth (Warren & Edwards, 2005). In the long tradition of American federalism, the states continue to serve as policy incubators for IDAs and other asset-building policies (Edwards & Mason, 2003).

Considerable policy progress has occurred outside the United States. About the same time as President Clinton (2000) highlighted IDAs in a State of the Union Address, a serious discussion of asset-based policy was underway the United Kingdom (Kelly & Lissauer, 2000; Nissan & LeGrand, 2000). Prime Minister Tony Blair proposed a universal and progressive Child Trust Fund for all children in the United Kingdom. He also proposed a demonstration of a Saving Gateway, a matched saving program for the poor modeled after IDAs (Blair, 2001). Discussion of asset building in the United States, especially CSD research on IDAs, was a strong influence on the UK proposals, including the first major policy speech on asset building in London by David Blunkett (2000), and CSD co-hosted an international conference on this topic in London in January 2001. It would be hard to overstate the importance of research data in opening the door and informing these policy discussions.

Center for Social Development has planned and hosted several major international meetings on asset building and participated in many others. The CSD team has continued to consult on the Saving Gateway and Child Trust Fund in the United Kingdom with the Offices of the Prime Minister and Chancellor of the Exchequer (H. M. Treasury, 2001, 2003; Kempson, McKay, & Collard, 2003, 2005; Paxton, 2003; Sherraden, 2002), Family Development Accounts in Taipei (Chen, 2003), IDAs and "Learn$ave" demonstration in Canada (Boshara., Cramer, & Sherraden, 2007), and matched savings programs for the poor in Australia, Uganda, Peru, and elsewhere. At present, CSD has applied research projects in Uganda and western China; is advising governments in Korea, Hong Kong, and Indonesia; and with several international partners, is planning a test of children's development accounts in up to six developing countries in Asia and Africa.

## The Role of Research in Policy Innovation

What happened, during the 1990s, to move the idea of asset building into national and international policy and practice? Primarily, it was research. This may come as a surprise to some (in all honestly, we continue to be a little amazed by this ourselves, and do not take it for granted). Too often research reports and publications are thought of as "collecting dust on shelves," but in the present case, research has been a primary mover of policy and practice. How did this come about?

During the last part of the 1990s CSD, in partnership with CFED (a self-described policy "do tank"), had designed, secured funds, and implemented a demonstration and research of IDAs around the country. CSD led the research. The early data showed that low-income people, at that time, had average net savings of $33 per month in IDAs. Moreover, controlling for other factors, the poorest IDA participants were saving about as much as others, and saving a higher proportion of their income (Schreiner & Sherraden, 2007). IDA research results also stimulated a policy discussion in London, with extensive exchange and consultation from CSD from 2000 onward, leading to Blair's adoption of the Child Trust Fund and Saving Gateway.

The basic point cannot be overstated—systematic data showing that very poor people can save and accumulate assets in IDAs has been enormously influential in policy and practice. Today the concept of asset building has moved from the margin to the mainstream. The terms "asset building," "wealth creation," "stakeholding," and "ownership society" have sprung up in both domestic and international policy discussions in developed as well as developing countries.

To be successful, research questions must be carefully selected, and research plans must be designed with policy and practice relevance in mind. Both short-term and long-term horizons are required. Research questions should have immediate policy and practice relevance. A single study is seldom enough; the aim should be to carry out a body of work that is substantial enough to build a foundation for policy and practice into the future.

## From Research to Policy and Practice in Individual Development Accounts

What research agenda has led to policy and practice impacts? The most important applied research initiative on IDAs has been the American Dream Demonstration (ADD), funded by a consortium of 11 foundations. ADD was among the largest policy demonstrations in the country during its duration, with 14 IDA demonstration sites across the country, a 4-year demonstration (1997–2001) and 6-year research period (to 2003), and multiple research methods.

This intensive research agenda has yielded informative results regarding many aspects of IDAs and generated major influence on public policy. Implementation assessment informed many starting IDA programs. Case studies of participants have yielded detail and stories for use by the media. Monitoring and periodic reporting on all participants and their savings patterns have been especially important in shaping policy. Using information technology to the fullest, CSD created software called MIS

IDA that could be used to manage IDA programs, track all program data, and download data immediately on savings patterns of all participants in all 14 IDA programs (Johnson & Hinterlong, 1998; Johnson, Hinterlong, & Sherraden, 2001). Regarding impact on practice, MIS IDA became the standard software for operating programs, helping ensure successful accounting, reporting, and other aspects of sound practice.[3]

Again, this has not been ideal. MIS IDA was developed primarily as a research tool and is far from perfect as a program management tool. While some practitioners give rave reviews, many others can provide a list of complaints about MIS IDA. Support for MIS IDA and obtaining quality data from IDA programs has required a telephone support line and systematic quality control checks. The lesson, as in so much of this work, is that challenges are inevitable and perfection is not likely. The core "take away" is that sincere effort and continual problem solving can be good enough to achieve useful results.

Beyond the research itself, the direct payoff of demonstration sites on policy cannot be underestimated. In other words, the mere *presence* of the policy innovation can matter. When senators and representatives know that an IDA program is succeeding in their district, they are much more likely to become advocates for the concept. In this regard, research is not something that happens after policy but is integral and essential to policy development at each step along the way.

## Theory, Research, Evidence: Building Knowledge for Application

What is the role of theory? Is it a kind of decoration for the empirical work and only for scholars to discuss, or does it add to understanding that improves practical application? We definitely think the latter. Application cannot be separated from theory because the "how," when documented with evidence, provides guidance to policy and practice. Two general theoretical statements underlie this work. The first is that saving and asset accumulation are shaped by institutions, not merely individual preferences. In our research on IDAs, CSD has identified the following institutional factors that may affect saving and asset accumulation: (1) access, (2) expectations, (3) information, (4) incentives, (5) facilitation, (6) restrictions, and (7) security (Beverly & Sherraden, 1999; Sherraden, 1991; Sherraden, Schreiner, & Beverly, 2003; Sherraden & Barr, 2005). These constructs are useful in explaining saving outcomes.

For example, we have found in research on IDAs that, controlling for many other factors, the monthly saving target (i.e., expectation) is associated with a 40- to 50-cent increase in average saving for every dollar the target is increased—a huge effect. We have found that financial education (i.e., information) up to approximately 10 hours is associated with increased saving performance, but after 10 hours no effect is noted. Because financial education is expensive, this is important to know. We found that increasing the saving match (i.e., incentive) keeps people saving in the IDA program, but among the "savers"[4] does not increase amounts saved. We found that direct deposit (i.e., facilitation) also keeps people saving but among savers does not increase amounts saved (Schreiner, Clancy, & Sherraden, 2002; Schreiner & Sherraden, 2007).

We have found that IDA participants see the program as an opportunity (i.e., access) that they would not otherwise have because few are offered retirement plans at work. In a focus group, one potential IDA participant insightfully said "This is like a 401(k)—only for us." We found that IDA participants like the fact that their matched saving account is "off limits" and can be used only for specific purposes (i.e., restrictions), even though this is contrary to mainstream economic theory, which assumes that people prefer as much choice as possible (Sherraden, McBride, Hanson, & Johnson, 2005a). These and other results from IDA research have direct relevance for saving policy, program, and product design.

One of the most important findings in ADD is that, controlling for many individual and program variables, participant income was only weakly associated with saving outcomes; that is, the poorest participants saved about as much as those who were not as poor and saved a higher proportion of their income (Schreiner et al., 2002; Schreiner & Sherraden, 2007). This finding suggests that saving by the very poor is not determined solely by income levels and should not be dismissed in public policy.

The second theoretical statement is that assets have multiple positive effects, not merely deferred consumption. To take one example, home ownership creates financial equity in housing as well as more stable and more committed citizens. Theory regarding effects of asset holding—when specified, tested, and supported by evidence—has the potential to provide a solid rationale for inclusive asset-based policy. The possible effects of asset holding are to (1) improve household stability; (2) create orientation toward the future; (3) stimulate enhancement of assets; (4) enable focus and specialization; (5) provide a foundation for risk taking; (6) increase personal efficacy; (7) increase social connectedness and influence; (8) increase political participation; and (9) enhance the well-being of offspring (Sherraden, 1991). A broad range of research in economics, sociology, political science, anthropology, and social work provides evidence generally in support of these propositions (Page-Adams & Sherraden, 1997; Scanlon & Page-Adams, 2001), although the number of rigorous policy tests is still modest.

In a test of assets on multiple outcomes, Yadama and Sherraden (1996) use the *Panel Study of Income Dynamics* (PSID) and simultaneous equation modeling to test alternative theories within the same study. The focal explanation is that assets lead to positive attitudes and behaviors. The two alternative explanations are that (1) positive attitudes and behaviors lead to assets and (2) income leads to positive attitudes and behaviors. All three explanations are supported to some extent in the analysis, but the focal explanation has the strongest support. The findings that assets lead to positive attitudes and behaviors, and positive attitudes and behaviors lead to assets, may be a glimpse of a virtuous cycle in which household development is a reinforcing feedback loop. Arguably, the most efficient use of public policy is to find such cycles and support them.

Experimental results from ADD indicate that, compared with a randomly assigned control group, IDA participants increased their rate of home ownership and the value of real and total assets. Positive effects appear to be stronger for African-Americans (perhaps because past practices have discriminated against African-Americans in home ownership, leading to greater demand). The IDA program may not have increased net worth over the period of the study, though this is not yet clear (Grinstein-Weiss & Sherraden, 2005; Mills, Patterson, Orr, & DeMarco, 2004; Mills, Gale, Patterson, R., & Apostolov, 2006). Nor do we yet know if these results will persist over time, increase, or decrease and follow-up research would be very informative in this regard

While all of this may sound ideal in summary, in reality the research has been far from smooth and fraught with innumerable pitfalls and shortcomings. Applied social research by its very nature is a great challenge in design, methods, and data quality. Despite the very best intentions and efforts, neither the process nor the results can be ideal. Shortcomings are inevitable and must be acknowledged. Results sometimes do not support hypotheses and must be stated objectively. Over the course of this research program, we have learned—again somewhat to our surprise—that full transparency in research shortcomings and objectivity in reporting results have very little downside. Both academic and policy audiences value quality and honesty in scholarship and reporting.

## A Promising Direction for Asset-Based Policy: Children's Development Accounts

How do we move from a beginning policy discussion to a large-scale policy? Perhaps the most promising pathway in this case is by beginning with children. Universal and progressive accounts for all children at birth have been proposed in the United States by Sherraden (1991), Lindsey (1994), Boshara and Sherraden (2003), Cramer (2004), and Goldberg (2005).[5] Children's development accounts (CDAs) may be a promising pathway to inclusive asset building in the United States. The United States is one of the few economically advanced nations without a children's allowance (monthly cash payment to all families with children). The average children's allowance in Western Europe is 1.8% of gross domestic product (GDP). The United States is unlikely, for ideological and political reasons, to adopt a children's allowance, but a CDA is much more likely. Even 0.1% of U.S. GDP today would be enough for a $3000 "start in life" account for every newborn (Curley and Sherraden, 2000). A visionary and bipartisan ASPIRE Act, which would create a savings account for every newborn in the United States, has been introduced and deliberated in the Congress since 2004.[6]

The potential of CDAs as a long-term pathway to inclusive asset building may be great because (1) lifetime accumulation and compounded earnings will lead to greater asset accumulation; (2) having an account from birth will likely create positive psychological and behavioral effects for both parents and children; (3) important reasons exist to save for education and home ownership, in addition to retirement (education and home ownership are ultimately retirement strategies as well); and (4) newborns are in some ways more politically appealing than adults. Regarding the last point, investing in children can be a bipartisan effort even in partisan times (note, for example, the bipartisan support for the ASPIRE Act and other current proposals for CDAs).

Looking to the future, CDAs also may have appeal in developing countries and for international aid. Although possibly considered farfetched today, when information technology is developed enough an account could be generated for every newborn on the planet. No other single strategy would have a greater impact on economic development. Accounts for every child could receive international aid that goes directly to children—avoiding dictators, mismanagement, and corruption.

## CDA Research and Demonstration: The Seed Initiative

How can applied research inform and influence a potential universal policy for children in the United States? The next major research project in asset building is a large-scale test of CDAs. The Ford Foundation and several other foundations are now in the process of demonstrating and testing an inclusive CDA in the form of the Saving for Education, Entrepreneurship, and Downpayment (SEED) initiative. SEED is a demonstration and research partnership among CFED, CSD, the New America Foundation, the Institute for Financial Security of the Aspen Institute, Research Triangle Institute, and other organizations. The goal of SEED is to model, test, and inform a universal CDA policy for the United States.

At the end of 7 years, we hypothesize positive impacts of SEED savings for parental attitudes and behaviors related to education, cognitive and educational development of children and, within the 7-year window of the study, children's educational achievement. These hypotheses are founded in theory and some (though varying) empirical support. The key effects to be tested in the SEED experiment over the first 7 years (2007–2013) are savings for children's education; total household savings; other household assets, liabilities, and net worth; parents' financial knowledge; children's financial knowledge; parents' aspirations for children; children's aspirations, especially for education; children's cognitive development; children's socio-emotional development; and children's preschool and early school performance.

Saving for Education, Entrepreneurship, and Downpayment was launched in 2004 as a national policy, practice, and research initiative to test the efficacy of offering matched asset-building accounts for children and youth. A 10-year research endeavor is taking place in 12 community sites. The initiative has many components: different locations, multiple ages, and multiple account structures. The research design attempts to capture all the nuances of the initiative and provide knowledge and lessons for a wide audience.

There is good reason to believe that research can assess these effects and that they may matter for long-term development of children. In in-depth interviews with IDA participants and controls, participants say that a major incentive and purpose for saving is the well-being of their children, even though IDAs are not targeted toward uses for children (Sherraden et al., 2005b). There is evidence that very young children can connect savings with going to college, and that saving is positively associated with their aspirations and expectations (Elliott, 2007).

Studies using the PSID to look at the impact of wealth on child developmental outcomes found that, controlling for many other factors, parental wealth is positively associated with cognitive development, physical health, and socio-emotional behavior of children (Williams 2003; Williams Shanks 2007). Consistent with this, Conley (1999) used the PSID to look at the influence of childhood household wealth on adult outcomes. He found that parental wealth in childhood helps predict both high-school graduation and college graduation. Effects of wealth are stronger than the effects of income.

Our perspective is larger than elimination of income poverty in the short term; we are focused on young people reaching their potential in early childhood development and education. This, in turn, may lead to continuing positive outcomes in education, later improved employment and higher incomes, and reduced income poverty in the long term. If hypothesized cognitive and behavioral changes occur within the first 7 years, we anticipate that these will put the child and family on a pathway of sustained improvements in educational performance. Of course, this could turn out not to be so. Based on previous social experiments, no predictable relation has been found between short-term and longer-term results. However, in this case, accounts will continue to be in place *after* seven years; that is, the intervention continues, and therefore short-term gains may not fade away. The fact that this experiment can be revisited, say at ages 12, 18, and 25 years, is a huge added value. Initial research investments will yield long-term returns in knowledge building. Long-term outcomes from policy experiments are uncommon and considered in the policy analysis community to be highly valuable.

As with ADD, the SEED demonstration is generating significant policy interest long before final research results are in. At this writing, at least five bills are in the U.S. Congress for creation of CDAs, each supported by both Republicans and Democrats.

## Informing Public Policy

Now, returning to the larger context, how does all this fit together? How do we use these research findings to inform and achieve an inclusive asset-based policy? In America, we have very well developed financial services and efficient, transparent, and secure financial markets. These markets are a huge national and global resource. In any savings policy, not using private markets for investments would almost be foolish. Nonetheless, the public sector has an essential role in an inclusive savings policy. Although sometimes called "private" or "privatized," asset building in the form of defined contribution individual accounts (the most likely vehicle for this policy) is generated, defined, and regulated by government, often with large public subsidies through the tax system, as in 401(k) plans. In that sense, these are *public* policies. Large-scale, inclusive asset building cannot occur through private corporations or nonprofit organizations. Government will be required for establishing the (1) institutional framework that brings everyone into the asset building and keeps costs low; (2) legal protections and regulation; and (3) resources for inclusive asset building.

The technical capability to create universal asset accounts is rapidly developing. Information technology will one day make it possible to give everyone an account, with instantaneous and secure investment options in any of the financial markets in the world. This technical capacity, one aspect of globalization in the information age, has the potential to sweep the entire planet into social and economic development more completely than has heretofore occurred.[7]

Academically, the knowledge base for how to shape asset-based policy and its likely effects is also developing, as illustrated by research examples in this chapter. More work must be done in specifying and testing theory and drawing policy implications. To keep this in perspective, however, we can already say with confidence that asset holding is likely to have multiple positive effects. Some of the most important effects may be regarding the development of children. Moreover, we believe that institutional factors such as access, information, incentives, and facilitation affect saving and asset accumulation, which have direct policy relevance (as discussed earlier).

Still other considerations are political. Creating an inclusive asset-based policy will require visionary leadership, raising asset building to the level of a long-term national project. This project would be, in the most basic sense, creation of a universal system of accounts, an infrastructure to promote asset accumulation. This is perhaps analogous to creation of a national system of highways to promote transportation. Once the infrastructure is in place, development will occur.[8] Political leaders and planners would have to understand asset building in these expansive terms. Once established, such a policy would likely generate strong political support, such as that for the exceptionally popular Central Provident Fund of Singapore (Sherraden, Nair, Vasoo, Ngiam, & Sherraden, 1995; Vasoo & Lee, 2001).

### Toward a Policy Vehicle: Not Just Saving Products—a Saving Plan

Application requires a policy vehicle. One of the great lessons in this body of work is that it is probably easier to build from an existing policy than create a whole new policy instrument. In the present case, research evidence also strongly indicates that, if saving and asset building are to be inclusive, the policy must be in the form of a *savings plan,* such as a 401(k) or 403(b) plan, the Federal Thrift Savings Plan, or a College Savings (529) plan. Such plans are how most Americans are able to save. Savings plans (contractual savings) have important features that lend themselves to inclusion. These features are centralized and efficient accounting, outreach and education, a limited number of low-cost investment options, low initial and ongoing deposit requirements, automatic deposits, and opportunities to establish other practices and defaults that increase saving performance. These include automatic enrollment, savings matching, matching caps (amount of savings that can be matched), a default low-cost fund, and automatic increases in savings deposits with pay raises. During the payout period, a required minimal annuitization may be desirable for income protection.[9] For these very good reasons the ASPIRE Act calls for a plan structure similar to the Federal Thrift Savings Plan.

At the CSD, we find considerable potential in using College Savings (529) plans as a platform for inclusion in asset building. To be sure, some state 529 plans have high fees and high investment costs; but some state 529 plans keep costs low, have very low deposit requirements, provide outreach to state residents, and match savings for the poorest savers. These state plans or something similar have the potential to be a platform for an inclusive children's savings account (Clancy & Sherraden, 2003; Clancy, Cramer, & Parrish, 2005; Clancy, Orszag, & Sherraden, 2004).

If properly designed as an inclusive and low-cost savings plan, an inclusive asset-based policy would be a large-scale public good in which all citizens could benefit. With this in mind, a major agenda for CSD in the coming years is continuing research on inclusive features of 529 plans so they could be considered a platform for a universal CDA. Why are we doing this? CSD is purposefully taking responsibility for testing accounts as well as a policy vehicle for delivering the accounts. Unless this is done, an enacted policy may be built on an inadequate or dysfunctional platform; public policy is full of such mistakes. Our aim is to inform policy to enhance the likelihood for success.

Of course, not everyone has the same policy vision. Our close partners in think tanks, philanthropic organizations, financial institutions, and elsewhere share a wide range of policy visions. Again, evidence can help to inform these decisions. An agenda at CSD is to study the progressive potential of saving plans and make these results available. At the end of the day, the vigorous policy process in American social policy will sort out evidence and interests and reach a policy formulation. No researcher can be in control of this; we can only inform the process.

## Some Reflections and Lessons for Consideration

What can we gather from this experience? In thinking about the emergence of IDAs and CDAs, many general points could be made concerning policy and practice innovations and influence as carried out from an academic setting. The observations listed here are perhaps among the most important.

*Think creatively and boldly.* Academic work of any kind requires enormous energy and effort. Even the brightest and most efficient scholars can undertake only a few major projects in their entire careers. This being the case, scholars should not choose small. To be sure, being an "incrementalist" and a compromiser in the policy world is necessary, but this practical stance should be in the context of a larger vision. Spending one's talents and energies on small matters that do not connect with and build a larger agenda is a strategic mistake.

*Start with a compelling idea and good research question(s).* Related to the above-mentioned, not all ideas and research questions are of equal value for application to policy and practice. The social or economic issue should be compelling, and the research question(s) should be incisive, easily communicated, built on scientific knowledge and/or practical initiatives, and have potential to address a social or economic challenge. In the case of IDAs and CDAs, several books were published in the late 1980s and early 1990s that documented the extent of wealth inequality and the theorized potential of policies to enable low-income households to build assets.

*Focus on both the problem and the potential solution.* A majority of applied social scientists spend most of their time examining and detailing problems. This activity is overdone. Of course problems must be accurately described; but from a practical standpoint, equally intensive effort should be devoted to examining capacities and potential solutions. This activity is often neglected by social scientists, even those who are in applied and professional areas of inquiry. To be effective in the policy-making process, the involved scholar must have a strong research program in both problems and potential solutions.

*Develop intellectual and applied aspects of the policy innovation simultaneously.* Some people respond to ideas and others respond to concrete proposals. Either way, beginning with one soon involves the other. Applied social researchers should do their homework in both areas and be ready to speak on different ground to different individuals, as necessary in any particular situation. Intellectual aspects of a policy or practice innovation include identification of key questions, knowledge of relevant theory, tests and reformulation of theory, empirical evidence, and a future research agenda. Applied aspects of policy innovation include simple and clear statements of the rationale, general strategy, specific policy applications, and evaluative data.

*When possible, envision and carry out major applied research projects.* In this chapter we have described two very large research projects—ADD (completed) and SEED (now underway). ADD focused on adults and IDAs, whereas SEED focuses on children and CDAs. In addition to testing policy ideas, these large initiatives attract considerable attention from policy makers and also build a practitioner base of persons who understand and can run these programs throughout the country. Thus, on a large scale, the research itself can be used to build capacity.

*Be grounded.* Success at policy and practice innovation is not possible from the ivory tower. Asset-based policy evolved from talking with welfare recipients, and we have continued to work with "real people" at the state and local level when implementing IDAs. Two policy points should be made. First, policy thinking will be much better if it is from the ground up, based in the realities of people's everyday lives. Second, the innovator is in a much stronger political position when able to show evidence of working with real people. At annual IDA conferences, the most impressive and memorable sessions are by IDA participants who simply tell their own stories. Enormous expertise is found in the target population, and applied social researchers make a huge mistake if they ignore it.

*Build partnerships.* Clearly, planning and implementing a large, multisite, multiyear, multi-method demonstration is an enormous undertaking. In so huge a task, partnerships are essential. Ideally, partners include the program implementers along with policy organizations so that the applied scholar can concentrate primarily on research.

*Implement a high-quality demonstration with policy goals in mind.* As a demonstration, ADD explicitly set out to (1) test whether asset building could be implemented within a low-income population; and (2) document how participants were able to save, how much, and impacts of this saving and asset accumulation. The demonstration aimed to create a quality program that delivered what was promised. To do this, practitioners had to gather regularly to ask questions, share concerns, and document successes. In addition to convening and overseeing those running the programs, these gatherings also include organizational partners not involved in the details of running the program. Some partners and staff can be devoted to getting new people interested in the idea, attracting decision makers, and generating greater impact through policy or strategic alliances. Thus, as the work moves forward, progress can continue on building practitioners in the field, improving practice, generating knowledge, and influencing policy.

*Convene, connect, and promote policy and practice networks.* Organize and participate in conferences to connect research to policy and create networks of policy makers and practitioners. CSD and other organizations host numerous conferences for these groups at which research findings are presented and discussed among people who will apply this work.

*Engage in policy advising, legislative testimony, and policy and practice review.* Engaging in policy process can be done in many ways (Sherraden, 2007a, 2007b; Williams Shanks, 2005a, 2005b, 2005c). Legislative forums, speeches in national and international meetings, and individual consultations with policy leaders are required to do this work successfully. In these settings research evidence can be directly woven into policy proposals.

*Speak to a broad political audience and refuse to be partisan.* The primary allegiance of a scholar is to objective knowledge. The political process is naturally partisan, and political actors will almost automatically attempt to connect any "policy expert" with a viewpoint consistent with their political party. In the short term, falling into this partisan pattern would be very easy. However, in the long term, partisan political activity is a disservice to one's independence as a scholar and to policy and practice research. This cannot be overstated. Once associated with a particular political party, a proposal is weakened and the objectivity of the "policy expert" is suspect. The applied scholar must learn to engage vigorously in policy and practice, but never in politics.

*Never sacrifice academic quality.* Above all, applied social researchers must undertake sound scholarship. In the long run, this is what matters most. Policy and practice, and accompanying public relations, mean very little in the absence of a strong academic foundation. Without a strong foundation, the policy or practice innovation is unlikely to last. The applied scholar can and should guard against bias by exposing key ideas to scrutiny, by creating opportunities to be disproved, by including academics with different perspectives, and by proactively inviting criticism. If objectivity is fiercely protected and research methods are sound, the applied scholar brings rich understanding and dedication to the subject matter that can potentially produce better research. We have found that objective inquiry, even when results do not support the original hypotheses, is by far the most productive approach. Honest scholarship with implications is deeply valued, while unfounded advocacy is not.

*Get a foothold, and then go for scale.* Big ideas *do* come in small and medium. A demonstration may cover only a few thousand people, but if well designed it can inform a large-scale policy. In the case of IDAs, the agenda is to show that matched saving for the poor is a sound public investment so that this principle can be incorporated into larger asset-based policies. In fact, asset-based policies are growing rapidly and may become dominant in the twenty-first century. The challenge will be to include those at the bottom. Data from relatively small IDA projects have affected progressive policy proposals in the United States and elsewhere.

*Set the rudder and stay the course.* The applied social researcher should be clear about the key hypotheses and, if supported by research, implications for long-term policy and practice. The first part of this principle is to set the rudder—that is, to know the long-term goal. In the present case, it is to inform a large-scale, inclusive, and progressive asset-based policy. The second part is to stay the course and roll with the inevitable high seas. In applied social research and policy and practice innovation, a scholar and her craft will be tossed about and the boat will take on some water. By definition, if it is innovation, smooth sailing is unlikely. Innovation requires a clearly charted direction, a compass, skill, and the ability to redirect as the data indicate.

## Conclusion

In closing, we reflect briefly on two historical examples of social research that have led to meaningful, positive changes. First is the careful documentation of lynchings of freed African-Americans in the late nineteenth century by Ida B. Wells (Wells-Barnett, 1969). In some respects Wells was inventing applied social research. She documented that the purported crimes that led to lynchings, typically accusations regarding the black lynching victim and a white woman, were often fabricated. Instead, many of the incidents were white responses to perceived competition for property or power. With systematic data in hand, Wells then used her considerable skills in writing and political organizing to bring research results to the attention of national leaders. When American leaders would not listen, she went to Europe and used her data to create international pressure that eventually reduced lynchings in the United States.

As another example, perhaps the greatest single body of policy and practice innovation in U.S. social policy resulted from the extraordinary work of Jane Addams and the women of Hull House, who made contributions in many areas of community life, public health, and social protection (Addams, 1910). Their work was always based in empirical data in the community, and they applied these data not merely to local and state innovations, but to national policy proposals as well. Much of their groundbreaking work influenced federal policies over the next three decades, shaping much of what became known as welfare state policies of the twentieth century.

Applied social researchers who have the far-reaching effects of Wells and Addams come along rarely. But all of us can aim for research that builds relevant knowledge and makes positive contributions to policy and practice. In this regard, we have attempted to draw some lessons for applied social researchers by using the example of asset building. We hope these lessons can be useful in other areas of policy and practice innovation as well. As Wells and Addams exemplified in their careers, research is fundamental to this process.

The key point of this chapter is that application cannot be an afterthought. If research results are neglected and unused, the problem lies not with practitioners and policy makers, but with the researcher. After all, if it is applied research, application is integral to the purpose. Application should follow directly from the research question, theory, design, data collection, analysis, and reporting. To emphasize this point, imagine Ida B. Wells or Jane Addams carrying out a research project and only later asking how the findings might be used in the world. Of course this never occurred—they knew exactly how the findings could be used. None of us is likely to be as great as Wells and Addams, but we can learn from them. A good starting point is simply to ask the following question (one that we regularly pose to our doctoral students): If this study is carried out as proposed and the hypotheses are supported, what exactly will happen? An applied social researcher should be able to predict—based on knowledge of existing conditions and sound reasoning—the desired and potential outcomes. This is the essence of research influence on policy and practice.

## Notes

1 CSD was created in 1994. A major reason for starting the center was to build capacity to carry out research on asset building, which was

becoming an agenda too large for a single researcher. Since that time, CSD has broadened its research to civic engagement, productive aging, community economic development, and other areas, but asset building remains a major program of study.

2 CFED is the organization formerly known as Corporation for Enterprise Development.

3 During this period, MIS IDA software became the most licensed software product of Washington University in St. Louis and also created a significant revenue stream for CSD. CSD did not aim to be in the software business, and we have since transitioned out of it, but this period was productive in many ways for IDA practice capacity and CSD's long-term stability.

4 A "saver" in ADD is someone who had a minimum of $100 in net savings. We use this somewhat arbitrary definition to sort out participants who did very poorly from those who saved more successfully. In fact, most of the "savers" had much more than $100 in net savings.

5 Discussions of CDAs in the United States go back at least to the George H.W. Bush administration. Goldberg was a proponent of CDAs in the Bush senior administration. At the request of the Bush White House, Sherraden created a plan for a CDA with an initial deposit of $1000 for all children in the United States.

6 An important background paper for what became the ASPIRE Act was written by Reid Cramer (2004). Ray Boshara and his team at the Asset Building Program at the New America Foundation have been instrumental in organizing the introduction of the ASPIRE Act. New America Foundation has been an invaluable policy partner in this work.

7 This optimistic statement does not consider some large political and regulatory challenges. Global access to financial services must solve innumerable challenges of nationalism and central banking. Still, in the long run, global access to internet-based financial services seems likely, in the same way that access to any form of information seems likely.

8 For this insight on universal asset accounts as an overall infrastructure and public good, we are indebted to Fred Goldberg.

9 These plan features are expressions of institutional constructs for saving, as discussed above.

---

## POINTS TO REMEMBER

- Applied social research requires purposeful development of an idea along two fronts: the academic and the applied
  - Research must be well conceived to lead to news articles, speeches, testimony, advising, op-eds, and legislation.

  - Build partnerships for application with legislatures and policy organization.
  - Research data opens the doors and informs policy decisions.
  - Design research questions and plans with policy and practice relevance in mind.
  - Build a body of work substantial enough to be foundation for policy and practice.
  - Be transparent in research shortcomings and objectively report results.
- Lessons
  - Start with a compelling idea and good research question(s).
  - Focus on both the problem and the potential solution.
  - Develop intellectual and applied aspects of the policy innovation simultaneously.
  - When possible, envision and carry out major applied research projects.
  - Be grounded.
  - Build partnerships.
  - Implement a high-quality demonstration with policy goals in mind.
  - Convene, connect, and promote policy and practice networks.
  - Engage in policy advising, legislative testimony, and policy and practice review.
  - Speak to a broad political audience and refuse to be partisan.
  - Never sacrifice academic quality.
  - Get a foothold, and then go for scale.
  - Set the rudder and stay the course.

## REFERENCES

Addams, J. (1910). *Twenty years at Hull House*. New York: Macmillan Co.

Beverly, S. G., & Sherraden, M. (1999). Institutional determinants of saving: Implications for low-income households and public policy. *Journal of Socio-economics*, *28*, 457–473.

Blair, T. (2001). *Savings and assets for all*, speech. London: 10 Downing Street, April 26.

Blunkett, D. (2000). *On your side: The new welfare state as an engine of prosperity*, speech, London: Department of Education and Employment, June 7.

Boshara, R., Cramer, R., & Sherraden, M. (2007). The politics of aspirations: Creating inclusive asset-based policies in the US. In J. Robson &

P. Nares (Eds.), *Wealth and well-being: New directions for social policy in Canada* (pp. 1–18). North York, Ontario: SEDI.

Boshara, R., & Sherraden, M. (2003). For every child, a stake in America. *New York Times*, July 23.

Chen L. C. (2003). Developing family development accounts in Taipei: Policy innovation from income to assets. *Social Development Issues, 25*(1&2), 106–117.

Clancy, M., Cramer, R., & Parrish, L. (2005). *Section 529 savings plans, access to post-secondary education, and universal asset building.* Washington, DC: New American Foundation.

Clancy, M., Orszag, P., & Sherraden, M. (2004). *College savings plans: A platform for inclusive savings policy?* St. Louis: Center for Social Development, Washington University in St. Louis.

Clancy, M., & Sherraden, M. (2003). *The potential for inclusion in 529 savings plans: Report of a survey of states*. St. Louis: Center for Social Development, Washington University in St. Louis.

Clinton, W. J. (2000). *State of the Union address*. Washington, DC: U.S. Executive Office of the President.

Conley, D. (1999). *Being black, living in the red: Race, wealth, and social policy in America*. Berkeley, CA: University of California Press.

Corporation for Enterprise Development. (2004). *Hidden in plain sight: A look at the $335 billion federal asset-building budget*. Washington, DC: Corporation for Enterprise Development.

Cramer, R. (2004). *Accounts at birth: Creating a national system of savings and asset building with children's savings accounts*. Washington, DC: New America Foundation.

Cramer, R., Parrish, L., & Boshara, R. (2005). *Federal assets policy report and outlook*. Washington, DC: New America Foundation.

Curley, J., & Sherraden, M. (2000). Policy lessons from children's allowances for children's savings accounts. *Child Welfare, 79*(6), 661–687.

Edwards, K. (2005). *Temporary Assistance to Needy Families (TANF) and Individual Development Accounts (IDAs): A good match?* Policy report 05-02. St. Louis: Center for Social Development, Washington University.

Edwards, K., & Mason, L. M. (2003). State policy trends for individual development accounts in the United States, 1993–2003, *Social Development Issues, 25*(1&2), 118–129.

Elliott, W. (2007). Examining minority and poor youth's college aspirations and expectations: The potential role of college savings (CSD Working Paper 07-07). St. Louis: Washington University, Center for Social Development.

Goldberg, F. (2005). The universal piggy bank: Designing and implementing a system of savings accounts for children. In M. Sherraden (Ed.), *Inclusion in the American Dream: Assets, poverty, and public policy*. New York: Oxford University Press (pp. 303–322).

Grinstein-Weiss, M., & Sherraden, M. (2005). Racial differences in saving performance in individual development accounts. *Journal of Income Distribution, 13*(3–4), 98–111.

H. M. Treasury. (2001). Saving and assets for all: The modernisation of Britain's tax and benefit system, number eight. London: H. M. Treasury.

H. M. Treasury. (2003). Details of the Child Trust Fund. London: H. M. Treasury.

Howard, C. (1997). *The hidden welfare state: Tax expenditures and social policy in the United States*. Princeton, NJ: Princeton University Press.

Johnson, E., & Hinterlong, J. (1998). *Management information system for individual development accounts*, Version 2.0, software. St. Louis: Center for Social Development, Washington University.

Johnson, E., Hinterlong, J., & Sherraden, M. (2001). New advances in technology for social work education and practice. *Journal of Technology in Human Services, 18*(3/4), 5–22.

Kelly, G., & Lissauer, R. (2000). *Ownership for all*. London: Institute for Public Policy Research.

Kempson, E., McKay, S., & Collard, S. (2003). *Evaluation of the CFLI and Saving Gateway pilot projects*. Bristol, UK: Personal Finance Research Centre, University of Bristol.

Kempson, E., McKay, S., & Collard, S. (2005). Incentives to save: Encouraging saving among low-income households. Bristol, UK: Personal Finance Research Centre, University of Bristol.

Lindsey, D. (1994). *The welfare of children*. New York: Oxford University Press.

Midgley, J. (1999). Growth, redistribution, and welfare: Towards social investment, *Social Service Review, 77*(1), 3–21.

Mills, G., Gale, W. G., Patterson, R., & Apostolov, E. (2006). *What do individual development accounts do? Evidence from a controlled experiment*, working paper. Washington, DC: Brookings Institution.

Mills, G., Patterson, R., Orr, L., & DeMarco, D. (2004). *Evaluation of the American Dream Demonstration*, final evaluation report. Cambridge, MA: Abt Associates.

Nam, Y. (2008). Welfare reform, asset limits, and financial asset accumulation among low-income households, *Social Science Quarterly, 89*(1), 133–154.

New America Foundation. (2007). *Savings accounts for kids: Current federal proposals*. Washington, DC: New America Foundation.

Nissan, D., & LeGrand, J. (2000). *A capital idea: Start-up grants for young people,* policy report no. 49. London: Fabian Society.

Organization for Economic Co-operation and Development (OECD). (2003). *Asset building and the escape from poverty: A new welfare policy debate*. Paris: Author.

Page-Adams, D., & Sherraden, M. (1997). Asset building as a community revitalization strategy, *Social Work, 42*(5), 423–434.

Paxton, W. (Ed.). (2003). *Equal shares? Building a progressive and coherent asset-based welfare policy*. London: Institute for Public Policy Research.

Powers, E. T. (1998). Does means-testing welfare discourage saving? Evidence from a change in AFCD policy in the United States. *Journal of Public Economics, 68*, 33–53.

Rossi, P. (1980). Presidential address to the American Sociological Association. *American Sociological Review, 45*, 889–904.

Scanlon, E., & Page-Adams, D. (2001). Effects of asset holding on neighborhoods, families, and children. In R. Boshara (Ed.), *Building assets* (pp. 25–49). Washington, DC: Corporation for Enterprise Development.

Schreiner, M., Clancy, M., & Sherraden, M. (2002). *Saving performance in the American Dream Demonstration,* research report. St. Louis: Center for Social Development, Washington University.

Schreiner, M., & Sherraden, M. (2005). IDAs and drop-out: Prediction and prevention. *Financial Services Review, 14*(1), 37–54.

Schreiner, M., & Sherraden, M. (2007). *Can the poor save? Savings and asset building in individual development accounts*. New Brunswick, NJ: Transaction.

Seidman, L. (2001). Assets and the tax code. In T. Shapiro & E. N. Wolff (Eds.), *Assets for the poor: Benefits and mechanisms of spreading asset ownership* (pp. 324–356). New York: Russell Sage Foundation.

Sen, A. (1993). Capability and well-being. In M. Nussbaum & A. Sen (Eds.), *The quality of life* (pp. 30–53). Oxford, UK: Clarendon Press.

Sen, A. (1999). *Development as freedom*. New York: Knopf.

Sherraden, M. (1988). Rethinking social welfare: Toward assets. *Social Policy 18*(3), 37–43.

Sherraden, M. (1991). *Assets and the poor: A new American welfare policy.* Armonk, NY: M. E. Sharpe.

Sherraden, M. (2000). From research to policy: Lessons from individual development accounts. *Journal of Consumer Affairs, 34*(2), 159–181.

Sherraden, M. (2002). *Opportunity and assets: The role of the Child Trust Fund,* notes for seminar organized by Prime Minister Tony Blair, 10 Downing Street, and dinner speech with Chancellor of the Exchequer Gordon Brown, 11 Downing Street, London, September 19.

Sherraden, M. (2007a). Assets for all: Toward universal, progressive, lifelong accounts. In J. R. Edwards, M. Crain, & A. L. Kallenberg (Eds.), *Ending poverty in America: How to restore the American Dream* (pp. 151–164). New York: New Press.

Sherraden, M. (2007b). *Overcoming barriers to saving and asset-building for poor people in developing countries,* keynote speech, Consultative Group to Assist the World's Poorest, World Bank, January.

Sherraden, M., & Barr, M. S. (2005). Institutions and inclusion in saving policy. In N. Retsinas & E. Belsky (Eds.), *Building assets, building credit: Creating wealth in low-income communities* (pp. 286–315). Washington, DC: Brookings Institution Press.

Sherraden, M., Nair, S., Vasoo, S., Ngiam, T. L., & Sherraden, M. S. (1995). Social policy based on assets: The impact of Singapore's Central Provident Fund. *Asian Journal of Political Science, 3*(2), 112–133.

Sherraden, M., Schreiner, M., & Beverly, S. (2003). Income, institutions, and saving performance in individual development accounts. *Economic Development Quarterly, 17*(1), 95–112.

Sherraden, M. S., McBride, A. M, Hanson, S., & Johnson, L. (2005a). The meaning of saving in low-income households. *Journal of Income Distribution, 13*(3–4), 76–97.

Sherraden, M. S, McBride, A. M., Johnson, E., Hanson, S., Ssewamala, F., & Shanks, T. (2005b). *Saving in low-income households: Evidence from interviews with participants in the American Dream Demonstration,* research report. St. Louis: Washington University, Center for Social Development.

Sherraden, M. S., Slosar, B., & Sherraden, M. (2002). Innovation in social policy: Researcher, practitioner, advocate, and student collaboration. *Social Work, 47*(3), 209–224.

U.S. Agency for International Development (USAID). (2004). Conference on asset building and sustainable livelihoods, Washington, DC, January.

Vasoo, S., & Lee J. (2001). Singapore: Social development, housing, and the Central Provident Fund. *International Journal of Social Welfare, 10*(4), 276–283.

Warren, N., & Edwards, K. (2005). *State IDA policy survey report* (CSD Policy Report 05-03). St. Louis: Washington University, Center for Social Development.

Wells-Barnett, I. B. [1892, 1894, 1890] (1969). *On lynchings.* New York: Arno Press.

Williams, T. R. (2003). *The impact of household wealth and poverty on child outcomes: Examining asset effects,* doctoral dissertation, St. Louis: Washington University in St. Louis.

Williams Shanks, T. R. (2005a). *Doing it for ourselves: Creating sustainable wealth in African American communities*, Congressional Black Caucus Foundation 35th Annual Legislative Conference, Washington, DC, September 21.

Williams Shanks, T. R. (2005b). *The impact of wealth on child development outcomes*, invited testimony. Washington, DC: Subcommittee on Social Security and Family Policy of the Senate Finance Committee, April 28.

Williams Shanks, T. R. (2005c). *Income poverty, asset poverty, and their consequences for children*, Congressional Black Caucus Foundation Research Policy Roundtable, Washington, DC, November 17.

Williams Shanks, T. R. (2007). The impacts of household wealth on child development. *Journal of Poverty*, *11*(2), 93–116.

World Economic Forum. (2003). Conference on asset-based social policy, Geneva, Switzerland, November.

Yadama, G., & Sherraden, M. (1996). Effects of assets on attitudes and behaviors: Advance test of a social policy proposal. *Social Work Research*, *20*(1), 3–11.

Ziliak, J. P. (1999). *Income transfers and assets of the poor,* discussion paper no. 1202-99. Madison, WI: University of Wisconsin, Institute for Research on Poverty.

# 11

# Disseminating Results and Sharing Data and Publications

*Matthew Owen Howard and Michael G. Vaughn*

> For knowledge itself is power.
> —Francis Bacon, Meditationes sacrae 11, de haeresibus, 1597

> New knowledge is the most valuable commodity on earth. The more truth we have to work with, the richer we become.
> —Kurt Vonnegut, *Breakfast of Champions*, 1973

Successful completion of a field investigation often requires years of diligent effort on the part of many individuals. Despite this substantial investment of time, energy, and money, few researchers adequately plan project dissemination activities (Shanley, Lodge, & Mattick, 1996). All too often, dissemination efforts are ad hoc in nature, haphazardly conducted, poorly informed, and initiated precisely at the point project funds have all but expired (Tarrier, Barrowclough, Haddock, & McGovern, 1999). A study may be interesting in its aims, rigorous in its methods, and important in its implications, but this is all for naught if project findings are not appropriately disseminated to funding entities, scientists, and practitioners. Early in his career, the lead author (M.H.) was a junior member of a research team that conducted a multimillion dollar federally funded project that ultimately yielded fewer than a handful of published empirical reports. This lamentable outcome was distressing and is one that is increasingly less tolerated by funding organizations. For ethical and professional reasons, field researchers should carefully plan project dissemination activities contemporaneously with the development of project aims and methods.

There are several methods by which field research findings are commonly communicated. These include publication of journal articles, monographs, and books and presentation of project results at professional conferences and other meetings (Bourke & Butler, 1996). Numerous complexities attend each of these dissemination activities, many of which are addressed subsequently. We have largely drawn examples from the alcohol and drug abuse literature—our area of expertise—but the lessons conveyed by these cases illustrate more general principles of scientific publication practices.

Dissemination of project findings can be the most gratifying stage in field research, representing the culmination and embodiment of all the work that has been done on behalf of the project. Published results may also offer improved understanding or the possibility of more effective professional practice in the area targeted by the field research. Dissemination activities will have the greatest likelihood of success if they are well-informed and proactively planned. Key considerations pertaining to the dissemination of project findings are discussed subsequently.

## Publication of Scientific Articles, Monographs, and Books

*Prioritizing publication efforts.* Most datasets offer abundant publication opportunities. Thus, the first task that confronts the field researcher is deciding which papers to publish first; that is, which papers are the most substantively important (i.e., critical to the advancement of the scientific literature and/or professional practice in a given area). In grant- or contract-funded research, this decision is relatively straightforward. The papers to be published first are those that flow directly from the specific aims of the funded project. For example, in a recently funded proposal I submitted to the National Institute on Drug Abuse, I wrote that

> The proposed research will constitute the most comprehensive set of population-based investigations of inhalant use heretofore undertaken and will greatly increase the knowledge base upon which new inhalant prevention and clinical interventions can be developed by: 1) identifying risk factors for inhalant use and inhalant use disorders and developing an empirically-based

> taxonomy identifying subtypes of inhalant users, 2) describing the natural history of inhalant use disorders in the general population, including factors predicting onset, escalation, persistence, and stable remission of inhalant use and inhalant use disorders, 3) estimating the prevalence of psychiatric disorders in individuals with lifetime inhalant use/inhalant use disorders and the prevalence of inhalant use/inhalant use disorders among individuals with different psychiatric disorders, and conducting case-control studies evaluating the extent to which inhalant users are at elevated risk for psychiatric disorders compared to other drug abusers and respondents without substance use disorders, 4) evaluating the independent contribution of inhalant use and inhalant use disorders to the development of serious medical conditions, adverse social consequences, and health and mental health-related functional impairments, and 5) describing the substance-related and general medical health care treatment experiences of inhalant users including types of treatment sought, utilization rates, perceived barriers to treatment, and differences between inhalant users and other drug users with regard to these factors.

As in most funded research, each of the five specific aims enumerated here was associated with corollary set of proposed analyses with which to examine questions related to that aim. In grant and contract-funded research, substantial prior effort has been devoted by investigators to identifying key research questions in advance; investigator-identified research questions have then subsequently passed muster with funding body review groups comprised of experts in the area of proposed research. For these reasons, field researchers with funded research should prioritize those publications that derive directly from the specific aims of the funded research.

Once core project papers have been published, investigators can then select subsequent papers on the basis of their novelty, substantive importance to an active area of research other than that directly addressed in the funded proposal, or because the topical area of the paper interests a potential author (often an important consideration if funds are low to pay analysts). Some investigators also choose to write papers that will appeal to a large number of journal publication outlets. This approach enhances the likelihood of eventual publication of project findings.

For example, We are currently preparing a paper addressing asthma inhaler abuse among incarcerated youth, findings of which are suitable for publication in medical, substance abuse, criminology, psychiatric, or adolescence focused journals. It is critical, for future funding purposes and reputational reasons that project investigators attempt to “mine” a study dataset for all core papers and those of obvious importance to related research areas. Once this point is reached, project investigators may then decide to archive the dataset (discussed subsequently) and make it available to other scientists who are interested in publishing reports based on the dataset.

The process of prioritizing papers for publication is similar for field research studies that are not funded by governmental or foundation funding organizations. Almost invariably, these projects have been implemented to answer key research questions. Well-designed studies also include a set of explicit null and/or directional hypotheses to be formally tested in an effort to answer core research questions. Reports addressing these issues should be prepared first and then secondary papers prepared according to the considerations elucidated earlier.

Key project papers, if sufficiently rigorous, should be targeted to top-tier journals as should well-conducted investigations of important issues ancillary to the core aims of the field project in question. For various reasons (e.g., missing data, limited statistical power, less interesting research questions, etc.), some reports derived from a project will be less appealing to journal editors than others. Often, these reports warrant publication in a mid- or lower-tier journal. Thus, field researchers need to be aware of the full gamut of journal publication outlets that can realistically be targeted for project publications. A brief perusal of virtually any top scientist’s CV will reveal a number of such lower- and mid-tier journal publications; it is simply a fact of life that not all research efforts culminate in a uniform set of highly interesting and important findings suitable for publication only in top-tier journals. Many scientists know of several “fall back” journals they can count on to accept their less than stellar articles if all else fail. It is important, however, in most cases to consider only those journals for publication that submit all manuscripts to a formal peer review process. Journal outlets are available across many scientific domains that have no or only a very limited peer review process. These “vanity presses” are generally held in low esteem.

A practice We have adopted in recent years is to target three to five possible journals for a given manuscript submission, ranging from the most prestigious (usually journals with the largest circulations and highest Journal Impact Factors—a measure of the extent to which articles published in a journal are subsequently cited), down to mid- and lower-tier outlets. If, upon submission, reviews of the submitted paper are severely critical and irremediable, then it can reduce stress and wasted time to have several publication outlets already identified to whom one can then submit the manuscript.

One additional issue concerns invited publications. Authors can often pique journal editors' interest in publishing a paper and garner a formal invitation to submit an article. Invited articles may or may not be subjected to a formal peer review process and are generally labeled as such in the table of contents of the journal in which the article is published and often on the published version of the article itself. Occasionally, editors will invite publication of articles without prior contact with the author. We were once contacted by the editor of the journal *Alcohol & Alcoholism* to prepare an invited article addressing an issue we had raised in a letter to the editor we wrote that was published by the *Journal of the American Medical Association*. Authors may also solicit invitations or be invited by journal editors to "guest edit" a particular issue of a journal. When the lead author (M.H.) was editor of *The Journal of Social Service Research*, he was contacted by one project principal investigator who eventually served as a guest editor of a journal issue entirely devoted to presenting findings from his field project. Although invited articles may, in some cases, not undergo formal peer review, they are often positively perceived because their authors are presumably selected on the basis of their expertise in a given area.

Typically, project investigators will publish core papers in professional journals rather than monographs or books. The publication lag associated with monographs and books exceeds, on average, that of journal articles (Hopewell, Clarke, Stewart, & Tierney, 2007). Further, study findings relevant to specific issues can be published sequentially as journal articles, whereas monographs and books typically require integration of a more comprehensive set of findings and therefore take longer to prepare and process for publication. Once the findings of a field research project have been published as a discrete number of professional journal articles, it is not uncommon for such papers to be "collected" and published in book or monograph format. For example, the voluminous findings from the Epidemiologic Catchment Area survey of the prevalence of psychiatric disorders in the United States were originally published in article format (e.g., Blazer, Hughes, & George, 1987; Cottler, Zipp, Robins, & Spitznagel, 1987), but later partly collected in book form (Robins & Regier, 1990). When they are devoted to field research, monographs usually address methodological or technical issues that are likely to have limited readerships. If project findings collected in book form are of general interest value, they may be published by trade, academic, or scientific publishing houses; monographs are usually published by university or academic presses.

Although seldom formally acknowledged, it is usually the case that data-based original reports are more highly regarded by the scientific community than narrative literature reviews, "concept" papers, or descriptive reports (e.g., papers describing implementation of a project). Even within publication categories, there are differences in the esteem in which a given publication is held. Among data-based papers, randomized controlled trials (if otherwise well conducted) are among the most favorably regarded publication types, whereas uncontrolled one-group designs and case studies are regarded least favorably. Project papers published in book or chapter format are often held in less esteem than similar journal articles, because they often have not passed through a formal peer review process. If articles published in book or chapter format are subjected to peer review, the peer review accorded such works is not generally considered as rigorous as that which attends submission to a leading journal (Wolfe, 1990).

*Selecting journals for manuscript submission.* Scientific journals differ dramatically in orientation and with regard to the nature and size of their readerships, publication practices, scientific influence, and prestige. Professional journals publish research relating to the professional practice of medicine, dentistry, pharmacy, nursing, law, social work, and other professions. Disciplinary journals publish results pertaining to academic disciplines such as sociology, economics, and archeology. Specialty journals publish select articles in relatively narrow substantive areas (e.g., suicide and life-threatening behavior, child abuse and neglect).

*Circulation sizes can differ dramatically across journals.* Current subscribers to prominent professional journals such as *Journal of the American Medical Association* or *American Psychologist*, official organs of the American Medical Association and American Psychological Association, number 350,000 and 108,203, respectively (American Psychological Association, 2006; Fontanarosa & DeAngelis, 2007).

Disciplinary journals like the *American Sociological Review* or *American Anthropologist,* which count 11,500 and 13,000 subscribers, can vary widely in circulation size, depending on the relative size of the discipline involved (Advertising and Mailing List Rentals, 2006–2007, p. 3; University of California Press, 2006). Many specialty journals, on the other hand, have limited circulations. For example, only 2 of the 35 English-language journals in the addiction field identified by Arciniega and Miller (1997) surpassed 1800 subscribers. Prominent professional and disciplinary journals often have large circulations and considerable research and practice influence, whereas specialty journals often have far smaller circulations and relatively limited influence. Several years ago, the lead author and his brother (a reference librarian) I published a study comparing the average citation impact of different kinds of alcohol and drug articles published in professional, disciplinary, and specialty journals (Howard & Howard, 1992). We found that alcohol and drug abuse articles published in professional or disciplinary journals were significantly more heavily cited than comparable articles published in substance abuse specialty journals. It is often more difficult to publish a paper addressing a given issue, say alcohol dependence treatment, in a professional or disciplinary journal than in a specialty journal. This is because professional and disciplinary journals try to publish a wide variety of articles to meet the diverse interests of their readership, whereas all of the articles in specialty journals are devoted to a comparatively narrow range of topics. For example, articles published in *The Journal of Studies on Alcohol* have historically focused exclusively on alcohol, whereas only a small fraction of articles published in *The Journal of Consulting and Clinical Psychology* address alcohol-related issues.

Professional journals with large circulations often publish papers that are widely cited and influential, but this is not invariably the case. For example, the journal *Social Work* has a circulation size in excess of 160,000, but a Journal Impact Factor (see following text) of only 0.78. This curious finding is probably due to the nature of the *Social Work* readership, which is largely comprised of master's-level practitioners who may apply the findings to their professional practices, but who do not subsequently cite the articles they read. If field researchers are interested in reaching audiences of practitioners, they may forego publication in journals with larger circulations and relatively high article mean citation counts, if publishing in the practice journal holds the promise of significantly influencing contemporary practice patterns.

What, then, should authors consider in addition to circulation size when they are selecting journals as potential publication outlets for a planned or prepared manuscript? First, it is useful to know how often the journal is published and the average number of articles published per issue. For example, the *Journal of Alcohol & Drug Education*, publishes approximately three issues per year and 18 articles annually, whereas *Alcoholism: Clinical and Experimental Research* and *Addiction*, publish 200 and 134 articles a year, respectively, on 6- and 12-issue annual publication schedules.

In general, publication possibilities are enhanced when journals publish comparatively large numbers of articles on a comparatively frequent basis, but there are notable exceptions to this rule. For example, the *Journal of the American Medical Association* is published 48 times a year and receives more than 5,300 manuscripts annually, but accepts only 8 % of submitted articles. To a significant extent, the degree of competition authors face in publishing their articles is a function of the number of investigators currently active in their research area, the number of available journal outlets for the publication of work in a specific area, and the rigor of the journal to which an article is submitted (i.e., percentage of submissions ultimately accepted for publication). If many investigators are active in a given area, few publication outlets are available, and authors seek to publish in top journals, the odds of acceptance may be low. Fortunately (or unfortunately to those who bemoan the dramatic recent growth in the number of scientific journals), there are currently more than 7,500 scientific journals indexed by the Institute for Scientific Information and available to authors interested in publishing scientific research.

Authors should always make the effort to learn the Journal Impact Factor scores for journals they are considering for publication. Journal Impact Factor scores are calculated by "dividing the number of current year citations to the source items published in that journal during the previous two years" (www.biotechmedia.com/y2005-Impact-Factor-Def.html) and can be obtained by going to the Institute for Scientific Information website http://isiknowledge.com/) and then accessing the Journal Citation Reports link and searching for specific journals or by general areas of interest. Table 11.1 presents Journal Impact Factor scores for 24 specialty journals in the substance abuse field. Reference to the table readily reveals that the journals *Addiction* and *Drug and Alcohol Dependence* have the highest impact factors (4.1 and 3.2, respectively),

**Table 11.1** Journal Impact Factor scores for 24 Alcohol and Drug Abuse Specialty Journals (Ranked from highest to lowest journal impact factor)

| Journal Title | 2006 Impact Factor |
|---|---|
| *Addiction* | 4.1 |
| *Drug and Alcohol Dependence* | 3.2 |
| *Alcoholism: Clinical & Experimental Research* | 2.9 |
| *Psychology of Addictive Behaviors* | 2.2 |
| *Alcohol & Alcoholism* | 2.1 |
| *Alcohol* | 2.0 |
| *Journal of Substance Abuse Treatment* | 2.0 |
| *Journal of Studies on Alcohol* | 1.9 |
| *Addictive Behaviors* | 1.8 |
| *Addiction Biology* | 1.8 |
| *European Addiction Research* | 1.7 |
| *Journal of Addictive Diseases* | 1.4 |
| *American Journal on Addiction* | 1.4 |
| *Drug and Alcohol Review* | 1.4 |
| *Substance Abuse* | 1.4 |
| *Substance Use & Misuse* | 1.4 |
| *American Journal on Drug and Alcohol Abuse* | 1.1 |
| *Journal of Psychoactive Drugs* | 0.83 |
| *Journal of Drug Issues* | 0.76 |
| *Addiction Research and Theory* | 0.66 |
| *Drug Education and Prevention Policy* | 0.52 |
| *Alcohol, Research, and Health World* | 0.46 |
| *Journal of Drug Education* | 0.42 |
| *Journal of Child & Adolescent Substance Abuse* | 0.39 |

whereas *Journal of Child & Adolescent Substance Abuse* and *Journal of Drug Education* (0.39 and 0.42) have comparatively low impact factors. These figures indicate that articles published in *Addiction* have nearly 11 times the impact of articles published in *Journal of Child & Adolescent Substance Abuse*. In general, Journal Impact Factor scores within a given discipline or profession correlate highly with subjective notions regarding the prestige of various journals, but citation counts also reflect other factors specific to given areas of research (e.g., density of researchers active in a given area; Ha, Tan, & Soo, 2006). Thus, most observers caution against comparing journal citation reports across widely disparate fields (West & McIlwaine, 2004).

Ideally, to maximize dissemination of their scientific findings, field researchers hope to publish in widely circulated journals publishing frequently cited articles. That said, researchers must also be realists and consider two additional issues of key import: journal article acceptance rates and the fit between the article they have written or plan to write and the journals they are considering for manuscript submission.

Journals differ substantially in the proportion of articles they accept for publication. In the substance abuse area, acceptance rates ranged from 25% to 95% across the 35 journals examined by Arciniega and Miller (1997). Many of the most prestigious journals accept only 5% to 10% of submitted manuscripts. Editors and publishers use different methods to compute acceptance rates and have vested interests in portraying their journal in as favorable a light as possible; thus, published acceptance rates must be regarded with caution and are often not readily available.

Field researchers also must ensure that the manuscript they plan to publish is suitable for a given journal. The first step in making this determination is to evaluate the substantive appropriateness of the article for the journal. Journals typically define their scope of publication interests explicitly, both with respect to the types and substantive focus of articles they are interested in publishing, in the print and on-line versions of their journal. For example, the journal *Drug and Alcohol Dependence* was described by its editors and publishers in the following manner (cf., www.elsevier.com):

> Drug and Alcohol Dependence is an international journal devoted to publishing original research, scholarly reviews, commentaries, and policy analyses in the area of drug, alcohol and tobacco use and dependence. Articles range from studies of the chemistry of substances of abuse, their actions at molecular and cellular sites, in vitro and in vivo investigations of their biochemical, pharmacological and behavioural actions, laboratory-based and clinical research in humans, substance abuse treatment and prevention research, and studies employing methods from epidemiology, sociology, and economics. The rationale for this extensive coverage is the conviction that drug, alcohol and tobacco use/dependence cannot be understood in their entirety from a single perspective and that without an understanding of other areas of research, studies by individual investigators may be limited.

> The goal of the journal is to provide researchers, clinicians, and policy makers access to material from all perspectives in a single journal in a format that is understandable and which has received rigorous editorial review. The hope of its editors is to promote mutual understanding of the many facets of drug abuse to the benefit of all investigators involved in drug and alcohol research, and to facilitate the transfer of scientific findings to successful treatment and prevention practices.

A host of books, admittedly published some time ago, such as *Journals in Psychology* (American Psychological Association, 1997), *An Author's Guide to Social Work Journals* (National Association of Social Workers, 1997), *Author's Guide to Journals in Sociology* (Sussman, 1985), and others of their ilk are also available to authors interested in learning more about the publication interests and practices of large numbers of journals in given professional or disciplinary areas. Occasionally, journal articles will also examine publication practices and journal outlets in given specialty areas (e.g., Arciniega & Miller, 1997).

Many journals publish only literature reviews, such as *Drug and Alcohol Review,* while others define the scope of their publication interests largely in substantive (e.g., *Journal of Child and Adolescent Substance Abuse)* or disciplinary (*Experimental and Clinical Psychopharmacology)* terms. It is critical to avoid the delays and frustrations associated with inappropriate manuscript submissions. Authors should take steps to ensure that the substantive focus, research design, and type of article they have prepared or plan to submit is suitable for the journal they are considering for manuscript submission. A good practice for all field investigators is to peruse the most recent 2 to 3 years of issues of all journal publication outlets in their areas of professional interest. This process can reveal many subtleties in the publication preferences of professional journals. It should go without saying that once a journal has been selected for manuscript submission, the manuscript should be prepared in strict accordance with the "Note/Guide to Authors" information that publishers usually provide in print and on-line version of their journals. Such guidelines can range from the relatively brief to the highly detailed. One example of the latter is the "Information for Authors" document prepared by the editors of the *Annals of Internal Medicine* (www.annals.org/shared/author_info.html), which is 35 pages long and very detailed.

Finally, there are occasions when authors are most interested in reaching a particular professional audience. For example, a subscription to the journal *Drug and Alcohol Dependence* is automatically provided to all members of the venerable College on Problems of Drug Dependence. Likewise, all members of the Research Society on Alcoholism receive a subscription to the journal, *Alcoholism: Clinical and Experimental Research*. Researchers interested in having their research seen and read by members of these two important professional bodies in the substance abuse field may choose to submit their articles to these journals, even if other journals with larger circulations and higher Journal Impact Factor scores might otherwise be considered as publication outlets.

In summary, field researchers intending to publish reports from their studies, should familiarize themselves with leading journals in their research area. This process includes investigating journal circulation sizes, Journal Impact Factor scores, publication schedules, number of articles published per annum, and acceptance rates. Researchers should also review all published statements by journal editors regarding the scope, types, and preferred content of articles published and specific recommendations for the preparation and submission of journal articles. *Ulrich's International Periodicals Directory* available in print and online formats (www.ulrichsweb.com) is a good source for much of this information and also includes information pertaining to the number and types of indexing and abstracting services that cover each journal. For example, *American Psychologist* is covered by 21 such services including *Child Development Abstracts*, *Current Contents*, *Index Medicus*, *Social Science Citation Index*, and *Social Work Research & Abstracts*.

*Authorship Issues.* Across the social and natural sciences, there has been an increasing trend toward multiple authorship of scientific papers over the past 50 years (Gibelman & Gelman, 2000). A study the lead author conducted examining authorship of articles in the *Journal of Studies on Alcohol* over half a century provides a compelling case in point. We found that the modal number of authors per article had increased from one to three or more (Howard & Walker, 1996).

Factors other than the trend toward multiple authorship also require field researchers to address authorship issues early in the project planning process. Field research projects have grown increasingly large, complex, and costly in recent years. Multisite investigations staffed by large teams of investigators are increasingly common. Datasets derived from such field research projects are rightly regarded as precious commodities. Project MATCH (Matching Alcoholism Treatments to Client Heterogeneity) exemplifies one such project. Project MATCH was a multisite randomized controlled trial designed to examine the extent to which alcohol dependent patients with particular characteristics responded to various forms of treatment. Three staff members of the federal agency funding the project (i.e., the National Institute on Alcohol Abuse and Alcoholism) were involved in the study design and conduct, as well as a statistical consultant, members of a Research Coordinating Center at the University of Connecticut (including a principal investigator, co-principal investigator, project coordinator, data manager, administrative program manager, and support personnel), members of a Treatment Coordinating Center at Yale (including principal investigator, co-principal investigator, data manager, and research associate), and principal investigators and co-principal investigators for each of the nine clinical research units participating in the project across the United States. To date, more than 50 authors have published a total of more than 120 Project MATCH publications (http://www.commed.uchc.edu/match/pubs/journals.htm). Early in the implementation of Project MATCH, project investigators developed a protocol for assigning and otherwise rationalizing the publication process to ensure equity in the allocation of publication opportunities and to delimit who, and under what conditions, would be provided access to the dataset (Fuller et al., 1994).

Although many projects are not as large as Project MATCH, all field researchers should proactively develop a protocol for sharing data, allocating publication opportunities, and assigning authorship that are consistent with prevailing ethical standards and acceptable to project team members. If the project dataset is eventually made available to the public, as was the case in Project MATCH (cf., www.commed.uchc.edu/match/dataset), policies and procedures must be developed regarding use of the dataset and a process established whereby interested parties can become registered users. In Project MATCH, dataset users were required to sign a formal statement agreeing to protect the integrity of the dataset and study participants' confidentiality.

Other key issues pertaining to authorship include pertinent ethical standards, order of authorship considerations, variations in co-authorship agreements, and acknowledging support.

Many professional organizations have published guidelines for the ethical conduct of various publication-related activities. For example, the

American Psychological Association has published, “Ethical Standards for the Reporting and Publishing of Scientific Information” (American Psychological Association, 2001). In Section 6.23, Publication Credits, the standards state that

> (a) Psychologists take responsibility and credit, including authorship credit, only for work they have actually performed or to which they have contributed, (b) Principal authorship and other publication credits accurately reflect the relative scientific or professional contributions of the individuals involved, regardless of their relative status. Mere possession of an institutional position, such as Department Chair, does not justify authorship credit. Minor contributions to the research or to the writing for publications are appropriately acknowledged, such as in footnotes or in an introductory statement, (c) A student is usually listed as principal author on any multiple-authored article that is substantially based on the student’s dissertation or thesis. (pp. 395–396)

The *Publication Manual of the American Psychological Association* (American Psychological Association, 2001) also includes detailed ethical guidelines for publishing that address the proper reporting of results, plagiarism, duplicate publication of data, and the requirement that data be available for verification purposes. Similar guidelines have been published by a wide variety of organizations (e.g., Graf et al., 2007).

In addition to general guidelines for ethical publication practices, many scientific publishing houses and journals have prepared standards for practice in this area. For example, Elsevier, a publisher of more than 900 scientific journals, has offered useful “Ethical Guidelines for Journal Publication” (www.elsevier.com/wps/find/intro.cws_home/ethical_guidelines) that describe standards for the appropriate conduct of publication-related activities that are specific to editors, publishers, journal manuscript reviewers, and authors. With regard to authors, the guidelines address reporting standards, data access and retention requirements, originality and plagiarism, multiple, redundant, or concurrent publication, acknowledgment of sources, authorship, appropriate treatment of human subjects, disclosure of real or perceived conflicts of interest, and obligations pertaining to correction of important errors in published works.

Field researchers should become familiar with the general guidelines for ethical publication practices in their respective professions/disciplines and with those adumbrated by the journals in which they seek to publish. Many of these guidelines, unsurprisingly, share significant commonalities.

With regard to authorship, it should be noted that order of authorship can have somewhat different meaning across different research areas. In the biological sciences, laboratory directors are often listed last in the author list, whereas in the social sciences authors are typically listed in an order that reflects the extent of their contribution to the submitted manuscript. Authorship of important manuscripts is prized by junior investigators who are often struggling to make tenure. In some cases, it is considerate and appropriate for senior investigators to make prime publication opportunities available to junior investigators, but it is imperative, for ethical reasons, that lead authorship reflects a correspondingly significant body of work on the submitted manuscript.

Coauthorship styles and arrangements are manifold. In some cases, each coauthor contributes one or more component pieces of the submitted manuscript. Manuscripts prepared in this fashion often require a comprehensive final edit by the lead author of the paper; otherwise, they frequently lack transitional sentences, can reflect startling contrasts in writing styles, and may contain numerous redundancies and important omissions. Other arrangements may call for one or more authors to conduct analyses for the report, while other members of the research team write-up the results and prepare the actual manuscript. Scores of such arrangements are conceivable; thus, it is important that such coauthorship agreements are consensually established in advance of manuscript preparation and submission.

For ethical and practical purposes, it is important to acknowledge all persons and organizations that have contributed to the project, especially if their contribution is not reflected in authorship of the submitted manuscript. It is especially prudent to acknowledge funding organizations and staff and other individuals who have supported the project.

## Data Sharing and Archiving

In 2003, the National Institutes of Health (NIH) adopted a formal data sharing policy (http://grants.nih.gov/grants/policy/data_sharing/data_sharing guidance.htm), because it reinforces open scientific inquiry, encourages diversity of analysis and opinion, promotes new research, makes possible the testing of new or alternative methods and hypotheses and methods of analysis, supports studies on data collection methods and measurements, facilitates the education of new researchers, enables the investigation of topics not envisioned by the initial investigators, and permits the creation of new datasets when data from multiple sources are combined. (p. 1)

Although NIH specifically mandated that projects requesting $500,000 or more of directs costs in any year submit a data sharing plan, the new policy stated that "all data should be considered for data sharing... [and] made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidentiality and proprietary data" (NIH, 2003, p. 1).

The NIH data sharing policy notes that project "final research data" can include raw data and derived variables included in electronic datasets and described in associated documents. The policy further underscored NIH's formal requirement that all funded projects must retain all data for a minimum of three years following termination of grant or contract-funded research and that data sharing must be timely (i.e., occur no later than coincidentally with the time at which the main study findings are accepted for publication).

Prior to data sharing, project datasets must be deidentified. That is, all personal identifiers such as names, addresses, telephone numbers, social security numbers, must be removed from the dataset, as well as indirect identifiers that can compromise project participants' identities via "deductive disclosure." Investigators interested in using project data, typically are required to sign formal data sharing agreements designed to protect the confidentiality of project participants. Secondary users of project data are in no way obligated to include original project investigators as co-authors, though they may choose to do so.

In addition to publishing articles in scientific journals, project investigators may directly provide data (e.g., via web access or CD-ROM) to other interested researchers; create data enclaves (i.e., secure, regulated environments wherein registered users can access project data and perform statistical analyses); or establish data archives (i.e., repositories where machine readable data are acquired, fully documented, and distributed; cf. NIH, 2003). Mixed modes of data sharing are not uncommon. Examples of data sharing agreements are available at the following sites: Agency for Health Care Research and Quality National Inpatient Sample (http://ahcpr.gov/data/hcup/datause.htm) and Center for Medicare and Medicaid Services Data (http://hrsonline.isr.umich.edu/rda/userdocs/

cmsdua.pdf; cf. NIH, 2003). A linked data sharing workbook is also available at the NIH Data Sharing Policy and Implementation Guideline website listed for field researchers interested in learning the methods by which other investigators have shared data.

All field investigators should ensure that they have adequately documented all project data collection activities for the benefit of other researchers. At a minimum, this will include developing a project code book defining study variables, data pertaining to variable frequency counts and distributions, and documenting the methods by which data were gathered and organized. Applicants for NIH grants can formally request funds for data sharing activities in their budgets. Field investigators, in general, would be well advised to consider data sharing arrangements in the early stages of project planning, including the timing, format, and mode of data sharing and the types of data sharing agreements that will be required.

The NIH data sharing policy provides significant direction to field researchers conducting research projects of all sizes and funded under the auspices of different funding bodies. Like project publication and dissemination activities, data sharing procedures are best developed early in the project planning process and not on an ad hoc basis.

Investigators, particularly those conducting large field projects, are increasingly interested in formally archiving their project datasets. Data archives "can be particularly attractive for investigators concerned about a large volume of requests, vetting frivolous or inappropriate requests, or providing technical assistance for users seeking help with analyses" (NIH, 2003, p. 4). One of the largest archives of quantitative social science datasets in the world is the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan, created in 1962, with over 550 participating member institutions. At present, ICPSR's data archive includes 450,000 individual files, across a broad range of substantive areas. Field researchers interested in exploring data archiving procedures should review the ICPSR website (www.icpsr.umich.edu) for information regarding data depositing and data use procedures and member benefits and policies.

## Conference Presentations

Academic conferences provide regular opportunities for presentation of project findings. Conferences may be general in nature, sponsored by national and international learned societies, and covering a diversity of topics. Smaller conferences are often devoted to a particular issue or associated with a particular “conference theme.”

Conferences typically offer a variety of presentation formats, including orally presented papers, poster sessions (wherein study methods and findings are concisely presented on poster board during designated “poster sessions”), informal roundtables, and symposia incorporating a number of related oral paper presentations. Researchers interested in presenting at a conference normally submit abstracts of their presentation, which are then reviewed vis-à-vis their scientific merit and interest value to conference attendees.

Field researchers interested in presenting at conferences should carefully read print and online conference brochures to learn where and when conferences will be held, which presentation formats are available to them at different conferences, and the appropriate methods by which they should submit their materials for review.

Newly independent researchers are often interested in learning how to identify appropriate conferences for presentation of project findings. Colleagues engaged in similar research often have useful suggestions as to well-attended conferences that are frequented by influential scientific and professional decision makers. Annual conferences are also often sponsored by professional organizations active in one’s field. As discussed earlier, the College on Problems of Drug Dependence and Research Society on Alcoholism, both sponsor large annual conferences that cover a full range of issues pertinent to drug and alcohol abuse, respectively. Conferences also frequently make travel award monies available to doctoral students, postdoctoral students, junior faculty, and community- or agency-based researchers. Many schools and professional organizations provide extensive directories of professional associations. For example, the Thayer School of Engineering at Dartmouth presents a compelling discussion of reasons why students should seek involvement with professional associations and lists more than 60 professional associations active in the engineering area (http://engineering.dartmouth.edu/career/students/associations.html).

Several online services are also available for identifying professional conferences and linking to their web sites. These services may be subscriber or nonsubscriber based, charge a fee or not, use widely different categories to index conferences, and provide ongoing e-mail alerts regarding upcoming conferences tailored to users’ keyword descriptions of their particular research interests. These services include *Conference Alerts* (www.conferencealerts.com), *Papers Invited* (www.papersinvited.com), *All Conferences.com* (www.all conferences.com/), and *Confabb.com* (www.comfabb.com/). A quick search of upcoming alcohol and drug conferences across these four conference web services identified four such conferences (of which the authors had not previously been aware). Search engines differ across these sites; thus, it is advisable for researchers to search for conferences using multiple sites. Once researchers have identified key conferences they want to attend, they can then link to the conference web sites themselves to learn more about the requirements for registration and presentation of findings.

Conference presentations are important means by which project findings can be disseminated. Conferences offer the opportunity to present findings often long before they enter the published literature. In addition, each conference serves a somewhat different audience. Targeted presentation of findings can ensure that research results are presented to key project constituencies, including fellow scientists, practitioners, and policy makers. Our recommendation would be to seek out colleagues and ask them which conferences they attend, join important professional associations in your research area and present findings at their conferences, and use available online services to identify additional presentation opportunities. In addition, upcoming professional conferences are often advertised in the pages of professional journals; thus, it is important for researchers to review print copies of research journals on an intermittent basis, paying particular attention to "Call for Papers" announcements. Regular presentation of project findings at professional conferences can ensure that your research will have the impact you desire. On occasion, researchers will be invited by organizations in the community, governmental bodies, and other groups to present the findings of their research. In all cases, they should take advantage of these and other unforeseen opportunities to further disseminate the findings of their research.

## Conclusion

Field research dissemination plans often take a back seat to more exigent project concerns. That dissemination plans are so often neglected is unfortunate because a project is apt to contribute little to the scientific literature or professional practice in their absence. Publication and presentation of project findings can be the most rewarding of project activities, especially when results obtained point the way to increased understanding of key phenomena or to improved methods for professional practice. Proactive planning of project research dissemination activities will return the investment of time and energy many times over and ensure the eventual success of the project.

---

## POINTS TO REMEMBER

- Field research findings are commonly communicated as journal articles, monographs, books, presentations at professional conferences and other meetings
- Publication of scientific articles, monographs, and books
  - Prioritize those that derive directly from the aims, then select subsequent papers.
    - Target key papers to top-tier journals, less appealing to mid- or lower-tier journals.
    - Target three to five possible journals for a given submission, from most prestigious down.
    - Publish core papers in professional journals rather than monographs or books.
      a. Can be published sequentially as journal articles.
      b. Once published as journal articles, collect and publish in book or monograph.
  - Levels of prestige.
    - Data-based original reports most highly regarded.
    - Randomized controlled trials most favorably regarded.
    - Project papers in book or chapter format held in less esteem than in journal articles.
  - Selecting journals for manuscript submission.
    - Size of readerships.
      - Prominent professional and disciplinary journals: large circulations and considerable influence.
      - Specialty journals: far smaller circulations and limited influence.
    - Publication practices.
      - How often the journal is published.
      - The average number of articles published per issue.
    - Scientific influence.
      - Articles in professional or disciplinary journals most heavily cited.
    - Prestige.

      - Journal Impact Factor scores.
      - Journal citation reports.
      - Maximize dissemination.
      - Additional issues.
        - Journal article acceptance rates.
        - Fit between the article and the journals.
        - To be read by members of professional bodies submit articles to the free journals sent with membership.
- Authorship issues
  - Address authorship issues early in the project planning process.
  - Develop a protocol for sharing data, allocating publications, and assigning authorship.
  - Develop policies and procedures regarding use of the dataset and how others can use.
- Conference presentations
  - Types of presentations.
    - Orally presented papers.
    - Poster sessions.
    - Informal roundtables.
    - Symposia.
- Choosing conferences.
    - Often sponsored by professional organizations active in one's field.
    - Frequently make travel award monies available.
    - Schools and professional organizations provide directories of professional associations.
    - Online services for identifying professional conferences and linking to their websites.
- Conference advantages.
      - Offer the opportunity to present findings before published.
      - Each conference services a somewhat different audience.
      - Regular presentation at professional conferences can ensure impact.
- Data Sharing and Archiving
    - Prior to data sharing, project datasets must be deidentified.
    - Adequately document all project data collection activities for the benefit of other researchers.
      - Project code book defining study variables.
      - Variable frequencies.
      - Document the methods by which data were gathered and organized.

## References

Advertising and Mailing List Rentals 2006–2007 (2007). *Journals: American Sociological Review*, 3.

American Psychological Association (1997). *Journals in psychology: A resource listing for authors, fifth edition*. Washington, D.C.: American Psychological Association

American Psychological Association (2001). *Publication manual of the American Psychological Association, fifth edition*. Washington, D.C.: American Psychological Association.

American Psychological Association (2006). Updated summary report of journal operations, 2005. *American Psychologist, 61(9)*, 1019–1020.

Arciniega, L. T., & Miller, W. R. (1997). Where to publish? Some considerations among English-language addiction journals. *Addiction, 92*(12), 1639–1648.

Blazer, D., Hughes, D., & George, L. K. (1987). Stressful life events and the onset of a generalized anxiety syndrome. *American Journal of Psychiatry, 144*(9), 1178–1183.

Bourke, P., & Butler, L. (1996). Publication types, citation rates, and evaluations. *Scientometrics, 37*(3), 473–494.

Cottler, L. B., Zipp, J. F., Robins, L. N., & Spitznagel, E. L. (1987). Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *American Journal of Epidemiology, 125*(2), 329–339.

Fontanarosa, P. B., & DeAngelis, C. D. (2007). Thank you to JAMA peer reviewers. *Journal of the American Medical Association, 297*(8), 875–876.

Fuller, R. K., Mattson, M. E., Allen, J. P., Randall, C. L., Anton, R. F., & Babor, T. F. (1994). Multisite clinical trials in alcoholism treatment research: Organizational, methodological, and management issues. *Journal of Studies on Alcohol Supplement, 12*, 30–37.

Graf, C., Wager, E., Bowman, A., Fiack, S., Scott-Lichter, D., & Robinson, A. (2007). Best practices guidelines on publication ethics: A publisher's perspective. *International Journal of Clinical Practice, 61*(152), 1–26.

Ha, T. C., Tan, S. B., & Soo, K. C. (2006). The journal impact factor: Too much of an impact? *Annals of the Academy of Medicine Singapore, 35*, 911–916.

Hopewell, S., Clarke, M., Stewart, L., & Tierney, J. (2007). Time to publication for results of clinical trials. *Cochrane Database of Systematic Reviews, Issue 2. Art. No.: MT000011. DOI: 10.1002/14651858.MR000011.pub2*.

Howard, M. O., & Howard, D. A. (1992). Citation analysis of 541 articles published in drug and alcohol journals: 1984–1988. *Journal of Studies on Alcohol, 53*(5), 427–434.

Howard, M. O., & Walker, R. D. (1996). Multiple authorship: Trends over 50 years in the *Journal of Studies on Alcohol*. *Journal of Studies on Alcohol, 57*(1), 105–106.

National Association of Social Workers (1997). *An author's guide to social work journals*. Washington, D.C.: National Association of Social Workers Press.

National Institutes of Health (2003). *NIH Data Sharing Policy and Implem*entation *Guidance, March 2003*. See pages 1–8 of http://www.grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

Robins, L. N., & Regier, D. A. (1990). *Psychiatric disorders in America*. New York: Free Press.

Shanley, C., Lodge, M., & Mattick, R. (1996). Dissemination of research findings to alcohol and other drug practitioners. *Drug and Alcohol Review, 15*(1), 89–94.

Sussman, M. B. (1985). *Author's guide to journals in sociology*. St. Louis: Haworth.

Tarrier, N., Barrowclough, C., Haddock, G., & McGovern, J. (1999). The dissemination of innovative cognitive-behavioural psychosocial treatments for schizophrenia. *Journal of Mental Health, 8*(6), 569–582.

University of California Press (2006). Publisher of *American Anthropologist*.

Vonnegut, Kurt. (1973) *Breakfast of* Champions, New York: Dial Press.

West, R., & McIlwaine, A. (2004). What do citation counts count for in the field of addiction? An empirical evaluation of citation counts and their link with peer ratings of quality. *Addiction, 97*(3), 501–504.

Wolfe, A. (1990). Books vs. articles: Two ways of publishing in sociology. *Sociological Forum, 5*(3), 477–489.

# INDEX

Aboriginal, 158, 169–171
Abuse:
  CAPTA (Child Abuse Prevention and Treatment Act), 107
  child, 106–108, 110–111, 113–114
  mandated reporting, 107
Accountability, 47, 161
ADD (American Dream Demonstration), 212–215, 222–223, 230–231
AIM-HI (American Indian Multisector Help Inquiry), 108–109, 114, 116
Applicants, 39–40, 42–44, 46
Application, 213, 220, 222
  research application, 71, 130–136, 138–142, 165
Applied research, 208, 211, 215, 222. *See also* Social research
Archiving, 59–60, 77–79, 237, 248–250. *See also* Data
  documentation, 100–101
ASPIRE act, 216, 220
Assets, 206, 208–210, 214–215, 219, 222
  accumulation, 207–208, 213, 216, 219–220
  building, 205–212, 216–217, 219–220, 226–227
  policy, 207–208, 210, 216, 219–221, 223–225. *See also* Public policy
Attrition, 190. *See also* Tracking
Audit trails, 78
Authorship, 121, 194, 245–248

Belmont Report, 105. *See also* Human subjects
Bias, 182, 190–191
Boundaries, 7–8, 44–45, 163, 204
Budget, 10–11, 65–67, 77–78, 137, 164, 250

CASA (Child and Adolescent Services Assessment), 28
CBPR (Community-based Participatory Research), 147, 150–151
CDIS (Computerized Diagnostic Interview Schedule), 61–62
Certificate of Confidentiality, 165
CSD (Center for Social Development), 209–213

CHIPTS (Center for HIV Identification, Prevention, Treatment and Services), 50, 52, 54
CID (Community Intervention Development Model), 13
Circulation, 239–241. *See also* Publication
Citations, 130, 240–242
Clients, 14, 178
Clinics, 8–9, 12–13, 15–16
  clinicians, 6, 8–9, 13–15, 177, 244
Collaboration, 6–7, 11, 13, 46, 107. *See also* Community; Investigators; Team members
Community:
  advisory boards, 6–7, 137–138, 163
  capacity, 151, 172
  collaborations, 146–147, 149–157, 161–169, 172–173
  leaders, 6–7
  members, 109, 151, 154, 159, 162–167
  partners, 9–11, 150–152, 154–161, 167
  partnerships, 110–112, 147, 211, 223
Comprehensive Community Mental Health Service Program for Children and their Families, 180
Compromise, 7, 10, 70, 111–112
Conditions, controlled, 13–14
Confidentiality, 39, 44, 50, 105–108, 113–114, 121, 165, 249. *See also* Human subjects
Conflicts, 3, 8, 104–105, 107–109, 112, 138–139
Consultants, 84, 127–128, 131, 134–142, 144–145
Consultation, 137, 212, 224
Context, 15–16, 24, 31, 36
  practice, 4–9, 13–16
Contracts, 180, 182, 192
Control, 12, 168–169, 188, 190–191, 194
Costs:
  staffing, 41, 59–61
  study, 24, 65–67, 84, 190–191
Culture, 61, 105–106, 111–112, 168
  issues, 104–105, 107–109, 119

D.A.R.E. (Drug Abuse Resistance Education), 183
Data:
  archiving, 77–79
  cleaning, 60, 101
  codebooks, 72, 89–90, 98, 250
  codes, 77–78, 87, 89–90, 93–95
  collection, 60, 69, 76–77, 80, 188
  management, 88
  managers, 63, 77, 246
  missing, 60, 63–64, 90–91, 93–97, 101
  naming, 71–72, 77, 83, 87–88, 98–99
  out-of-range values, 85, 92–93, 96
  preanalysis checklist, 91–92, 98
  raw, 98–99, 103, 249
  sharing, 237, 249–251
  spreadsheet, 97–100
  standards, 82–85, 87, 99, 101
  waves, 76–77, 87, 90, 97
Data entry, 65–66, 69, 72–75, 82–86, 91–93, 95–96
  corrections, 86, 92–93, 95, 97
  errors, 72–73, 75, 86
  personnel, 69–70, 73–74, 85–86
  program, 83–85, 96, 102
  screens, 63–66, 73–74, 76–77
Database, 77, 84–86, 88–89, 92–93, 97–98
  ACCESS, 48, 86
Datasets, 76–77, 83–84, 86–87, 89–93, 97–101

- longitudinal, 76, 83, 101
- project, 246, 249–250

Decolonizing, 153–154, 157–158
Design, 23, 141–142, 187, 191–193, 208
- pipeline study, 188

Development accounts:
- children's, 216–218, 221–222
- individual, 205, 207–214, 218, 221–222

Diagnostic and Statistical Manual of Mental Disorders, 62
DIS (Diagnostic Interview Schedule), 61–62, 64
Dissemination:
- data, 250
- products, 13, 16, 30–32
- results, 156, 234–235, 243

Document, 89–90, 92–93, 99–100, 126
Drug, 135, 175, 197

ECA (Epidemiologic Catchment Area), 238
Editing, 66, 81
Effect size, 186–187
Effectiveness, 6, 27, 177–178, 180–181, 197, 199. *See also* Outcome
Employees, 39, 41–46, 49–54
Errors, 63, 72–75, 85, 96–97
Ethical, 14, 104, 121, 234, 247–248
Ethics, 105, 107, 132, 134
- issues, 106, 108–109, 112–113, 115

Evaluation, 175–181, 183–192, 194–199
Exploitation, 149–150, 162–163

Failure, 174, 196. *See also* Null results
Fidelity, 15–16, 178, 195
Field research, 3, 5, 44–47, 52–54, 128–130, 136
Follow up, rates, 68
Formatting, 62, 67–75. *See also* Data; Instruments
- consistency, 25, 60
- IDs, 76–78, 89, 92, 98
- right-side coding, 68–71, 74
- row, 68, 76, 85, 100

Fort Bragg Demonstration Evaluation, 199–201
Funded research, 127, 143, 236
Funding, 10–11, 132–133, 164–165, 190
- dissemination, 30
- mechanisms, 129–130, 132
- problems, 180

Genetic research, 149
Guidelines, 107, 119, 154, 244, 246–248

Harm, 14, 106, 113
HIPPA (Health Insurance Portability and Accountability Act), 50
HIV/AIDS, 39, 50, 110, 159–160
HONOR (Honor Our Nations, Our Relations) project, 147, 159–160, 166–168
Human subjects, 105. *See also* Confidentiality; IRB
- permission, 48, 51, 57, 117

ICBPR (Indigenist Community-Based Partnership), 159
IDAs (Individual Development Accounts), 207–214
Implementation:
- intervention, 13, 175–176, 178–179, 185–186, 189, 195–196
- study, 43, 110

Incentives, 191, 213–214
Income, 41–42, 206–208, 214–215
Indigenous:
  communities, 147–158, 165–166, 168–169
  research, 152, 158–159, 161, 163–164, 167
Informed consent, 49, 55, 106
Instruments. *See also* Formatting
  adapting and modifying, 25, 59, 60–65, 77
  administration, 24
  behavioral, 47, 87–88
  formatting, 60, 68–69, 73, 85
  instrumentation, 24–26
  items, 25, 61–64, 87–88, 93–96, 100–101
  length, 61–62, 65–66, 87–88
  new, 28–30
  scales, 13, 25, 61–64, 70–71, 94
  scores, 62–63, 248
  service use, 28–32
  skip patterns, 66–68, 74, 91–92, 94–95, 112–113
  standardized, 61–62
  triage questions, 114–115, 117–118
Integrity, 106, 184–185, 192, 246
Interdisciplinary, 127–135, 137, 161
Intervention, 13–15, 129–130, 176–179, 183, 186–188
Interviews, 50–53, 65–68, 71–72, 111–114, 116–119
  computerized, 61–62, 75
  interviewee, 40
  interviewers, 38–45, 47–55, 57, 167–168
  safety, 52–53, 105
  laptop, 65–68
  mock, 50–51. *See also* Pretest
Investigators. *See also* Collaboration; Community
  principal, 67, 131–132, 136–139
  team, 130–131, 135–136, 160
IRB (Institutional Review Board), 41, 104–105, 107. *See also* Certificate of Confidentiality; Human subjects
IRG (Initial Review Group), 138
IQ, 179

Journals, 182, 237–245, 247–248. *See also* Authors; Publishing
  publication bias, 180, 182

Logic modeling, 194

Management information systems, 47–48, 57
MATCH (Matching Alcoholism Treatments to Client Heterogeneity), 246
Measurement, 25, 187
Measures, modifying, 25, 59, 61–64
Mediators, 15, 193
Mental health services research, 133, 138
Meta-analysis, 180–182
MIS IDA, 213
Moderators, 15, 193–194
Multisystemic Therapy, 160–161

National Institutes of Health (NIH), 121, 127–134, 138, 156, 248–250
National Institutes of Mental Health (NIMH), 28, 31, 133, 135–136
Native American, 146–147, 154, 162. *See also* American Indian
NCAI (National Congress of American Indians, 149
Null results, 176–187, 189, 191–193, 193, 195–200

Oregon Social Learning Center, 141
Outcomes, 15, 161, 177–178, 180, 184, 195, 198

Panel Study of Income Dynamics (PSID), 215, 218
Parents, 9, 28–30, 108, 117–118
Participatory research, 147, 150–151, 159, 171–172
Personnel, 137, 163, 167
Phases:
  study, 49, 83, 134, 188–189
  translation, 129, 144
Pretest, 28–30, 37. *See also* Interviews; Mock
Prevention, 116, 119, 184
Program evaluation, 181, 184
Project MATCH, 246
Protocol, 13–16, 91, 94–95
Psychometrics, 24, 29, 93
Public policy, 207–209, 212, 214–215, 219, 221. *See also* Assets; Policy
Publishing, 121, 180, 182, 194–195, 240–241, 243, 245–247, 248. *See also* Authors; Journals
  File-drawer effect, 182

Questionnaire, 85–87, 95. *See also* Instruments
Questions, triage, 114–115, 117–118

Reciprocity, 140, 154, 156
Recruitment of subjects, 48, 188–189
Reliability, 24–25, 29, 59
Research partnerships, 151, 162, 217. *See also* Collaboration; Partnerships
Research questions, 112, 156–157, 191, 197, 212, 222
Retention of subjects, 47–48, 191
Retention of interviewers, 54–55
Review process, 138–139
  reviewers, 130, 132, 134–136
Roadmap Initiative, 128–129
RSPI (Referral Sequence & Problem Interview), 28

SACA (Services Assessment for Children and Adolescents), 29–31
Sample:
  size, 186–187
  random, 141, 190, 193–194
Saving plans, 220
SCA-PI (Services for Children and Adolescents), 28
SEED (Saving for Education, Entrepreneurship and Down payment), 217, 222
Service system, 5, 28, 137–138
Services research, 10, 138, 141
Service use, 26
Sex, 110–111, 147, 159–161
Social research, 205–206, 208, 215, 225. *See also* Applied research
Staff, 8–9, 38–39, 42, 44, 49–55, 161. *See also* Interviews
  follow through, 47–48
  hiring, 38–39, 41–42, 45–47, 49, 53–54, 160
  payment, 139, 162
  supervisors, 6, 8, 50–51
  training, 49–50, 156, 160
Stakeholders, 6, 109–110, 112, 187. *See also* Collaboration; Community
Statistics, 16, 192
Substance Abuse and Mental Health Services (SAMSHA), 181, 199
Suicide, 115–118, 126, 239
Survey, 59–63, 68, 74–75, 77–80, 110

Survey development. *See* Instruments; Formatting; Questions

Team members, 132, 162, 167. *See also* Collaboration; Investigators
Theory, 177, 213
  program theory, 178, 194
Tracking, 78–79, 82, 91–92, 97, 99–100. *See also* Retention
Training, 49–50, 156, 160
Translational, 128–129, 143–144
Trials, 14–15. *See also* Interviews, Pretest; Mock

Validity, 24–25, 28–30, 59, 191
Value labels, 88, 93, 97, 99
Variables, 12, 96–97
  constructed, 77–78
  names, 63, 71–72, 74, 76–77, 86–90