

Kant and the Metaphysics of Causality

ERIC WATKINS

CAMBRIDGE

CAMBRIDGE

more information - www.cambridge.org/9780521835671

This page intentionally left blank

Kant and the Metaphysics of Causality

This is a book about Kant's views on causality as understood in their proper historical context. Specifically, Eric Watkins argues that a grasp of Leibnizian and anti-Leibnizian thought in eighteenth-century Germany helps one to see how the Critical Kant argued for causal principles that have both metaphysical and epistemological elements. On this reading, Kant's model of causality consists not of events, but rather of substances endowed with causal powers that are exercised according to their natures and circumstances.

This innovative conception of Kant's view of causality casts a light on Kant's philosophical beliefs in general, such as his account of temporality, his explanation of the reconciliation of freedom and determinism, and his response to the skeptical arguments of Hume.

Eric Watkins is Associate Professor of Philosophy at the University of California, San Diego.

For Teresa, Christa, and Nicholas

Kant and the Metaphysics of Causality

ERIC WATKINS

University of California, San Diego



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521835671

© Eric Watkins 2005

This book is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format

ISBN-13 978-0-511-08217-7 eBook (NetLibrary)

ISBN-10 0-511-08217-7 eBook (NetLibrary)

ISBN-13 978-0-521-83567-1 hardback

ISBN-10 0-521-83567-4 hardback

ISBN-13 978-0-521-54361-3 paperback

ISBN-10 0-521-54361-4 paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this book, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

| | |
|--|----------|
| <i>Acknowledgments</i> | page vii |
| <i>A Note on Citations and Translations</i> | xi |
| Introduction | 1 |
| PART I: CAUSALITY IN CONTEXT | |
| 1 Pre-established Harmony versus Physical Influx | 23 |
| <i>Introduction</i> | 23 |
| <i>Leibniz and the Many Faces of Pre-established Harmony</i> | 24 |
| <i>Christian Wolff and Pre-established Harmony in Eighteenth-Century Germany</i> | 38 |
| <i>Knutzen's Leibnizian Arguments for Physical Influx</i> | 50 |
| <i>Baumgarten and Meier: A New Case for Pre-established Harmony</i> | 73 |
| <i>Crusius and Fundamental Powers</i> | 81 |
| <i>Conclusion</i> | 93 |
| 2 Kant's Pre-Critical Theory of Causality | 101 |
| <i>Introduction</i> | 101 |
| <i>Kant's Concept of Force in the True Estimation and Physical Monadology</i> | 104 |
| <i>The Nova dilucidatio and the Principle of Succession</i> | 112 |
| <i>The Nova dilucidatio and the Principle of Coexistence</i> | 140 |
| <i>Kant's Pre-Critical Reaction to Hume</i> | 160 |
| <i>New Developments in the Inaugural Dissertation</i> | 170 |
| <i>Conclusion</i> | 177 |

PART II: CAUSALITY IN THE CRITICAL PERIOD

| | | |
|---|--|-----|
| 3 | Kant's Second and Third Analogies of Experience | 185 |
| | <i>Introduction</i> | 185 |
| | <i>The Context of the Second and Third Analogies</i> | 185 |
| | <i>The Argument of the Second Analogy</i> | 203 |
| | <i>The Argument of the Third Analogy</i> | 217 |
| | <i>Conclusion</i> | 229 |
| 4 | Kant's Model of Causality | 230 |
| | <i>Introduction</i> | 230 |
| | <i>Events and Event-Based Models of Causality</i> | 232 |
| | <i>Grounds, Causal Powers, and Determinations</i> | 243 |
| | <i>Making Sense of Activity, or the "Causality of the Cause"</i> | 265 |
| | <i>Implications for the Second and Third Analogies and for the "Critical Turn"</i> | 282 |
| | <i>Conclusion</i> | 296 |

PART III: CAUSALITY AND CONSEQUENCES

| | | |
|---|---|-----|
| 5 | The Metaphysics of Freedom | 301 |
| | <i>Introduction</i> | 301 |
| | <i>The Third Antinomy</i> | 304 |
| | <i>Aspects of Transcendental Idealism</i> | 317 |
| | <i>Freedom and Causality</i> | 343 |
| | <i>Conclusion</i> | 360 |
| 6 | Kant's Reply to Hume: Historical and Contemporary Considerations | 362 |
| | <i>Introduction</i> | 362 |
| | <i>The Historical Question of Kant's Reply to Hume</i> | 363 |
| | <i>Contemporary Considerations</i> | 389 |
| | <i>Conclusion</i> | 421 |
| | <i>Conclusion</i> | 423 |
| | <i>Bibliography</i> | 431 |
| | <i>Index</i> | 441 |

Acknowledgments

I gratefully acknowledge the help and support of numerous people and institutions while I was working on this book. In some sense, I have been thinking about the issues discussed in this book ever since I first read Kant's *Kritik der reinen Vernunft* from beginning to end in a series of seminars that Lorenz Krüger taught over several semesters at the Freie Universität Berlin starting in 1984. Krüger's interest in understanding Kant's theoretical philosophy both in its own terms and in light of the relevant scientific context has served as a useful *Leitfaden* guiding my own interpretive approach ever since. Subsequently, I have benefited from interactions with an increasing number of German Kant scholars – with Reinhardt Brandt, Wolfgang Carl, Martin Carrier, Konrad Cramer, Heiner Klemme, Manfred Kuehn, Wolfgang Lefevre, Georg Mohr, Werner Stark, Thomas Sturm, and Dieter Sturma deserving special mention. While the German and American philosophical traditions have, at times, used different methods to pursue divergent paths of inquiry, I have profited tremendously from what is distinctive in the German tradition and have attempted, in a sustained way over the years, to see how combining its unique insights with lines of investigation prominent on this side of the Atlantic can lead to a broader and more balanced understanding of what Kant was trying to say.

Uncountably – or at least unmentionably – many members of the American philosophical community have been important to me in both direct and indirect ways in writing this book as well. While any list of reasonable length would be egregiously incomplete, I nonetheless want to acknowledge helpful discussions over many years and in many different contexts with Henry Allison, Richard Aquila, Fred Beiser, Eckart

Förster, Michael Friedman, Daniel Garber, Hannah Ginsborg, Michelle Grier, Paul Guyer, Patricia Kitcher, Pauline Kleingeld, Rudolf Makkreel, Steve Naragon, and Allen Wood. Of course, past colleagues and former students at Virginia Tech and Yale have had an important influence on me and my work – in particular, Robert Adams, Roger Ariew, Dick Burian, Andrew Chignell, John Christman, Patrick Croskery, Michael Della Rocca, Mark Fisher, Bill FitzPatrick, Desmond Hogan, Sukjae Lee, L. A. Paul, Joseph Pitt, and Peter Thielke. A number of philosophical friends have been both stimulating and enjoyable interlocutors for many years, especially Jan Cover, Pat Kain, Paul Lodge, Trenton Merricks, Mike Murray, Fred Rauscher, Jim Schmidt, Dave Vessey, Thomas Williams, and Falk Wunderlich. My current colleagues at the University of California, San Diego, have influenced my thinking about Kant as well. I have had discussions with every one of them about various ideas contained in the book, and many have been so generous as to read and provide comments on several chapters – Nancy Cartwright on Chapter 4; Craig Callender on sections of Chapter 6; Dana Nelkin on Chapter 5 and other parts of Chapter 6; Sam Rickless, Wayne Martin, Michael Hardimon, and Rick Grush on Chapter 4; and Don Rutherford has been particularly invaluable with detailed feedback on several versions of Chapters 1, 2, and 4. I should especially like to thank Karl Ameriks, Paul Guyer, Robert Hanna, and Tad Schmaltz for reading and providing me with extensive comments on the penultimate draft of the entire book, and Terence Moore for being an excellent editor at every step in the process. I also thank James Messina for preparing the index. My greatest intellectual debt is to Karl Ameriks, whose incredible patience and invariably wise remarks have been especially invaluable to me for more than two decades. Most important, however, is my gratitude toward my family – especially my wife, Teresa, and my children, Christa and Nicholas, all of whom provided constant encouragement, joyous distractions, and abundant love.

While I started writing this book in 1999, much of it was completed during a year-long Fellowship awarded by the National Endowment for the Humanities in 2001. I am extremely grateful for the financial support that enabled me to focus exclusively on my research at that time. I am also grateful to several publishers for their permission to print in this book some of the ideas that appeared previously in the form of articles. “The Development of Physical Influx in Early Eighteenth Century Germany: Gottsched, Knutzen, and Crusius,” *Review of Metaphysics* 49 (1995): 295–339, and “From Pre-established Harmony to Physical Influx: Leibniz’s Reception in Early 18th Century Germany,” *Perspectives on Science* 6 (1998):

136–203, were expanded in some places, shortened in others, and incorporated into the broader framework of Chapter 1; “Forces and Causes in Kant’s Early Pre-Critical Writings,” *Studies in History and Philosophy of Science* 34 (2003): 5–27, was reworked into the first third of Chapter 2, and elements of “Kant’s Theory of Physical Influx,” *Archiv für Geschichte der Philosophie* 77 (1995): 285–324, are causal ancestors of ideas found at the end of that chapter; “Kant’s Third Analogy of Experience,” *Kant-Studien* 88 (1997): 406–441, provided the original impetus for what became one section of Chapter 3; and “Kant’s Model of Causality: Causal Powers, Laws, and Kant’s Reply to Hume,” *Journal of the History of Philosophy* (2004), contains ideas that are expressed primarily in Chapter 4.

A Note on Citations and Translations

As is customary, citations from Kant's *Critique of Pure Reason* are located by the pagination of the first edition of 1781 (A) and/or the second edition of 1787 (B). All other citations from Kant's writings are located by the volume and page number of *Kants Gesammelte Schriften, herausgegeben von der Deutschen Akademie der Wissenschaften*, 29 vols. (Berlin: de Gruyter, 1902–), such that “1:21” refers to page 21 of volume 1 of *Kants Gesammelte Schriften*. I have typically used the English translations of Kant's writings insofar as they are currently available in the Cambridge Edition of the Works of Immanuel Kant. Specifically, I have made use of Immanuel Kant, *Critique of Pure Reason*, ed. and trans. P. Guyer and A. Wood (New York: Cambridge University Press, 1998); Immanuel Kant, *Lectures on Metaphysics*, ed. and trans. K. Ameriks and S. Naragon (New York: Cambridge University Press, 1997); Immanuel Kant, *Practical Philosophy*, ed. and trans. M. Gregor (New York: Cambridge University Press, 1996); Immanuel Kant, *Theoretical Philosophy, 1755–1770*, ed. and trans. D. Walford (New York: Cambridge University Press, 1992); and Immanuel Kant, *Theoretical Philosophy after 1781*, ed. H. Allison and P. Heath, trans. G. Hatfield, M. Friedman, H. Allison, and P. Heath (New York: Cambridge University Press, 2002). In some instances, I have modified these translations. For all other primary texts, if an English translation is not cited, the translation is my own.

Introduction

The goal of this book is to present a comprehensive account of Kant's views on causality in their proper historical context. Since what I take that context to be departs from the standard view in significant ways, in the first part of this introduction I sketch the familiar contours of the standard view and present very general historiographical and historical reasons that show why we might consider rejecting central aspects of such a view. I also suggest that we can discern the main features of a more adequate account by approaching the topic in a fuller, contextualist manner. In the second part of this introduction I summarize each chapter in this book, illustrating how such an approach makes it possible to provide a more satisfying historical and philosophical account of Kant's central views on causality.

Within general histories of modern philosophy, one can find a narrative concerning the specific issue of causality whose main story line is repeated with remarkable consistency, even if it is embellished on each occasion with slightly different details.¹ Told in the most general of terms,

¹ It is true that certain (primarily epistemological) aspects of a more general history of modern philosophy – of which what is described here is merely a part – were discredited long ago, e.g., by Louis Loeb, *From Descartes to Hume: Continental Metaphysics and the Development of Modern Philosophy* (Ithaca: Cornell University Press, 1981). However, despite the fact that a tremendous amount of excellent scholarly work has contributed to our understanding of early modern philosophers and the specific topics that they address, no consensus has emerged about what general narrative ought to take the place of the standard view. In fact, histories of modern philosophy that would be comprehensive in scope have rarely been offered in recent times. Since the narrative given here has not yet been replaced by an alternative view, it still represents the best available account on this issue.

the story is that philosophers in the early modern period, such as René Descartes and John Locke, attempted to articulate a novel metaphysical account of causality that could support the claims of the “new sciences” of mathematical physics and corpuscularianism discovered by Galileo Galilei, Isaac Newton, and Robert Boyle. Descartes and Locke, as the founders of radically opposed views within modern philosophy (“rationalism” and “empiricism,” respectively) disagreed about many substantive issues, such as the existence of innate ideas and the role of sensations, but their accounts of causality nonetheless revealed remarkable similarities (due in part to the fact that they shared a common opponent, namely medieval and early modern Aristotelians). For both accounts involved purely quantitative properties and exact laws of nature that invoke only efficient and, in fact, mechanistic causation in explaining how a cause necessarily brings about its effect, rather than qualitative features that occur merely “for the most part” and according to final or teleological causes, as many Aristotelians held.

Typically, it is then reported that Descartes’s position came under attack from later “rationalists,” such as Nicolas Malebranche, Benedict Spinoza, and Gottfried Wilhelm Leibniz. They objected that Descartes could not explain how causal relations obtain between the mind and the body, since the mind and the body are, on his account, distinct substances with radically different natures, and they thus concluded that his attempt at providing a truly comprehensive account of causality failed. These figures then developed their own positive accounts of causality in a way that would avoid this objection. Malebranche did so by denying that finite substances, such as the mind and the body, could act at all and by asserting that only an infinite substance (i.e., God) could truly be a cause. Spinoza argued that the mind and the body are not really distinct substances, but rather modes of one all-encompassing substance (God), so that causation between the mind and the body is a relation not between substances with different natures, but rather between modes of a single substance.² Finally, Leibniz asserted that a finite substance can act, but

² Spinoza’s position is more complex than this brief description might suggest insofar as he asserts (1) that each attribute of a substance (e.g., thought and extension) must be conceived through itself and (2) that the order of ideas in the mind parallels the order of (bodily) things. As a result of the first claim, he seems to deny that we could understand how the mind and the body, *described as such*, can act on each other, since understanding such an action would require us to conceive of a mode under one attribute as following causally from a mode conceived of under a different attribute. Yet as a result of the second claim, Spinoza seems to be committed to a parallelism between what occurs in the mind and what happens in the body, which one might otherwise try to explain as the result of causal interaction.

only on itself, and that God, prior to creation, programmed all substances with such extraordinary wisdom and care that their states *merely seem* to be the result of their acting on each other causally. Insofar as each of these three alternatives might appear, at first sight, to be at least as counter-intuitive as Descartes's view was problematic, the "rationalist" line of inquiry concerning causality looks to be, at best, a superfluous curiosity and, at worst, a dead end that simply distracts from the main story line. Since the primary role played by these rationalists in the story of causality is that of critics of Descartes's account of mind-body interaction rather than that of figures who contribute something of lasting positive value to the philosophical tradition that we inherit today, no significant harm would be done if their constructive views were given short shrift and not pursued in any further detail.

If the rationalists' positive views on causality are thus philosophically barren, empiricists would need to step up and play a more prominent role, should there be an interesting story to tell about causality in the modern period. George Berkeley, the next modern philosopher typically considered an empiricist, did not make especially important contributions to discussions of causality, since his views were quite close to Malebranche's. However, David Hume, the final empiricist of the modern period, delivered a truly spectacular performance, even if his very first critics mistakenly panned it. For in the course of following the fundamental assumptions of empiricism to their logical conclusion, so the story goes, he developed extraordinarily powerful criticisms of the very foundations of early modern accounts of causation, arguing, among other things, that the basic notion of causality invoked in their accounts does not possess the sense of necessity claimed for it. All that causality could be for us, according to Hume, is the constant conjunction of two events and a "subjective feeling of the mind," or expectation, following repeated observations of their regular occurrence together in the past, that they will be correlated in the future as well. For Hume asserts that one can have no impression from internal or external sensation of a necessary connection between any one event (the cause) and any other (its effect) and thus that no corresponding meaning can be attached to the terms that are commonly used to describe this kind of connection, such as "force," "power," or "bringing about." As a result, Hume famously showed that the new sciences do not require as robust a metaphysical account of causality as Descartes and Locke had thought, since all that is needed are mere regularities between distinct events rather than necessary connections between substances and their states.

It is at this point that Immanuel Kant enters the standard story, claiming to have a reply both to the rationalists' (overly) ambitious claims to knowledge of God, freedom, and the immortality of the soul and to Hume's skeptical doubts about causality. In one of the most famous passages to be found in the history of philosophy – the Second Analogy of Experience – Kant is supposed to refute Hume's position by showing that the notion of causality Hume had called into question is not in fact dispensable at all, but is rather absolutely necessary as a condition of the very possibility of experience. Specifically, the category of causality is necessary because it makes possible knowledge of *objective* succession as something distinct from the merely *subjective* flow of our representations in consciousness.

The story fails to have an entirely happy ending, however, since after more than two centuries of sustained exegetical and philosophical attempts, no consensus has emerged about *what* Kant's argument in the Second Analogy is and *how* it is supposed to refute Hume's position. Instead, one typically faces some version of the following dilemma. Either one cannot find a valid argument that is actually successful in refuting Hume's position (i.e., many different reconstructions are proposed on Kant's behalf, but clever Humeans then take delight in pointing out the fallacies they involve) or one can identify a valid argument, but then, on closer inspection, one must admit that, in some subtle way, it draws on assumptions that a Humean could easily reject, so that the search for a cogent argument that does not beg the question against Hume continues.

Although failing to find an argument in Kant's Second Analogy of Experience that both successfully refutes Hume's position and does so on his own terms causes the story to end on a disappointing note, one's evaluation of how unhappy the ending is may depend, to a certain extent, on one's own philosophical outlook. Obviously enough, contemporary Kantians who are attracted to the idea that there might be substantive conditions of the possibility of experience continue to search for an argument that would win the day but are severely burdened with the worry that the lack of consensus is due not to the obscurity that might naturally be thought to accompany an argument so profound in its insight, but rather to the fact that no such argument is there to be found in the first place. By contrast, when present-day Humeans are in an optimistic mood, they might draw support from the Kantians' oft-repeated failure to produce the sought after goods and view the lack of consensus about his argument as representing the history of philosophy's verdict on the issue

and thus as constituting an important piece of evidence in their overall case for empiricism. However, without a definitive analysis of Kant's exact intentions and argument in the Second Analogy, Humeans cannot rest fully content in their views insofar as they must live in constant fear (or at least with the prospect of sporadic fits of melancholy) that the latest argument offered on Kant's behalf could turn out to be decisive.

The primary aim of this book is to tell a far different and, I hope, much more satisfying narrative about Kant's account of causality and its place in the history of modern philosophy. By taking some general historiographical considerations into account, we can provide both a compelling analysis of why the standard view goes wrong and positive guidance about how a more satisfactory story can be told. Though part of the standard view's appeal surely stems from the tidy way in which it can relate the history of modern philosophy in a simple, linear fashion (with each major figure in a given tradition improving in some way on the views of his immediate predecessor), its primary motivation in the case of causality is that it seems to allow Kant to speak directly to our contemporary philosophical interests because the issue of causality can be used to serve as a paradigm case for addressing the question of whether there are systematic grounds that would suffice to refute empiricism.

However, there are obvious dangers in an approach that simply assumes that Kant's interests are identical to our own so that we can immediately reconstruct and evaluate Kant's argument without having to bother with much else (beyond his unique terminology and odd architectonic, which add more than enough seemingly unnecessary difficulties on their own). It clearly runs the risk of leading one to read "foreign" arguments into an author's texts and it is perhaps not too surprising when the fate of these arguments turns out not to be consonant with the reputation of the philosopher in whose name they are offered. Less obviously, but more importantly, it can also lead to a narrowing of focus. In the current case, since Hume is typically considered the prime representative of empiricism and, as such, develops arguments about causality that are quite attractive to many today, and since Kant is explicitly critical of Hume in the Second Analogy, this approach makes it seem clear that Kant's primary interest in the Second Analogy of Experience lies in refuting Hume's skeptical doubts. Unless it bears directly on the Second Analogy, whatever else Kant says (even about causality) can be discussed later.

We can avoid these dangers if we first try to understand Kant's views and arguments within their proper historical context, before determining

whether and how it makes sense to use them in addressing our contemporary questions and concerns. The primary disadvantage of this historiographical approach is that one cannot guarantee in advance that Kant will have anything interesting to say to us. In response, one might simply appeal to the fact that the value of Kant's position to contemporary philosophy has been consistently documented over such a wide range of issues in the past that there is no reason to think that the issue of causality should be any different. Yet one can also point out that the standard view is, in reality, in no better shape on this issue. For simply assuming without question that what Kant says directly addresses our contemporary concerns does not entail that it actually does so, a fact made clear by the repeated failure of previous reconstructions of Kant's argument in the Second Analogy to refute empiricism.

But what exactly does it mean to say that we should understand Kant's views on causality in their proper historical context? To assess the standard story's assumption that refuting Hume's position in the Second Analogy of Experience is Kant's primary concern regarding causality, three general points are immediately relevant. First, it would be a mistake to assume that one can focus on the Second Analogy to the exclusion of other passages within the *Critique of Pure Reason*. For however one interprets Kant's argument in the Second Analogy, it must be consistent with the main thrust of his other arguments in the *Critique*, especially those that deal directly with the issue of causality, such as that of the Third Analogy and the Third Antinomy – *regardless* of whether or not their arguments appear, at first glance, to be a lost cause from a contemporary philosophical perspective, since, according to the historiographical approach being recommended, our primary task is simply to *understand* what Kant's views are.

Second, it would be preferable if Kant's views on causality in the *Critique* were to fit in naturally with his remarks about causality in other contexts, for example, with those that occur in works written during his "pre-Critical" period, in the *Reflexionen* that give insight into his private thoughts at the time, and in the transcripts from the metaphysics lectures he held throughout his career. It is true that these passages involve various complex interpretive issues (e.g., involving the "Critical turn" and Transcendental Idealism), but addressing such complex issues may ultimately provide indispensable help as opposed to problematic obstacles to our inquiry (as, I think, the narrowness of the standard view ultimately does). Thus whatever the "proper historical context" might amount to exactly, we should prefer an account of causality that is more comprehensive in

scope than is suggested by the standard story's exclusive focus on the Second Analogy.³

Finally, as soon as we extend our focus beyond the Second Analogy of Experience, it becomes even more imperative that we ask whom Kant intended to address with the *Critique of Pure Reason*. That is, for whom did Kant write this book and whose views did he intend to attack with it? It is evident from its title that he wanted to criticize views that he thought of as being supported by "pure reason" alone. Hume, who is a famous opponent of reason, would seem to be Kant's ally in such an endeavor rather than his enemy. Moreover, the fact that Kant wrote the *Critique* in German and not Latin suggests that his intended audience was not primarily European (whether it be French or English), but rather German. Nor ought one underestimate the consequences of the fact that Kant was educated and then lectured and wrote throughout his entire career only in Germany (or East Prussia, to be more exact). It should not be surprising, but rather to be expected that an exclusively German education and career would influence in significant ways both his fundamental aims and the particular ways in which he tries to achieve them. Fairly general historical considerations thus suggest that Kant would be directing his views at Leibniz and his various rationalist "followers," such as Christian Wolff, Moses Mendelssohn, Martin Knutzen, Alexander Baumgarten, and Georg Friedrich Meier, and even the briefest familiarity with Kant's pre-Critical writings reveals that he was extremely interested in the views of Christian August Crusius as well. Accordingly, if these figures influenced Kant, then the rationalists' views on causality may not be the dead end that the standard view maintains.

Thus, to understand Kant's views on causality in their proper historical context we must undertake several specific tasks before beginning to think about how his views might be adapted to address our contemporary interests. First, we must establish what range of substantive views on causality Kant would have been familiar with from his education and the first part of his career. Then, we must consider what his initial reaction to these views was during the roughly two and a half decades of his pre-Critical period (1746–1770). Only then can we look to the *Critique* in order to determine what his intent and arguments are. At that point we must take

³ Following this approach to its logical conclusion, one ought also to consider Kant's views on teleological causation in the *Critique of Judgment* as well his views in physics in the *Metaphysical Foundations of Natural Science*. In light of the scope of our current investigation, however, for the present, we set aside such considerations for the most part.

into account more than just the Second Analogy of Experience; at the very least, we must give careful consideration to Kant's Third Analogy of Experience, which asserts the necessity of mutual interaction, just as the Second Analogy does with respect to causality. However, we must also be open to the possibility that Kant's Third Antinomy (which addresses the consistency of freedom and natural causality) could add significant content to his views on causality. Only after having completed these tasks can we turn to evaluating Kant's arguments and consider whether and how they might be appropriated for other contexts. That is, only at the end of such an investigation, and not at its beginning, are we in a position to determine how Kant is replying to Hume and in what respects Kant's views on causality might be relevant to our contemporary questions and interests.

The structure of the following investigation into Kant's views on causality and their place in the history of modern philosophy is organized according to the contextualist historiographical approach described above. Part I ("Causality in Context") begins by presenting a detailed historical account both of which views of causality Kant would have been familiar with and of what his own immediate reaction to these views was in his pre-Critical writings. Chapter 1 ("Pre-established Harmony versus Physical Influx") focuses on the first of these tasks by presenting the views of Leibniz, Wolff, Knutzen, Baumgarten, Meier, and Crusius on causality. Though most of these figures suffer complete neglect in standard histories of philosophy (and considerably less attention in specialist histories than one might expect), their discussions of causality are often interesting in their own right and revealing about what issues were considered important at the time. In addition to addressing topics that are standard fare in the history of modern philosophy (such as the mind-body problem and the problem of the conservation of motion or living force), their primary focus on the issue of causality – which happened to be, for perhaps independent reasons, the central philosophical topic of the day – took the form of a debate about whether to accept Leibniz's doctrine of "pre-established harmony" or to develop a version of a doctrine he dubbed "physical influx," which allows for causation between finite substances.

Although Leibnizian-Wolffian philosophy is often presented as if it were a monolithic view, it turns out that many of those to whom this term refers developed their views with a considerable degree of independence from Leibniz. Accordingly, while Leibniz originally proposed pre-established harmony in the context of his idealism and his

view of the relationship between the primitive forces of monads and the derivative forces of bodies, Wolff, against the background of the radical shift in intellectual context that occurred in Germany between the late seventeenth century and the first part of eighteenth century, restricted it to the mind-body relationship (as opposed to considering it as a doctrine that pertains to all finite substances). As a result, despite his agreement with Leibniz about the necessity of ultimately real simple substances that result in the composite bodies we see, Wolff was agnostic about whether *all* simple substances must have representational powers or whether some might be endowed with physical forces instead.

Though Knutzen is a Leibnizian, just as Wolff is, namely in virtue of accepting simple substances that are endowed with the power of representation and that bear ultimate responsibility for the physical properties of bodies, he mounted an extensive case in favor of physical influx rather than pre-established harmony. While his case involves several distinct arguments, its main thrust relies on the idea that if a simple substance either has the capacity to move itself or is impenetrable (i.e., can resist the attempt of a distinct substance to penetrate the space it occupies), then it must also be able to act on other substances. Baumgarten and Meier, who were more orthodox Leibnizians in virtue of accepting pre-established harmony and many of Leibniz's other views, articulated new arguments for pre-established harmony, arguments that were based on intricate terminological considerations pertaining to what relations are required for substances to belong to the most perfect world and on some unusual metaphysical assumptions about the notion of an action (including that of a smallest action).

By contrast, Crusius, the leading Pietist philosopher of his generation, rejected many of the principles that were considered fundamental to any Leibnizian system. As part of his anti-Leibnizian program, he developed an interesting independent project that placed causality (in the guise of his notion of a "fundamental power") at its very core. In the course of carrying out this project, he argued that real rather than ideal relations are required to explain why certain substances belong to one and the same world, that substances can be related to each other by means of their very existence, and that God's will rather than his intellect plays an essential role in explaining why substances do not merely correspond to, but also depend on each other. Thus, the historical background to Kant's account of causality as formed by the views of Leibniz, Wolff, Knutzen, Baumgarten, Meier, and Crusius in the first half of the eighteenth century is much more diverse and interesting than one might have surmised from

the way that it is treated or, more typically, passed over in most histories of modern philosophy.⁴

Chapter 2 (“Kant’s Pre-Critical Theory of Causality”) then shows in detail how Kant’s views during his pre-Critical period can be understood against the background of the positions presented in Chapter 1. Though one might not have expected someone like Kant simply to toe the standard Wolffian line, it is striking to see just how creative his contributions to the debate about causality were. Early on Kant cultivated a deep and abiding interest in metaphysical aspects of causality by defining (in 1746/1747) the concept of force in terms of activity (rather than motion, as “certain Wolffians” had) and by developing (in 1755) an intricate argument designed to show that change in the intrinsic determinations of substances is possible only if they stand in causal connections with each other, that is, only if physical influx is true. While Kant was thus highly critical of the Wolffian position in several respects, he also developed a nuanced attack on Crusius’s position. He agreed with Crusius (against Baumgarten and Meier) that only a real relation can allow substances to belong to one and the same world, but then argued against Crusius that substances cannot stand in real relations by means of their existence alone. As a result, in attempting to chart a middle course between the positions of the Wolffians and Crusius, Kant developed a sophisticated metaphysical account of causality, according to which (contra the Wolffians) substances can act on each other by means of the grounds that constitute their immutable essences, and (contra Crusius) grounds must be understood in terms of the activities rather than the mere existence of these substances.

It is crucial to note, however, that Kant’s attention during this period was not limited solely to German philosophers, even if they were clearly his main focus in developing his distinctive account of causality. For after Hume’s *Enquiry Concerning Human Understanding* was translated into German in 1755, Kant reacted by introducing a new metaphysical distinction between real and logical grounds, reinterpreting the ontological principles he had developed earlier in terms of real grounds and making the notion of a real ground fundamental to several principles that became central parts of his overall position in the early 1760s. By the time of his

⁴ Lewis White Beck’s classic *Early German Philosophy: Kant and His Predecessors* (Cambridge: Harvard University Press, 1969) and Max Wundt’s *Die deutsche Schulphilosophie im Zeitalter der Aufklärung* (Hildesheim: Georg Olms, [Tübingen, 1945] 1992) are both excellent counterexamples to this claim, though both accounts, due to their breadth (which is quite extraordinary in Beck’s case), cannot focus on a single specific issue such as causality.

Inaugural Dissertation in 1770, however, Kant had apparently finished incorporating into his account the changes that he thought Hume's objections required in a direct way, even if he had still not completely worked out all of the implications that followed from his immediate reaction to them. Instead, by this time he had begun to reflect on more general issues in metaphysics, such as whether the world has an essential form (in addition to the form it happens to have in virtue of its actual interactions), how to understand the unique principles of the sensible world, and what the possible consequences are of not maintaining a strict separation of the principles of the sensible world from those of the intelligible world. Over the next decade, Kant would continue to reflect on these and other issues that, taken together, amount to what is commonly referred to as the "Critical turn."

This detailed description of what accounts of causality Kant was familiar with early in his career and of how he reacted to them and developed his own thoughts further throughout his pre-Critical period puts us in a position to turn to our main task in Part II ("Causality in the Critical Period"), namely that of understanding Kant's intentions regarding causality in their proper historical context by presenting his account of causality as it was developed in the *Critique of Pure Reason* and other relevant texts from that period in light of the results of Part I. Chapter 3 ("Kant's Second and Third Analogies of Experience") reconstructs Kant's central arguments in the Second and Third Analogies of Experience. On the one hand, these arguments are fully "Critical" in the sense that they are not simply remnants left over from his pre-Critical period and then added on as extraneous elements to his project in the *Critique*, but rather integral components of his project. As Analogies of Experience they play a central part in the *Critique's* systematic goals by establishing the necessity of two particularly important categories, namely causality and mutual interaction. By solving the problem of time-determination that arises for our knowledge of temporal relations, they also reveal how experience of a particularly fundamental kind is possible. Accordingly, Kant sees these arguments as making a major contribution to his primary goal in the *Critique* of establishing the conditions of the possibility of experience.

On the other hand, when the arguments of the Second and Third Analogies are interpreted in light of their proper historical context, several features immediately stand out. First, a number of Kant's crucial premises depend on concepts and principles that derive from his pre-Critical period. The Second Analogy's claim that causality is required to determine the successive states of an object is justified in part by noting,

just as Kant does in his pre-Critical period, that any determination requires a real determining ground, which is just a different name for a cause. Similarly, the Third Analogy's distinctive assumption that a substance cannot determine its own place in time finds its most plausible justification in a line of argument that is based on his pre-Critical principle that a substance cannot act on itself so as to change its own state.

Second, it is striking that Kant does not simply launch on an entirely new Critical line of argument in the Analogies, leaving his pre-Critical project completely behind. Rather, he combines certain aspects of his pre-Critical views (e.g., his interest in the connection between temporality and causality and in the concept of the world as a real whole) with a radically new metaphilosophical and methodological "Critical" framework. More specifically, Kant incorporates his pre-Critical view that mutual interaction, as a real causal relation, is necessary for substances to form a single world into his project of explaining how we can have a single, unified experience, that is, experience of a plurality of objects unified in a single time. Taking Kant's pre-Critical views into account thus allows us to see, in a way that was not obvious before, that, at least in the context of these central arguments of the *Critique*, Kant is neither an arch-epistemologist (who might be concerned solely with "epistemic conditions" or "inference tickets") nor a purely descriptive metaphysician (who would merely try to describe, on the basis of conceptual analysis, what the world must be like). Rather, he is interested in establishing a certain kind of metaphysical principle (concerning causality and mutual interaction at the phenomenal level) as the necessary presupposition of fundamental epistemological principles (which include our knowledge of succession and coexistence, that is, our unified experience of the world).

Chapter 4 ("Kant's Model of Causality") then considers what model of causality Kant is committed to on the basis of the arguments of the Second and Third Analogies, taken in conjunction with several other remarks concerning causality that he makes in the Critical period. What these arguments – and especially that of the Third Analogy of Experience – make clear is that Kant's model of causality cannot be that of one event causing another event, since it would be contradictory to assert that one event could mutually interact with, that is, be both the cause and the effect of, another. Since Kant presupposes a model of causality that is fundamentally different from Hume's event-event model, he cannot be using the arguments that are based on this model to refute Hume's position. Because Kant never explicitly asserts that he intends to refute Hume's position and because the fundamental structure of his argument is incompatible with

such an intention, we can now see quite clearly that the standard view's assumption about the intent of the Second Analogy must be mistaken.

If Kant's model of causality is not that of one event causing another, how should it be understood? Again, Kant's pre-Critical theory of causality provides important guidance insofar as one can understand Kant's Critical model as being similar (though certainly not identical) to his pre-Critical model in basic ways. For not only does Kant continue to accept the notion of a real ground that was fundamental to his earlier account, but he also continues to think that causality occurs if one substance determines the state of another by actively exercising its causal powers according to their natures and circumstances.

Given that causal powers are a standard feature of accounts of causality from Aristotle on, the fact that Kant's account of causality invokes causal powers might lead one to think that he has nothing new to offer. In fact, however, the ways in which he develops his position are quite innovative. First, Kant rejects the identification of causal powers with substances (an identification that Baumgarten explicitly endorsed), since that violates our standard way of talking about substances as *having* powers. Yet Kant also refuses to allow that causal powers might be simply relational determinations (or what we might call relational properties), since the grounds of relational determinations cannot themselves be relational determinations (on pain of infinite regress). Thus, they are irreducible to either substances or relational determinations, and must instead be accepted as a primitive relation "in between" substances and their determinations. Second, Kant also thinks of this irreducible relation as being essentially asymmetrical in virtue of the way that the active-passive distinction applies to it. If a cause determines its effect, it does so by actively determining some object that is passive with respect to that determination. This activity, or, as Kant sometimes puts it, the "causality of the cause," is not something that could itself be determined, since as something essentially active it can never be a passive determination. By incorporating an asymmetrical active-passive dichotomy into an irreducible causal relation in this way, Kant can represent his model of causality as distinct from more traditional accounts of causal powers.

But what sense can be attached to the notion of activity that is central to Kant's distinctive model of causality? Unfortunately, several of the passages that one might naturally look to for an answer, such as the *Metaphysical Deduction* and the *Schematism*, are of no help in clarifying the content of Kant's category of causality and thus the notion of activity it contains (beyond what was already clear from the arguments of the

Second and Third Analogies). Moreover, while Kant's account of physics fits in perfectly with such a notion insofar as it holds that a body *exercises* its attractive and repulsive forces in causing other bodies to move, it cannot add any clarification to that notion. For Kant can agree with Hume's insight, translated into Kantian terminology, that we do not have an intuition of the exercise of such forces (which must instead be represented by the categories), just as we do not literally see the "hitting" of one billiard ball by another. That is, it is difficult to identify in a clear way what there is to the activity of a cause apart from the empirical effects it produces; all that we seem to "see" are determinations (passive determinate states), not determinings (i.e., the processes by which the determinations are determined). Fortunately, however, Kant's account of self-consciousness and his distinction between apperception and inner sense can provide help on this point insofar as they show that even if we do not *know* the determining self through inner sense (since we can know only the determined self in this way), we can still be aware of activities in apperception in order to be able to be aware of representations as our own. As a result, Kant's model of causality not only differs in significant respects from both Hume's event-event and Leibniz's causal power models, but can also explain its fundamental components with a reasonable degree of clarity.

In Part III ("Causality and Consequences"), we turn to consider what consequences follow from Kant's Critical views on causality for closely related issues such as that of freedom and the question of what the nature of Kant's reply to Hume is and how it might be relevant to our contemporary philosophical interests. Chapter 5 ("The Metaphysics of Freedom") discusses the relations between Kant's views on natural causality and freedom, revealing that many of the basic metaphysical concepts employed in the model of phenomenal or natural causality described in Part II are of significant help in appreciating some of the less readily understood aspects of Kant's account of freedom. Specifically, just as natural causality is to be understood primarily not in terms of events, but rather in terms of a substance determining the state of another substance by means of an exercise of causal powers in accordance with its nature, so, too, freedom is to be understood not primarily in terms of desires (which are simply one kind of mental event), but rather in terms of an agent or, more metaphysically, a substance determining its actions according to its freely chosen character.

These conceptual and structural similarities allow one to see more clearly how Kant might hope to answer certain questions that naturally arise regarding the multifaceted problem of free will and determinism.

For example, by understanding causality in terms of the activity or determining ground of a substance rather than a determinate event, one can see that the radical difference in kind between our agency – which, as an activity, determines our states – and any complex hierarchy of desires that we might have – which, since they are determinate states, will always be determined or caused by something else and thus stand in need of further explanation – can help to stop the potentially infinite regress that seems to arise otherwise. They also allow one to understand Kant's resolution of the modal conflict that arises between the necessity of determinism and the contingency (apparently) entailed by free will. For if the laws of nature, from which the necessity of determinism derives, are contingent on the natures of things, including the natures that we freely choose for ourselves (which we typically call our character), then it turns out that the necessity of determinism does not ultimately conflict with, but rather depends on, the contingency of our free will. Accordingly, understanding Kant's general model of causality helps to clarify how Kant might think that he can respond to certain long-standing difficulties that arise in attempting to account for freedom.

However, it would be a mistake to think that clarification goes only in the one direction. For Kant's understanding of freedom helps us to appreciate certain aspects of his model of natural causality that might otherwise go unnoticed. One of Kant's most prominent discussions of freedom occurs in the Antinomies of Pure Reason and thus in the context of Transcendental Idealism, which, unlike Transcendental Realism, distinguishes between appearances and things in themselves, that is, between the phenomenal and noumenal worlds. This distinction is important for the problem of free will and determinism because it allows for the possibility that they could both be true (namely, if they hold for different worlds). Determinism must apply to the phenomenal world, because the cause of any phenomenal event presupposes a prior event in time, which, since it must be caused, requires a prior event, and so on. But we might be able to act freely in the noumenal world, because it is not temporal and noumenal causes are therefore not events that presuppose prior events that require causal determination.

While temporality thus forms a crucial aspect of Kant's resolution of this aspect of the problem of free will and determinism, it points to an even more fundamental difference between the phenomenal and noumenal worlds, namely that the noumenal world is completely determinate, whereas the phenomenal world (with times and causes going back indefinitely) cannot ever be completely determinate and must

therefore be indeterminate in some respect.⁵ The indeterminacy of the phenomenal world is important to the issue of freedom because we can freely determine our actions in the phenomenal world only if it is previously indeterminate with respect to that action. Because freedom is understood in terms of our agency, that is, in terms of our being grounds that determine our actions, the world in which our actions occur must be open to being determined in that way or, in short, must be indeterminate. But this point can be applied to Kant's model of phenomenal causality as well. For admitting indeterminacy in appearances creates the conceptual space that is necessary for Kant's model of phenomenal causality, which holds both that events can become determinate through the causality of a phenomenal cause (or determining ground) and that the activity or "causality of this cause" can itself be indeterminate and thus not an event. In this way, Kant's account of freedom can be used to clarify his general model of causality, just as his general model was used to highlight aspects of his views on freedom.

In Chapter 6 ("Kant's Reply to Hume: Historical and Contemporary Considerations") we can finally address the historical question of what Kant's reply to Hume is and how Kant's views on causality can be relevant to our contemporary interests. With regard to the historical question, it is helpful first to supplement our discussion of Kant's immediate reaction to Hume in the pre-Critical period (in Chapter 2) by considering how Hume's *Enquiry* was received more generally in Germany from 1755 to 1770. What emerges from this consideration of the reception of Hume (by Sulzer and Tetens) is that Kant would have been justified in assuming that most, if not all of his readers (but especially those enamored with "pure reason") would not have thought that a refutation of Hume's views on causality was necessary in the first place. But to understand the Critical Kant's views on Hume it is important to consider his explicit references to Hume in the first and second editions of the *Critique* as well. They suggest that Hume's views on causality were important to him not primarily in their own right, but rather as an illuminating illustration of Hume's more general skeptical approach, which, due to its inherent instability, should be replaced with his own Critical methodology.

Finally, reflection on the differences between Kant's and Hume's models of causality reveals a vast chasm. Hume's events are states of affairs

⁵ Kant's claim that the phenomenal world is essentially indeterminate follows from his analysis of how the condition-conditioned relationship applies in different ways to the phenomenal and noumenal worlds.

at instantaneous moments in time, whereas Kant's events are continuous changes of state over time. Hume accepts only events (or matters of fact), while Kant accepts noumenal and phenomenal substances as well as indeterminate entities such as inherence and "the causality of the cause," or the exercise of causal powers according to the natures and circumstances of the substances involved. Hume attempts to construct an account of the world solely on the basis of such discrete events, whereas Kant is concerned with explaining how we can know the temporally determinate states of objects within a single spatio-temporal world. Hume's events are distinct from each other, while Kant attempts to establish grounding or dependency relationships between substances and their determinate states. Hume's events are neither active nor passive, Kant's determinations are passive and his causes active. In short, Hume's and Kant's ontologies are radically different, and the lack of a shared vocabulary makes it impossible for one either to find a refutation of Hume's position in Kant's explicit arguments or to translate Kant's arguments into Humean terms such that the presuppositions required for a refutation to be possible would be satisfied.

If Kant is not attempting to refute Hume, then how should his project be understood in relation to Hume? From Kant's perspective, once Hume understands events the way he does and undertakes the project of attempting to construct the world out of them, he makes no mistake in inferring that we can be aware of no necessary connections in nature and there is no way to refute this view by pointing out some obvious move available within the empiricist framework that Hume simply overlooked. Rather, Kant recognizes that he must pursue a different project on the basis of a completely different set of ontological and epistemological presuppositions, and his account of causality should thus be understood accordingly as the result of his attempt to articulate a very different alternative to Hume's empiricist account.

If our answer to the historical question concerning Kant's reply to Hume is thus that Kant is attempting not to provide a refutation of Hume, but rather to elaborate an alternative philosophical account of causality, then Kant's views on causality can be relevant to our contemporary interests by suggesting not that we should look to Kant's views for a refutation of empiricism, but rather that we should see whether they can provide materials that help in the articulation of views that would be alternatives to empiricism. Accordingly, in the second half of Chapter 6 we can consider how Kant's model of causality and the way in which it contrasts with Hume's is directly relevant to three issues that have been

widely discussed in contemporary contexts: the metaphysics of causality, the nature of the laws of nature, and agent causation.

With respect to the contemporary debate about the metaphysics of causality, one can draw on Kant's notion of activity to explain what makes a causal power different from and irreducible to an event, an explanation that is indispensable for debate about the nature of causal relations to make sense in the first place.⁶ One can also see that Humean theories of events may not have the advantages sometimes claimed for them, since they presuppose, seemingly arbitrarily, that causal relations cannot themselves be "perfectly natural" events, to use Lewis's phrase.

Second, one can use the resources available in Kant's model of causality to defend a necessitarian conception of laws of nature (as developed by David Armstrong, among others) from two important objections that have been raised against such a conception.⁷ Specifically, because Kant holds that the laws of nature are based on the natures of substances and that substances must act in accordance with those natures, one can clarify both the *kind* of necessity that the laws of nature have (namely, natural rather than causal or logical necessity) and the sense in which a Kantian can say that the laws of nature *govern* what happens in the world without having to assert that the laws of nature are themselves directly causally efficacious.

Third, Kant's views on causality can be used to develop an account of agent causation that can respond to objections commonly raised against it. Against C. D. Broad's objection to the very idea of agent causation one can follow Kant in distinguishing between the activity of the agent and the circumstances under which the agent's activity is operative so that one can admit that there are datable factors relevant to a cause bringing about an effect without the cause itself being datable.⁸ Further, one can use Kant's

⁶ This literature on this topic is vast. For the purposes of this discussion, our attention focuses on the views of David Lewis, "Causation," in *Philosophical Papers* (New York: Oxford, 1986), vol. 2, pp. 159–213, "Causation as Influence," *Journal of Philosophy* 97 (2000): 182–197, and "Events," in *Philosophical Papers*, vol. 2, pp. 247–254; Rom Harré and Edward Madden, *Causal Powers: A Theory of Natural Necessity* (Oxford: Blackwell, 1975); and Nancy Cartwright, *Nature's Capacities and Their Measurement* (Oxford: Clarendon, 1989).

⁷ David Armstrong, *What Is a Law of Nature?* (New York: Cambridge University Press, 1983); Michael Tooley, "The Nature of Laws," *Canadian Journal of Philosophy* 7 (1977): 667–698; Fred Dretske, "Laws of Nature," *Philosophy of Science* 44 (1977): 248–268; Bas van Fraassen, *Laws and Symmetry* (New York: Oxford University Press, 1989); and Barry Loewer, "Humean Supervenience," *Philosophical Topics* 24 (1996): 101–127.

⁸ C. D. Broad, "Determinism, Indeterminism, and Libertarianism," in *Ethics and the History of Philosophy* (London: Routledge & Kegan Paul, 1952).

model of causality to develop an account of agent causation that, unlike Taylor's, Chisholm's, and O'Connor's, does not require the exercise of an agent's will to be viewed as a temporally determinate event.⁹ Such an understanding allows one to avoid the infinite regress that ensues if the exercise of an agent's will is itself caused, while also permitting one to deny that it is an uncaused event, which would seem to place it beyond the agent's control.

But what is most striking about Kant's account of causality is not merely its remarkable versatility in being able to contribute to several different contemporary philosophical debates, nor the way in which it appeals to the notion of an activity in each case, but rather the fact that it points to the possibility of a *comprehensive metaphysical* account of the world, one that covers both nature and freedom and that, by being based on a notion of activity, represents a *systematic* alternative to empiricist (e.g., Humean) positions. Insofar as an account with this kind of scope remains a desideratum for contemporary nonempiricist accounts, Kant's position represents an ideal that can provide significant guidance to us today, just as Kant himself had hoped to do over two centuries ago.

⁹ Richard Taylor, *Action and Purpose* (Englewood Cliffs, N.J.: Prentice Hall, 1966); Roderick Chisholm, "Human Freedom and the Self," in *On Metaphysics*, ed. R. Chisholm (Minneapolis: University of Minnesota Press, 1989), "Freedom and Action," in *Freedom and Determinism*, ed. K. Lehrer (New York: Random House, 1966), and *Person and Object* (LaSalle, Ill.: Open Court, 1976); and Timothy O'Connor, *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000).

PART ONE

CAUSALITY IN CONTEXT

The first part of this study presents the historical context to Kant's Critical views on causality. Chapter 1 ("Pre-established Harmony versus Physical Influx") discusses how the issue of causality was understood and developed in response to changing intellectual conditions in Germany in the first half of the eighteenth century by figures such as Leibniz, Wolff, Knutzen, Baumgarten, Meier, and Crusius. Chapter 2 ("Kant's Pre-Critical Theory of Causality") investigates how Kant reacted to this context and developed an increasingly independent position on the issue of causality throughout his pre-Critical period. These two chapters, taken together, provide the proper historical background for attaining an accurate understanding of Kant's views on causality in the Critical period, the subject of the second part of this study.

Pre-established Harmony versus Physical Influx

INTRODUCTION

To understand Kant's views on causality throughout his career, one should perhaps begin at the beginning. When Kant interrupted his studies at the university to settle his father's estate and was in the process of publishing his first book in 1746, what might his background knowledge on the issue of causality have included? What figures would have been important to him and what positions would have represented the primary options? What arguments, objections, and issues would he have thought of as standing in need of engagement or further analysis? What would have been perceived as a significant contribution to the state of the debate at this time? That is, what questions had not been answered satisfactorily and what questions still needed to be asked? It is the point of this chapter to explore some of the main currents in German philosophy pertaining to the issue of causality from Leibniz's death, in 1716, to Kant's first publication just over three decades later.

Since Leibniz's influence during the first half of the eighteenth century in Germany was pervasive, it is necessary to start with his views on causality in order to see the ways in which later figures, even his "followers," would appropriate different aspects of his position, modifying them to their own ends and specifying them in ways that he might not have condoned. On the issue of causality, Leibniz is the first or, at the very least, most famous proponent of a doctrine called "pre-established harmony." Its fundamental tenet is that finite substances can act not on each other, but only on themselves, and that in his infinite wisdom God programmed them prior to creation with all of their future states in such perfect

harmony that they nonetheless appear as if they were interacting with each other. Pre-established harmony represented not only one of the three main philosophical positions regarding causation during the last twenty years of Leibniz's lifetime and in the decades following his death, but also the orthodox position among German Enlightenment philosophers for a number of years in the first half of the eighteenth century. As a result, Leibniz's thoughts on pre-established harmony provide the crucial intellectual starting point for the most important philosophical reflection on causality in eighteenth-century Germany by figures such as Wolff, Knutzen, Baumgarten, Meier, and Crusius, who – along with Hume, whose views Kant would become familiar with by the early 1760s – form the most important background for Kant's views on causality throughout his career.

In this chapter, after explaining several features of Leibniz's doctrine of pre-established harmony and the fuller context in which it is situated within Leibniz's own philosophy in the seventeenth century, we consider how this doctrine was received in eighteenth-century Germany. What deserves special attention in this regard was the development of the main rival to pre-established harmony, a doctrine stating that simple finite substances can act on each other, just as they seem to do. With some notable exceptions (Wolff, Baumgarten, and Meier in particular), those who accepted pre-established harmony typically had little to say that went beyond what Leibniz had asserted, whereas those who rejected pre-established harmony (Knutzen and Crusius) were in a much better position to make contributions, since they needed to develop detailed reasons for rejecting that doctrine, reasons that required either a reinterpretation of Leibnizian metaphysics or a complete rejection of it. These developments thus reveal what issues Kant would have been focusing on in developing his own view of causality within that tradition.

LEIBNIZ AND THE MANY FACES OF PRE-ESTABLISHED HARMONY

In the second half of seventeenth-century Europe, three causal theories vied for acceptance. The first view was often associated with Descartes, Locke, and various (e.g., medieval) Aristotelians and asserted that finite substances can act on each other causally. Leibniz dubbed this view *influxus physicus*, or physical influx, which, if translated literally, would be “natural influence,” since the view holds that in causal interaction one

finite substance “flows into” or “influences” another by its very nature.¹ As this interactionist view encountered various difficulties – for example, in explaining how the mind and the body could act on each other or how it is that creatures could truly be said to act without detracting from their complete dependence on God – a second view, which held that God alone is a true or real cause, came to appear attractive. On this view, finite substances do not act at all, but rather are merely occasions for God to act in the specific ways that he does. This view was often called occasionalism or, following Leibniz, the way of assistance. To avoid what seemed to be the perpetual miracles implied by this second view, Leibniz developed his own view that finite substances must act causally (since such activity is required to distinguish finite substances from states of God), but not on each other. Rather, each substance acts only on itself according to the law God gave it before creation, so that it unfolds according to that plan in perfect harmony with the plans of all other substances. Since God establishes the harmonious plans of all finite substances prior to creation, Leibniz calls this theory “pre-established harmony.”

While Leibniz often uses the term “pre-established harmony” to assert that simple substances can act not on each other, but only on themselves, he also uses it in other senses in other contexts.² For example, he uses it to explain the mind-body relationship, which is not, properly speaking, a relationship between two simple substances at all insofar as he holds that bodies are merely “well-founded phenomena” and thus not themselves simple substances.³ Leibniz also mentions pre-established harmony at times in order to refer to the claim that the realm of efficient causality, which pertains to relations between bodies, harmonizes with the realm of final causality, which governs the workings of minds.⁴ And there are yet further meanings.⁵ Part of the reason why Leibniz uses the term “pre-established harmony” in such different ways is that he, perhaps more than any other modern philosopher, develops his own position while

¹ For discussion of which views Leibniz might have had in mind in thinking about *influxus physicus*, see Eileen O’Neill’s “*Influxus Physicus*,” in *Causation in Early Modern Philosophy*, ed. S. Nadler (University Park: Penn State University Press, 1993), pp. 27–55.

² See, e.g., “A New System of Nature,” in G. W. Leibniz, *Philosophical Essays*, ed. and trans. R. Ariew and D. Garber (Indianapolis: Hackett, 1989), p. 143.

³ See, e.g., *ibid.*, p. 144.

⁴ See, e.g., “A Specimen of Dynamics,” in *Philosophical Essays*, p. 126, but also “The Monadology,” props. 78–79, in *Philosophical Essays*, p. 223.

⁵ For example, he occasionally seems to mean by pre-established harmony the view that everything agrees with or expresses the entire world. See Leibniz, “New System of Nature,” p. 144.

criticizing the works of others and wants to invoke wherever he can the basic idea of “harmony” as representing a theoretical virtue of his position. Yet despite these differences, there are also important connections between these various uses, and it will prove to be helpful to understand both how they emerge from his detailed criticisms of the views of his predecessors and how they fit into the broader metaphysical context of his other philosophical doctrines.

Leibniz on the Mind-Body Problem in Descartes

When Leibniz discusses physical influx, he is often thinking of Descartes’s interactionist account of the mind-body relationship as one of its most prominent instances. According to Descartes, the mind and the body, as finite substances, are both capable of independent existence, yet their natures are radically distinct insofar as thought is the principal attribute of mind and extension the principal attribute of body. However, in human beings they stand in causal interactions with each other that are so close as to constitute, or at least to indicate, a special kind of metaphysical union. Via the pineal gland the body can cause sensations in the mind and the mind can change (the direction of) the motion of its body by exercising its free will. In light of these causal connections and the way in which the mind perceives its body (which is fundamentally different from the way in which a sailor perceives his ship), it is clear that the mind is conjoined to its body in a special manner.

In the “New System of Nature,” which Leibniz published in the *Journal des Savants* in 1695, he attempted to reveal some of the inadequacies of Descartes’s account of matter in order to present his own positive view of what else (beyond extension) is required for matter to exist – namely substantial forms, formal atoms, or metaphysical points (all of which possess a special kind of unity, a unity that extension lacks) – but he also explicitly rejected Descartes’s account of mind-body interaction. On the question of *how* minds and bodies could interact with each other at all, Leibniz remarked: “Descartes had given up the game at this point, as far as we can determine from his writings.”⁶

While it was common at the time to doubt whether substances with such radically different natures as thought and extension could act on each other, Leibniz developed these doubts in greater detail. First, it is not sufficient merely to cite the difference in nature between mind and body.

⁶ *Ibid.*, p. 143.

After all, why should the fact that two substances have radically distinct natures make any immediate difference as to what they can and cannot do? Rather, one must explain why a difference in the natures of these substances precludes their causal interaction. Leibniz, unlike many other critics of Cartesian interactionism, provided just such an explanation by noting that since their natures are radically distinct, there can be no proportion between them, and if there is no proportion between them, then there can be no intelligible connection between any state of the mind and any particular velocity of a body.⁷

Leibniz also objected that allowing mind and body to act on each other would violate the laws of nature as Descartes understands them, the law of the conservation of motion in particular. If the mind could exercise its free will by acting on the body, there would be more motion in the world after its exercise than beforehand.⁸ Similarly, if bodies caused thoughts in the mind, there would be less motion after such interaction than before it insofar as the body would expend some of its reality (its motion) in causing the thought. Leibniz thus developed two detailed objections to Descartes's account of mind-body interaction.⁹

However, Leibniz's disagreement with Descartes's position is even more fundamental than is suggested by these two objections. For Leibniz also argued that it is not even metaphysically possible for finite substances to act on each other causally, regardless of any differences between their principal attributes. If one substance acts on another so as to change its state, this implies that the first substance is causally responsible for a new accident in the other substance. How did that new accident get there? Did the first substance give up one of its accidents to the second substance (as when one billiard ball is said to "transfer" its motion to the other at the moment of collision)? Or did the first substance create such an accident

⁷ See his "Letters to Arnauld," in *Philosophical Essays*, p. 83. In this passage, Leibniz also suggests that this problem applies to occasionalism insofar as finite substances have no nature that would give God reason to take the state of the mind at one moment as an occasion for creating the body with any specific velocity rather than any other at the next moment in time.

⁸ According to Leibniz, Descartes attempted to distinguish between the speed and the direction of a given body in motion (so that the speed can be conserved even if the direction is not as would happen, e.g., in cases of free will). However, Leibniz objects to the distinction, claiming that speed and direction must be taken together. See Robert Sleigh, *Leibniz and Arnauld* (New Haven: Yale University Press, 1990), pp. 140–141, for a discussion of the textual basis for Leibniz's attribution of this distinction to Descartes.

⁹ In the "New System of Nature," Leibniz also suggests that elastic bodies do not act on each other in collisions; rather, so Leibniz asserts, each body is the cause of its own motion in the opposite direction.

ex nihilo? Neither option appears to represent a genuine metaphysical possibility.

As for the first option, Leibniz thought that it is untenable to maintain that an accident could literally “migrate” or be transferred from one substance to another, since accidents are defined as dependent beings and their dependency would conflict with the free-floating or independent state that they would be in “after” they had left the one substance and “before” they had arrived at the other. As Leibniz famously put it in “The Monadology”: “The monads have no windows through which something can enter or leave. Accidents cannot be detached, nor can they go about outside of substances.”¹⁰ Similarly, in “A New System of Nature,” Leibniz remarked that “the action of one substance on another is neither the emission nor the transplanting of an entity.”¹¹

But the second option would have appeared even less promising than the first, insofar as it would have been considered heretical to claim that finite substances might have quasi-divine powers of creation.¹² Yet if neither of these options is a real metaphysical possibility, then physical influx appears to be a theory that not only encounters significant difficulties in explaining one of its particularly important instances – interaction between the mind and the body – but also has a gaping hole at its very heart. Its main assertion is that finite substances can act on each other, but it seems unable to answer the first question one would want to pose: How is it even possible, metaphysically speaking, for one finite substance to act on another?

Leibniz’s Objections to Occasionalism

If physical influx is unacceptable – whether taken in the guise of Descartes’s mind-body interaction or considered in its most basic metaphysical sense – which alternative ought to be preferred? Historically, many Cartesians, most notably Malebranche, ended up adopting occasionalism. If finite substances cannot act on each other, it would be natural

¹⁰ Leibniz, “The Monadology,” prop. 7, *Philosophical Essays*, p. 214. A less famous, but equally nice statement of this point can be found in the “Primary Truths”: “no created substance exerts a metaphysical action or influx on any other thing. For . . . one cannot explain how something can pass from one thing into the substance of another,” *Philosophical Essays*, p. 33.

¹¹ *Philosophical Essays*, p. 145.

¹² In book six, part two, chapter 3 of his *The Search after Truth*, Malebranche remarks: “God cannot even communicate his power to creatures, if we follow the lights of reason; He cannot make true causes of them, He cannot make them gods.” Nicolas Malebranche, *Philosophical Selections*, ed. S. Nadler (Indianapolis: Hackett, 1992), p. 97.

to think, at least in the context of seventeenth-century Europe, that they cannot act at all and that God would be responsible for all causal activity. Malebranche seems to be motivated primarily by two lines of argument. First, if a cause necessarily brings about its effect, then only God could be a true cause, since only God's volitions are necessarily efficacious.¹³ Second, if one accepts the view that conservation is constant re-creation (i.e., that when God conserves the world, he does so by constantly re-creating it anew at each and every instant), then God is completely causally responsible for what occurs at every moment in time.¹⁴ Any contribution that a finite substance might make would be incompatible with God's infinite creative power and ought not to be accepted. Thus, on Malebranche's view, it would seem that finite substances are neither powerful enough nor even needed to be causes in the first place.

Though Leibniz was tempted by occasionalism much more than he ever was by physical influx, he nonetheless ultimately found it unacceptable.¹⁵ One objection to occasionalism, developed at length in "On Nature Itself," is based on the identity conditions of finite substances.¹⁶ If occasionalists maintain that God brings about every (noninitial and non-miraculous) state of every finite substance by means of his own (general) volitions (and not by means of the natures he might bestow on them), then God does not leave what Leibniz refers to as "traces" in substances that could be efficacious for their future states. The entire efficacy rests

¹³ In book six, part two, chapter 3, Malebranche argues as follows: "A true cause as I understand it is one such that the mind perceives a necessary connection between it and its effect. Now the mind perceives a necessary connection only between the will of an infinitely perfect being and its effects. Therefore it is only God who is the true cause." *Ibid.*, p. 96.

¹⁴ Malebranche offers this line of argument in his reply to the seventh proof contained in *Elucidation Fifteen*, which is appended to book six, part two, chapter 3 of his *The Search after Truth*: "Almost all theologians speak as follows: that the action of secondary causes is not different from the action by which God cooperates with them. For although they understand it in different ways, they hold that God acts in creatures through the same action as do creatures. And they are obliged to speak this way, it seems to me; for if creatures acted through an action God did not produce in them, their action qua efficacious action would be, it seems to me, independent; now they believe, as they must, that creatures depend immediately on God, not only for their being, but for their operation as well." Malebranche, *Philosophical Selections*, pp. 120–121.

¹⁵ In fact, Leibniz seems briefly to endorse occasionalism in 1675 in a letter to Foucher (see *Philosophical Essays*, pp. 1–5, esp. p. 3), contra Christia Mercer's reading in *Leibniz's Metaphysics: Its Origin and Development* (New York: Cambridge University Press, 2001). For a helpful discussion of when Leibniz came to accept pre-established harmony, see Paul Lodge's "Leibniz's Commitment to the Pre-established Harmony in the Late 1670s and Early 1680s," *Archiv für Geschichte der Philosophie* 80 (1998): 292–320.

¹⁶ For an extended discussion of the issues that surround this line of objection, see chapter 6 of Sleigh, *Leibniz and Arnauld*, pp. 116–136.

with God, not with the finite substances. But if God leaves no traces in the state of a substance that would lead it to its next state, then it is unclear what is supposed to connect those states so that they can be understood as different states of one and the same substance. As Leibniz puts it: How is it “that things themselves can endure through time,” if they do not act according to enduring natures that would be able to establish such a link?¹⁷ Moreover, Leibniz suggests that without invoking any causally active principle that would individuate finite substances, occasionalism collapses into Spinozism, which was considered a heretical or “profane” view at the time.¹⁸ For in Leibniz’s mind, substantiality and activity were necessarily connected, and if one dispensed with the causal activity of finite substances (as occasionalism does), one would also be forced to dispense with their substantiality as well, leaving God as the only active substance and the world as his states. Since Leibniz was convinced that causal activity is an essential feature of substantiality, he rejected occasionalism.

A second difficulty that Leibniz raised for occasionalism is that it seems to be inconsistent with human freedom and appears to entail the theologically unorthodox claim that God must be the cause of evil.¹⁹ Leibniz’s line of thought is that the exercise of free will is most naturally understood as a causal activity, in fact, a causal activity of a very special kind, which he dubs “spontaneity.” Accordingly, by denying that finite substances are causally efficacious at all, occasionalists are forced to deny, so Leibniz thinks, that human beings could act spontaneously or freely.²⁰ Moreover, if human beings lack spontaneity, freedom, or any causal activity at all, it would seem to follow that only God could be the cause of evil. These are positions that Leibniz, as an orthodox theist, cannot accept.

The objection that Leibniz raised against occasionalism more often than any other, however, is that occasionalism would require that God perform perpetual miracles. If one defines a miracle not in terms of frequency (as Hume and Clarke do from within an empiricist framework), but rather as an event that cannot be caused by finite substances according to their natures, then, because occasionalism holds that no events are caused by finite substances according to their natures and that God is the sole and immediate cause of all events, an occasionalist would

¹⁷ “On Nature Itself,” *Philosophical Essays*, p. 158.

¹⁸ *Ibid.*, p. 160.

¹⁹ See, e.g., *ibid.*, p. 161.

²⁰ This objection would seem to beg the question against occasionalists by presupposing that free will must be understood in terms of causal activity.

be committed to all events being miraculous (which would have been considered an unorthodox position at the time).²¹ To avoid this consequence Malebranche defined a miracle not as an event that is caused immediately by God, but rather as one that is caused by a *particular* volition in God. Natural events, by contrast, are caused by God's *general* volitions (which are thus supposed to be functionally equivalent to laws of nature). Malebranche thus attempted to defend occasionalism by distinguishing between natural and miraculous events in terms of general and particular volitions. Leibniz objected to Malebranche's distinction by denying that God could ever act according to particular volitions: "As God can do nothing without reasons, even when he acts miraculously, it follows that he has no will about individual events but what results from some general truth or will. Thus I would say that God never has a *particular will*."²² If this objection is correct, then occasionalism is saddled not so much with perpetual miracles as with something equally unacceptable, namely the impossibility of any miracle.

If God acts according to volitions that are absolutely general, Leibniz's own position might seem to be subject to the very same difficulty that he raised for occasionalism.²³ If miracles cannot be explained through recourse to particular volitions, how are they to be explained? And if they can be explained in terms of general volitions, why could an occasionalist not adopt that very same explanation in order to avoid Leibniz's charge of perpetual miracles? The difference between Leibniz and Malebranche on this point ultimately derives from their different conceptions of what a law of nature is. For Malebranche, any regularity that God would decide to bring about would presumably qualify as a law of nature. For Leibniz, by contrast, not only must a law of nature involve a regularity of some sort, but that regularity must also be based on the natures of things.²⁴

²¹ Hume defines miracles in terms of frequency in his *Inquiry Concerning Human Understanding*. Leibniz explicitly objects to Clarke's understanding of miracles in terms of "usualness" in his fourth letter to Clarke. See *Philosophical Essays*, p. 331.

²² §206. *Theodicy*, trans. E. M. Huggard (LaSalle, Ill.: Open Court, 1985), p. 256. Leibniz presents a similar line of thought in "On Nature Itself."

²³ See Sukjae Lee, "Leibniz on Individual Substances and Causation: An Account of Divine Concurrence," Ph.D. diss., Yale University, 2001, for a detailed discussion of how Leibniz attempts to distinguish his view from Malebranche's occasionalism.

²⁴ As Leibniz puts the point in the *Theodicy*, §207: "The distinguishing mark of miracles (taken in the strictest sense) is that they cannot be accounted for by the natures of created things. That is why, should God make a general law causing bodies to be attracted the one to the other, he could only achieve its operation by perpetual miracles." *Theodicy*, p. 257.

That is, there must be an intelligible, nonarbitrary connection between what happens to a thing and its nature (or the kind of thing that it is) in order for that event to be a natural event.²⁵ Given this account of the laws of nature, a state occurs naturally for Leibniz just in case it follows (at least in part) from the nature of a substance, while a miraculous event has no causal connection to what a finite substance is capable of producing according to its nature.

In asserting that a substance can act according to its own enduring nature, however, Leibniz was forced to come to terms with a problem occasionalists do not face. For Leibniz needed to explain how finite substances can make some sort of causal contribution to the world, despite the fact that God is supposed to create everything *ex nihilo* at all times. How can God and finite substances truly cooperate or engage in a joint causal venture, given what would appear to be the complete dependence of the one on the other? That substances owe their existence to God as the creator of everything that is real in them, but still somehow retain enough independence to be able to make a genuine causal contribution to the state of the world, was an old and orthodox idea, but articulating it with enough precision to allow one to distinguish it from occasionalism proved to be a difficult, even if stimulating task for Leibniz in his philosophical theology.

Leibniz's Positive Case for Pre-established Harmony

Since both physical influx and occasionalism encountered what Leibniz took to be insuperable difficulties, only his own doctrine of pre-established harmony remained as a viable option. But Leibniz also thought that a positive case could be made for it, a case that has at least three distinct lines of argument. The first, and perhaps most important, line of argument is based on Leibniz's conception of a substance.

²⁵ As Donald Rutherford helpfully points out ("Natures, Laws, and Miracles: The Roots of Leibniz's Critique of Occasionalism," in *Causation in Early Modern Philosophy*, pp. 146–147), the difference between Leibniz and Malebranche rests, in part, on a difference in what it means for the world to be intelligible. From Leibniz's perspective, Malebranche's account is defective since the world is chosen arbitrarily by God, whereas for Leibniz, God chooses to create the best possible world for a very good reason, namely the fact that it is fully intelligible in virtue of the natures that God bestows on things. There is also a connection between the intelligibility of natures and Leibniz's criticism of Cartesian interactionism, since it is precisely the difference between the natures of the mind and the body that precludes establishing a fully intelligible connection between them.

Like Descartes before him, Leibniz thought that a finite substance must be able to exist independently of any other finite substance. However, “independence” is a merely negative, relational characterization of substance, expressing the idea that one entity does *not* require *another* in order to exist. Such a negative relational characterization must ultimately be derivative of a positive, intrinsic characterization of substance. Accordingly, what Leibniz thought underlies the negative relational conception of substance he shared with Descartes is the idea that a substance must be self-sufficient. For it is precisely because a substance is self-sufficient that it is independent, that is, does not stand in need of any other finite entity. Yet Leibniz came to think that self-sufficiency implies pre-established harmony.²⁶ For if a substance is truly self-sufficient, then it will suffice to cause all of its own states, that is, no other substance is needed to bring about any of its states, which is exactly what pre-established harmony holds.

Leibniz also thought that pre-established harmony is supported by his doctrine of complete concepts along with his view that every predicate that is true of a thing must be grounded in that thing. In §8 of the *Discourse on Metaphysics*, in order to distinguish pre-established harmony from occasionalism, Leibniz begins by noting that every true predication requires that the subject term must contain the predicate term. He then suggests as a consequence of this that every individual substance must have a complete concept. A concept of a substance is complete if for every possible predicate, the substance contains either it or its complementary predicate. Since “every possible predicate” includes all predicates at all times, every complete concept must contain all the predicates that have been, are, or ever will be true of its individual substance. In §9 of the *Discourse on Metaphysics* Leibniz infers as a further consequence that no two substances can share the same complete concept, that is, that two things that share all of their predicates are numerically identical, a principle commonly known as the principle of the identity of indiscernibles. But if every substance has a complete concept and, moreover, contains within itself the grounds for each and every predicate contained in that concept, then it seems to follow that the substance is the cause of each and every one of its predicates or properties, which is exactly what pre-established harmony maintains. As Leibniz remarks of Alexander the Great, “when we consider carefully the connection of things, we can say

²⁶ See, e.g., “New System of Nature,” *Philosophical Essays*, p. 144.

that from all time in Alexander's soul there are vestiges of everything that has happened to him and marks of everything that will happen to him."²⁷ Accordingly, pre-established harmony follows from his complete concept theory of substance and his distinctive understanding of the grounds of predication.²⁸

Leibniz also pursues the following two-step line of argument in favor of pre-established harmony in "A New System of Nature." (1) He begins by suggesting that pre-established harmony is at least a *possible* metaphysical position. That is, God *could have* created the world as pre-established harmony describes it. Certainly, no orthodox theist would be eager to assert that God might lack either the power or the knowledge to create such a world, if he so desired. (2) Once the mere possibility of pre-established harmony has been conceded, the question then arises as to which of the three causal theories is more probable. Leibniz thinks that this question can be answered by determining which one has the most significant advantages. After raising some of the objections to occasionalism and physical influx that were discussed above (which would make those theories improbable), he suggests that pre-established harmony is more probable due to its numerous advantages: (i) pre-established harmony satisfies one of the necessary conditions of freedom, since no substance is constrained by the causal activity of any other finite substance; (ii) it is most conducive toward maintaining the immortality of the soul, given that nothing external to it (other than God) could act on it so as to destroy it; and (iii) it provides the materials for a new proof of the existence of God, since "the perfect agreement of so many substances which have no communication among them can only come from the common source."²⁹ For these reasons, Leibniz suggests, pre-established harmony is more probable and thus preferable to occasionalism and physical influx.

Pre-established Harmony in the Context of Leibniz's Broader Metaphysics

To appreciate, however, the full range of uses to which Leibniz puts the term "pre-established harmony," it is crucial to see how his views

²⁷ *Philosophical Essays*, p. 41.

²⁸ Leibniz presents a similar argument in the "Primary Truths," *Philosophical Essays*, pp. 30–33.

²⁹ "New System of Nature," *Philosophical Essays*, p. 145.

on the causality of finite substances fit into his broader metaphysics. First, although the matter is complex, especially during his middle years, Leibniz is ultimately an idealist. That is, Leibniz thinks that the basic constituents of everything that exists, that is, simple substances, are nothing other than minds, souls, or what, beginning in 1695, he calls monads. As such, a monad will necessarily have “perceptions and appetitions,” where an appetite is the teleological action of an internal principle of a monad whereby it changes from one perceptual state to another, striving to perfect itself in accordance with its conception of the good. Since all minds represent one and the same world in its entirety (a doctrine sometimes referred to as the universal expression thesis), what distinguishes monads (given that they all represent the same world and thus cannot be distinguished in terms of the *object* they represent) are the different points of view and/or the varying degrees of clarity and distinctness that are inherent in their representations of the world (i.e., features of their *subjective* standpoints). Further, Leibniz holds that the points of view from which monads represent the world are not literally physical. Nor could they be mathematical, since mathematical points are mere abstractions, in contrast to the reality of monads. As we saw above, Leibniz sometimes calls them “metaphysical points,” “substantial forms,” or “formal atoms,” depending on whether he is interested in appealing to notions stemming from mathematics, Aristotelian metaphysics, or Lockean/Newtonian physics.

In attempting to explain the ultimate metaphysical status of monads in greater detail, Leibniz asserts (most clearly in the “Specimen Dynamicum,” but also in “On Nature Itself”) that monads are constituted by primitive active and passive forces. As he puts it, “the very substance of things consists in a force for acting and being acted upon.”³⁰ To illustrate the distinction, Leibniz sometimes appeals to a modified version of the Aristotelian distinction between substantial forms and primary matter. A primitive active force is like a substantial form, since both are essential activities of an enduring substance. Similarly, a primitive passive force is like primary matter insofar as both are often understood as a principle of resisting the activity of substantial forms or primitive active forces. In light of Leibniz’s idealist commitments, it stands to reason that the force of appetite is identified with the primitive active force of a monad, since it is supposed to bring about its various perceptions and desires actively

³⁰ “On Nature Itself,” *Philosophical Essays*, p. 159.

according to the law of its nature (e.g., in a rational monad, according to its conception of the good). Similarly, it would be natural to think that primitive passive force would be responsible for the limitations of our representations, that is, for the fact that they are confused and obscure (especially in regards to what is truly good) and do not fully attain their object all at once.

Leibniz's account of primitive forces is intimately connected with his understanding of what he calls derivative forces. Derivative forces are to be understood, Leibniz tells us, as limitations or modifications of primitive forces. While derivative forces can thus be divided into active and passive kinds just as primitive forces are, what is particularly important about them is that they form the foundation of Leibniz's physics. Thus, in his "Specimen Dynamicum," Leibniz provides a mechanistic account of some of the most basic properties of bodies in terms of derivative forces. For example, he says that derivative active force causes a body's velocity and acceleration, while derivative passive force is responsible for a body's resistance to such motions and accelerations. Leibniz expands on his account of primitive passive force in this context by claiming that it "is that by virtue of which it happens that a body cannot be penetrated by another body."³¹ Elsewhere, he indicates that it is the source of extension and divisibility in addition to impenetrability.

Given the existence of both primitive and derivative forces, it is natural to investigate the precise nature of their relationship. In the "Specimen Dynamicum," Leibniz asserts that derivative forces can be seen as "resulting from a limitation of primitive force through the collision of bodies with one another."³² But what "resulting" means and in what sense a derivative force must be a "limitation" of a primitive force is not elaborated on. As a result, the crucial connection between what is ultimately real, namely monads, and the physical properties of the world investigated by natural science rests on notions for which Leibniz does not give a particularly clear or detailed public accounting.³³ However, Leibniz is convinced that whatever happens mechanically at the derivative level of bodies will harmonize with whatever happens with teleologically driven representations at the primitive level of monads. In other

³¹ *Philosophical Essays*, p. 120.

³² *Ibid.*, p. 119.

³³ For a discussion of the complexities of this issue in Leibniz, see Robert M. Adams *Leibniz: Determinist, Theist, Idealist* (New York: Oxford University Press, 1994), esp. chapter 13.

words, the realm of final causes (monads) must harmonize with the realm of efficient causes (bodies). While this view is clearly not identical to the claim that finite substances can act only on themselves, one can see, given Leibniz's idealism and his account of the relationship between the primitive forces of monads and the derivative forces of bodies, why he would be struck by the harmony between these realms, a harmony God must have pre-established in creating the best possible world.

Leibniz's idealism and his account of primitive and derivative forces also reveal why he might think that pre-established harmony would be a solution to the mind-body problem. Because a monad is a principle of unity and a body is a multitude (insofar as it is understood merely as extended), Leibniz argues that bodies can be real only if supplemented by the unity inherent in monads. As a result, Leibniz thinks that "we can thus understand how the soul has its seat in the body by an immediate presence which could not be greater, since the soul is in the body as unity is in the resultant of unities, which is a multitude."³⁴ In other words, Leibniz thinks that the mind-body relationship is no more problematic than any other relationship between a being that has unity and those that lack such unity. For Leibniz there is nothing particularly mysterious about that sort of relationship since it simply reflects the basic metaphysical principle that being and unity require each other.

By thus taking pre-established harmony in the broader context of Leibniz's metaphysics, one can understand his many uses of the term "pre-established harmony." Its primary use is that since finite substances can act not on each other, but only on themselves, the agreement in their states must be pre-established by God. However, it can also be used to express the relationship between the primitive forces that constitute monads and the derivative forces that pertain especially to bodies, insofar as God pre-establishes a harmony between the realms of final and efficient causality. Finally, it also applies to the mind-body problem, since that problem is simply a particular instance of the previous use, given his view that bodies governed by efficient causality are simply well-founded phenomena in virtue of being grounded in simple, immaterial substances that act according to inherent teleological principles.

³⁴ "New System of Nature," *Philosophical Essays*, p. 144.

CHRISTIAN WOLFF AND PRE-ESTABLISHED HARMONY IN
EIGHTEENTH-CENTURY GERMANY

A Radical Shift in Context: From Seventeenth-Century
Europe to Eighteenth-Century Germany

How was Leibniz's doctrine of pre-established harmony received in eighteenth-century Germany? Unsurprisingly, the first significant moment centers on Christian Wolff (1679–1754), for Wolff is inextricably linked with Leibniz. Leibniz and Wolff corresponded closely for many years about a variety of philosophical and mathematical issues; Leibniz was instrumental in Wolff being offered a professorship in mathematics at the university in Halle; Wolff was professionally useful to Leibniz both due to his friendly contributions to the *Acta Eruditorum* and in helping Leibniz to mount a defense in the latest stages of the priority dispute over the discovery of the calculus;³⁵ and, finally, their doctrinal differences were sometimes thought to be minimal enough that their views could be referred to jointly as the "Leibnizian-Wolffian philosophy."³⁶ At the same time, Wolff rarely fares well when his philosophical talents and achievements are compared with Leibniz's. Wolff is typically seen as nothing more than Leibniz's follower and as deserving of credit merely for the way in which he popularizes Leibniz's views. Moreover, he made no significant philosophical discoveries of his own and the level of his argumentation is entirely lacking in the rich, albeit often enigmatic profundity of Leibniz's texts.

Without suggesting that scholars have overlooked hidden philosophical talent in Wolff that requires a thorough reassessment of his place in the history of philosophy, we must nevertheless understand that a radical intellectual shift took place between the periods of Leibniz's and Wolff's mature thought, a shift that is crucial to attaining a proper appreciation of Wolff's achievements and the philosophical landscape in eighteenth-century Germany. To see the main contours of this shift, consider the differences between Leibniz's situation late in the seventeenth century,

³⁵ See C. I. Gerhardt, *Briefwechsel zwischen Leibniz und Christian Wolff* (1860; rpt. Hildesheim: Georg Olms Verlag, 1971), who claims (p. 11) that Wolff became "unentbehrlich" for Leibniz, esp. between 1711 and 1714, when the calculus dispute broke out again.

³⁶ This is not to say that Leibniz's and Wolff's positions are identical. Lewis White Beck, *Early German Philosophy: Kant and His Predecessors* (Cambridge: Harvard University Press, 1969), pp. 256–275, and Charles Corr, "C. Wolff and Leibniz," *Journal of the History of Ideas* 36 (1975): 241–262, both discuss ways in which Wolff's position may be unlike Leibniz's. As we see below, they differ significantly regarding pre-established harmony as well.

and then around the time of his death in 1716. It is striking that the defining features of Leibniz's thought stem mainly from his European connections outside Germany that originated and took shape in the seventeenth century. Perhaps his most formative experiences occurred during his so-called Paris years (1672–1676), since it was there that he was introduced to and interacted with many of the leading intellectuals of his day, encountering a wealth of stimulating and provocative new ideas. Many of his most important writings were directed at a European audience either by reacting to their views in the leading academic journals (such as the *Acta Eruditorum* and the *Journal des Savants*) or through personal correspondence.

Moreover, one must draw a distinction between the “public” and the “private” Leibniz, for there is a significant difference between his ultimate views and those views he chose to share with the broader philosophical readership.³⁷ Some of Leibniz's most remarkable and detailed statements of his position occur in writings that were not published during his lifetime: the “Discourse on Metaphysics,” the “Primary Truths,” “The Monadology,” and his correspondence with various prominent figures, including Arnauld, Huygens, De Volder, Des Bosses, and the Bernoullis. The only book he published in his lifetime, the *Théodicée*, was printed in Amsterdam in 1710, but its contents are clearly directed toward French intellectuals and stem from topics in philosophical theology most at home in the seventeenth century. Although some of his writings, for example, “The Monadology,” became available to the public after his death in the eighteenth century, this would in no way be sufficient to overthrow certain entrenched ways of interpreting Leibniz.

By contrast, the eighteenth-century was, by and large, much less kind to Leibniz. He had lost political favor, due in part to nationalistic fallout from his priority dispute with Newton about the discovery of the calculus and in part to the death in 1705 of Sophie Charlotte, queen of Prussia, who had been a constant supporter of his at court in Berlin. Philosophically, many of the figures with whom Leibniz had corresponded and against whose views he had made such penetrating and original criticisms had died or were no longer at the forefront of new philosophical developments. Also, Leibniz had not published any systematic statement of his

³⁷ See Emile Ravier, *Bibliographie de Oeuvres de Leibniz* (Hildesheim: Georg Olms, 1966), for a complete listing of works Leibniz published during his lifetime. Although Leibniz may have published a tremendous amount, which of these publications were well known among the broader philosophical public and how many of them conveyed his fundamental philosophical ideas are further questions.

views in a textbook that could be used for easy dissemination, for example, in the lecture setting of courses at the university.³⁸ Further, Leibniz had written primarily in French and Latin, rarely in German, which became increasingly important in both popular and academic settings in Germany.

Finally, as noted above, it is important to keep in mind how little of his work was published and generally known by this time. While his “Discourse on Metaphysics” and his correspondence with various prominent figures are standard sources for our current knowledge of Leibniz’s position, for the most part they had neither been published nor become public through other means in early to mid-eighteenth-century Germany. In addition to the *Théodicée*, Leibniz’s position became known in this particular context primarily as it was represented in his “Meditations on Knowledge, Truth, and Ideas” (1684), “On an Emendation of First Philosophy and the Notion of Substance” (1694), “A New System of Nature” (1695), Part I of the “Specimen Dynamicum” (1695), and “On Nature Itself” (1698), which were published in either the *Acta Eruditorum* or the *Journal des Savants*. The *Leibniz-Clarke Correspondence*, published after his death in 1717, was another regularly quoted source, one that would receive renewed attention later in the century after it was republished in 1768. His “Monadology” appeared in 1720 and the *New Essays* was published only in 1765.³⁹ Accordingly, due to the decline in his political standing, his failure to publish a systematic statement of his views in accessible German, and his focus on a seventeenth-century European (especially French) rather than eighteenth-century German audience, Leibniz had not left behind a body of work that would immediately bring with it a well-defined legacy for the eighteenth century.

As Germany rose in social, political, and economical prominence throughout the eighteenth century, the attention of German intellectuals shifted away from the rest of Europe and became much more internally focused, a development in which Wolff played a major role. Wolff shaped how both teaching and research in philosophy would be conducted at the university by publishing, from 1713 to 1725, a systematic and comprehensive series of seven philosophy textbooks in German (on logic, metaphysics, ethics, political science, cosmology or physics, natural theology

³⁸ The *Theodicy*, despite its length, is directed at a fairly narrow topic and would not have served well as a systematic introduction to his philosophy.

³⁹ Due to these limitations, the characterization of Leibniz’s account of pre-established harmony in the main text above focused on these works, to the exclusion of more informative texts that are now available.

or teleology, and biology) that would become exemplars for textbooks over the next several decades.⁴⁰

Thus, in metaphysics proper, which was considered to be foundational to all other disciplines, Wolff published *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt* (*Rational Thoughts on God, the World and the Soul of Human Beings, and All Things in General*) in 1720, which defined the structure of metaphysics for much of the rest of the century by systematically dividing it into four branches: ontology, psychology, cosmology, and theology.⁴¹ Ontology, as the science of being in general, not only clarified various primitive philosophical concepts (such as possibility, identity, necessity, contingency, order, quantity, quality, and composition), but also demonstrated the principles of contradiction and sufficient reason. Psychology, cosmology, and theology built on these principles by first providing a definition of one particular kind of being, whether it be of the soul, the world, or God, and then applying the fundamental principles of ontology to it. To extend his sphere of influence even more, Wolff reworked these textbooks into longer and more detailed Latin versions from 1728 to 1755, devoting separate volumes, for example, to rational and empirical psychology (which he had discussed in separate chapters in *Rational Thoughts*). Wolff's clear (albeit dull) prose, the comprehensive systematicity of his work, his importance at the university (both in Halle and throughout Germany), and his engaging teaching style not only contributed to the formidable influence he would have for the rest of the century, but also provided a framework within which scholars could easily locate and define their own interests.

It is also important to see that this shift in focus was not limited to philosophy and that it would be misleading to give the impression that Wolff's dominance was unopposed or even easily come by. In fact, much of the interest in philosophy was generated by a political struggle for intellectual domination at the university, as the university gained in importance for educating the leading social, political, and economic figures of the next generation. Specifically, in addition to conservative

⁴⁰ Beck's extensive presentation of earlier developments of philosophy in Germany (*Early German Philosophy*, e.g., pp. 184–189) shows that Wolff's aims were not unique, though they did attain a scope and diversity in content that were perhaps unparalleled.

⁴¹ The structure of Kant's *Transcendental Dialectic* in the first *Critique* notably follows Wolff's division. Christian Wolff, *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*, reprinted in Abt. 1, Bd. 2 of Christian Wolff's *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1983).

orthodox thinkers, two groups came to prominence: *Aufklärer*, that is, proponents of the Enlightenment, and the Pietists. Especially in Prussia (and Halle in particular), the Pietists had been quite successful in instituting educational reforms and establishing their advocates in positions of political importance (including positions within the university). It thus comes as no surprise that Wolff, as a leading Enlightenment figure, would come into conflict with the Pietists as his popularity grew. In 1717 the Pietists secretly sent students to Wolff's lectures in Halle to determine whether anything he said was hostile to Pietism and, if so, whether it could be used to make a case against him (as they had successfully done against Thomasius a few years earlier, restricting the range of the latter's lecture activities to topics in law). Although the material they gathered was apparently not enough to make Wolff alter his lectures, he was called in to meet with university officials and forced to apologize.

Not long thereafter, Wolff retaliated. In July of 1721 Wolff's duties as vice chancellor came to an end, and, as was standard procedure for such an occasion, he delivered a ceremonial address to celebrate the passing of the torch to his successor. The speech Wolff decided to give was entitled *Oratio de Sinarum philosophia practica*, on the practical philosophy of the Chinese, and the point of his speech was far from arbitrary. He argued that Confucius's practical philosophy was essentially based on reason alone and was thus worthy of praise and admiration. This was clearly an affront to the Pietists and must have outraged them.⁴² Since they stressed a person's individual suprarational relation to God and the impotence of reason concerning matters of spiritual weight, they could only interpret Wolff's address as it was intended, namely as a faintly disguised attack on their position. Further, the speech was not timed to please, inasmuch as the person who was to take over the vice chancellorship from Wolff was none other than one of the foremost Pietist leaders and the head of the theology faculty, Joachim Lange. Lange, furious about the content of the speech, demanded the right to censor Wolff's talk before its publication. Wolff refused on the grounds that the theology faculty was not competent to judge a philosophical work and thus had no authority

⁴² Georg Volckmar Hartmann notes: "So wurde, nachdem der Hr. R. R. Wolff am 12. Julii 1721 in einer öffentlichen Rede die Gleichheit seiner Lehr-Sätze mit der Sinesischen Philosophie gezeigt hatte, des Tages darauf so gleich öffentlich wieder ihn gepredigt," in *Anleitung zur Historie der Leibnizisch-Wolffischen Philosophie*, reprinted in Abt. 3, Bd. 4 of Christian Wolff's *Gesammelte Werke* (1737; rpt. Hildesheim: Georg Olms Verlag, 1973), p. 849.

over the philosophy faculty.⁴³ Needless to say, Wolff's actions did nothing to improve a situation that was already precarious.⁴⁴

For all practical purposes, at this point war between the Pietists and Wolff and his followers had been publicly declared, and heavy publication battles ensued over the next several years.⁴⁵ The Pietists repeatedly accused Wolff of determinism, fatalism, atheism, and Spinozism, claims against which Wolff repeatedly defended himself at length. What is of particular note in this dispute is that the Pietists based many of

⁴³ The story here is even more complicated. Due to the question of censorship, the publication of the speech was not a straightforward matter. It was first published without Wolff's knowledge by Jesuits in Rome, and later in different versions by Wolff and by Lange. For the complete history of the relationship between Lange and Wolff regarding the publication of this speech, see Michael Albrecht, "Editionsgeschichte," in Christian Wolff's *Rede über die praktische Philosophie der Chinesen* (Hamburg: Felix Meiner, 1985), pp. xc–ci.

⁴⁴ In "Freiheit gegen Fatalismus: Zu Joachim Langes Kritik an Wolff," in *Zentren der Aufklärung I, Halle: Aufklärung und Pietismus*, ed. N. Hinske (Heidelberg: Lambert Schneider Verlag, 1989), Bruno Bianco notes (p. 112) that Wolff's own behavior gave him no right to be described as a hero or a martyr, as Zeller does in his history of philosophy (Eduard Zeller, *Geschichte der deutschen Philosophie seit Leibniz* (München, 1873)).

⁴⁵ Christian Wolff published *Sicheres Mittel wieder unbegründete Verleumdungen, wie denselben am besten abzuwehren*, reprinted in Abt. 1, Bd. 21, 4 of Christian Wolff's *Gesammelte Werke* (1723; rpt. Hildesheim: G. Olms Verlag, 1981); *De differentia nexu reru sapientie et fatalis necessitatis, nec non systematis harmoniae praestabilitae et hypothesium Spinozae luculenta commentatio*, reprinted in Abt. 2, Bd. 9 of Christian Wolff's *Gesammelte Werke* (1723; rpt. Hildesheim: G. Olms Verlag, 1983); *Monitum ad commentationem luculentam de differentia nexu reru sapientie et fatalis necessitatis quo nonnulla sublimia metaphysica ac theologiae naturalis capita illustratur, auctore Christiano Wolffo*, reprinted in Abt. 2, Bd. 9 of *ibid.*, *Gründliche Antwort*, reprinted in Abt. 3, Bd. 23 of Christian Wolff's *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1986); *Anmerkungen über das Buddeische Bedenken von der Wolffischen Philosophie*, reprinted in Abt. 1, Bd. 17 of Christian Wolff's *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1980); *Nöthige Zugabe zu den Anmerkungen über Herrn D. Buddeus Bedenken von der Wolffischen Philosophie*, reprinted in Abt. 1, Bd. 21, 4 of Christian Wolff's *Gesammelte Werke* (1724; Hildesheim: G. Olms Verlag, 1981); *Der vernünftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen Überhaupt, Anderer Theil, Bestehend in Ausführlichen Anmerkungen*, reprinted in Abt. 1, Bd. 3 of Christian Wolff's *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1983); *Klarer Beweis, daß der Herr D. Budde die von ihm gemachten Vorwürffe einräumen und gestehen muß, er habe aus Übereilung die ungegründete Auflagen der Hällischen Widersacher recht gesprochen*, reprinted in Abt. 1, Bd. 19 of Christian Wolff's *Gesammelte Werke* (1725; rpt. Hildesheim: G. Olms Verlag, 1980).

Lange, alone, published *Causa Dei et religionis naturalis adversus atheismus* (Halle, 1723); *Modesta disquisitio novi philosophiae systematis de Deo, mundo et homine et praesertim de harmonia commercii inter animam et corpus praestabilita*, reprinted in Abt. 3, Bd. 23 of Christian Wolff's *Gesammelte Werke* (1723; rpt. Hildesheim: G. Olms Verlag, 1986); *Placidae vindiciae modestae disquisitionis de systemate philosophiae novo*, reprinted in Abt. 3, Bd. 23 of *ibid.*; *Anmerkungen über des Herrn Hoff-Raths und Professor Christian Wolffens*

their accusations on references to Leibniz's doctrine of pre-established harmony, presumably because that doctrine was distinctive and might have seemed most susceptible to attack. On November 8, 1723, a decisive blow was struck. Apparently, members of Frederick William I's "tobacco cabinet" – a group of military men with whom the king liked to smoke cigars – told him that according to pre-established harmony, a deserter from the army could not be held responsible for his action (given that it was "pre-established"), an idea with which they knew the king would have no sympathy.⁴⁶ Accordingly, at a certain level of appearances, it was due to his adherence to pre-established harmony that Wolff was removed from his professorship at Halle by the king and ordered to leave Prussia within forty-eight hours or be executed. Similar acts were undertaken against other Wolffians, such as Thümmig in Halle and Fisher in Königsberg, and it was officially forbidden to teach Wolff's textbooks in Prussia. Wolff was able to accept an offer at the university in Marburg, where he stayed until 1740. In 1733, the king saw that he had perhaps been too harsh on Wolff and invited him to return, with Wolff respectfully declining. Wolff would go back to Halle only after Frederick the Great, "the philosopher king," came to power in 1740, and even then Wolff declined the king's offer to be co-president of the Academy of Sciences in Berlin, since the person with whom he would have shared the presidency was Maupertuis, a staunch empiricist who openly expressed his hostility to Leibnizian-Wolffian rationalism in general and the monadology in particular.

The controversy surrounding Wolff should not hide the fact, however, that for many German thinkers in the 1720s and 1730s, his philosophy came to be accepted more than any other, and his followers were more likely than others to be pursuing the most novel lines of

Metaphysicam von denen darinnen befindlichen so genannten der Natürlichen und geoffenbarten Religion und Moralität entgegen stehenden Lehren, reprinted in Abt. 3, Bd. 23 of Christian Wolff's *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1986); *Bescheidene und ausführliche Entdeckung der falschen und schädlichen Philosophie in dem Wolffianischen Systemate Metaphysico von Gott, der Welt und dem Menschen* (Halle, 1724); *Ausführliche Recension der wider die Wolfianische Metaphysic auf 9. Universitäten und anderwärtig edirten sämtlichen 26. Schriften* (Halle, 1725); and *Nova Anatome* (Frankfurt, 1726).

⁴⁶ For the full story, see Wolff's *Eigene Lebensbeschreibung in Ausführliche Nachricht von seinen eigenen Schriften, die er in deutscher Sprache herausgegeben*, reprint of the 2nd ed. in Abt. 1, Bd. 9 of Christian Wolff's *Gesammelte Werke* (1726; rpt. Hildesheim: G. Olms Verlag, 1973), p. 28; Max Wundt, *Die deutsche Schulphilosophie im Zeitalter der Aufklärung* (Tübingen: Mohr Verlag, 1945), pp. 230–264; Eduard Zeller, "Wolffs Vertreibung aus Halle; der Kampf des Pietismus mit der Philosophie," *Preussische Jahrbücher* 10 (1862): 47–72.

research. His main metaphysics textbook, *Rational Thoughts on God, the World and the Soul of Human Beings, and All Things in General*, went through twelve editions between its initial publication in 1720 and 1752. During this period his most important followers (construed broadly) included Friedrich Christian Baumeister in Wittenberg; Alexander Baumgarten (whose work in aesthetics was extremely influential) in Frankfurt on the Oder; Georg Bernhard Bilfinger (who first became an important figure at the Academy of Sciences in St. Petersburg) in Tübingen; Christian Gabriel Fisher, Friedrich Albert Schultz, and Martin Knutzen in Königsberg; Johann Christoph Gottsched (whose elegant prose made his *Erste Gründe der Weltweisheit* rival Wolff's *Rational Thoughts* in popularity) in Leipzig; Johann Peter Reusch and Joachim Georg Darjes in Jena; and Georg Friedrich Meier and Ludwig Philipp Thümmig in Halle and then the latter in Kassel after his expulsion along with Wolff.

Pre-established Harmony in Wolff

If Wolff's textbooks defined the systematic framework in which particular philosophical issues were addressed in the first half of the eighteenth century in Germany and if his adherence to pre-established harmony was, at least nominally, at the center of heated philosophical (theological and political) debate, it is important to see how the general intellectual shift just described affected his attitude toward the issue of causality, in particular, toward pre-established harmony. The most significant point to note is that Wolff took pre-established harmony to be a doctrine whose systematic home lies in psychology. Although Leibniz clearly thought of the issue in its most general terms, which then had implications for a variety of particular instances, for Wolff pre-established harmony is nothing more than a particular solution to the mind-body problem and thus a doctrine whose relevance is limited to the restricted domain of psychology. Moreover, Wolff is quite clear that questions about the presence or absence of causal interaction between mind and body cannot be solved in *empirical* psychology, since he explicitly and repeatedly argues (in *Rational Thoughts* §§529, 534, 536, 761) – prior to Hume's *Treatise* and first *Enquiry* – that we directly perceive not the presence or absence of causal relations, but rather only that two things (or changes) coincide.⁴⁷ Thus the question of

⁴⁷ This is to suggest not that Wolff was the first to notice this point, but only that the point was not specific to Hume.

whether pre-established harmony, occasionalism, or physical influx is to be accepted as an explanation of the mind-body relationship is one that must be addressed in rational psychology.

This restriction in the scope of pre-established harmony explains not only why Wolff's objections to physical influx and occasionalism are fewer in number and less well developed than Leibniz's, but also why he thinks that he can be agnostic about the precise nature of the forces of those simples that compose or result in bodies.⁴⁸ In the chapter of *Rational Thoughts* that treats of ontological issues, after first showing (in §§51 and 76) that the objects of our perception are composites that must ultimately consist of simples, Wolff argues (§125) that every simple element must have a basic force that is responsible for the changes that it undergoes. In the chapter on cosmology, he argues (§594) that since everything has its sufficient reason, there must be a reason why each simple is related to every other simple in the way that it is, concluding that the reason for each such relational property must lie in the inner state of each simple thing. In light of this universal interrelation between simples via their inner states, Wolff (§§595–596) can account for the perfection of the world that is composed of these simple things. This line of argument is fully consistent with Leibniz's expression thesis (which states that every monad represents every other monad).

At this point, however, Wolff's argument takes what might otherwise appear to be a curious turn. For despite his agreement with Leibniz about the principle of sufficient reason, the necessity of simples (given our perception of composites), their complete reciprocal interrelation, and the ground of that interrelation in the inner states or forces of

⁴⁸ In §762 of the chapter on rational psychology from *Vernünfftige Gedanken*, Wolff argues that one cannot understand how causal relations could obtain between the mind and the body on the basis of the concepts of them that we have. Yet he does not think that this is sufficient to reject physical influx. For he acknowledges that we may not be able to comprehend everything that exists. Thus, he feels the need to investigate further and finds that accepting "natural influence" or physical influx would violate the law of the conservation of motion. In this section Wolff adds the idea that it would be implausible to claim that a mental force could transform itself into a motive force if the force of the mind "went" into the body (and vice versa). In §764 Wolff objects to occasionalism that it does not allow for a sufficiently robust distinction between creatures and God, it is inconsistent with the nature of simple beings (as enduring), and it requires perpetual miracles. In this section, Wolff also explicitly considers Descartes's distinction between the velocity and the direction of motion, noting that Huygens has shown that both motive force and direction are conserved. In §§766–767 Wolff argues that pre-established harmony is possible, explains how it is possible, and then adds, just as Leibniz does in the "New System of Nature," that it provides a proof of the existence of God.

these simples, Wolff does not require that the force of all simples be representational:

Now that I have distinctly established that the internal state of every simple thing refers to all the rest that exist in the world (§596) and Herr von Leibniz explains this in such a way that the whole world is represented in each simple thing according to the point where it is (§599), one can also understand further how everything in the world down to the smallest thing harmonizes with every other according to his opinion, and accordingly what he advances with his universal harmony of things, which, like all the rest that he has presented in this regard, appears to many as a puzzle that they believe to be unsolvable, since he has neither explained nor proved it sufficiently. However, because at the present time we do not want to decide what it really means for the inner state of simple things to refer to everything in the world, we shall let it remain undecided for the present in what the universal harmony of things consists, and it is enough for us that we have shown that it is present and that it can be explained in an intelligible manner according to the sense of Herr von Leibniz.⁴⁹

While Wolff is not widely known for his subtlety of thought, in this case he is suggesting not that he rejects Leibniz's idealism, but merely that he sees no ground for committing himself to it at this point.⁵⁰

In fact, Wolff's hesitation on this point is not completely unwarranted. For even if Leibniz does establish both the necessity of principles of unity (in the guise of simples) and the fact that my mind displays the kind of unity that would be required by (the multitudes of) my body, it does not follow that my mind (or even anything mental) is in fact what provides the requisite unity in all cases. Leibniz implicitly recognizes this point when he characterizes the principle of unity in numerous ways (formal atoms, substantial forms, metaphysical points) that do not have mentalistic connotations. As long as Leibniz does not present an argument for this further claim, it is not unwise to reserve judgment on this point.

In his so-called *Anmerkungen zur Deutschen Metaphysik*, which contains comments on some of the most important and controversial paragraphs of his *Rational Thoughts*, Wolff, rather than changing his position on this point, explicitly remarks that he had asked Leibniz for a proof of monads, but that Leibniz, though claiming in response that he could provide a proper demonstration, never actually did so. Wolff then comments that

⁴⁹ Christian Wolff, *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*, reprint of the 11th ed. in Abt. 1, Bd. 2 of Christian Wolff's *Gesammelte Werke* (1720; rpt. Hildesheim: G. Olms Verlag, 1983), pp. 370–371, §600.

⁵⁰ See Martin Schönfeld, *The Philosophy of the Young Kant: The Precritical Project* (New York: Oxford University Press, 2000), pp. 142–143, for a slightly different interpretation of Wolff's position on this point.

while he accepts that Leibniz has demonstratively proved the existence of simple things whose forces are responsible for their own changes, he still does “not yet see the necessity of why all simple things must have one and the same kind of force and suspects rather that a force would have to be found in the elements of corporeal things from which the forces of bodies . . . can be derived in an intelligible manner.”⁵¹ Wolff presents no considerations that would determine what kind of force that might be, whether mental or not, though he may be concerned that any connection between mental forces of representation and physical properties of bodies might not be sufficiently intelligible to count as “derivation.” As a result of Wolff’s agnosticism about the precise nature of the forces of simple elements, he must also be agnostic both about idealism (since the forces could be nonmental) and perhaps even about causal interactions between the simples (since the forces could be either immanent or transeunt).⁵²

Wolff’s point here is completely consistent with remarks he makes in other passages in the *Rational Thoughts*. In the first-edition preface, Wolff explains that initially he wanted to leave the question of pre-established harmony open.⁵³ However, he notes that when he came to discuss the mind-body problem in the chapter on rational psychology (which follows the chapters on ontology and cosmology), he found that contrary to his initial expectations, he had been led (§765) completely naturally to pre-established harmony by the principles adopted in earlier chapters. His argument runs as follows. By their very nature, bodies can neither think (§738) nor receive the power of thought (§739). Accordingly, the soul, which thinks, can be neither bodily nor composed of bodies (§742) and must therefore be a simple, self-subsistent thing (§743). But since, as he had argued earlier in the *Rational Thoughts*, every simple thing must have one and only one force that is directed to everything else in the world, the soul’s representational force must bring about, as its effects, not only its representations of the entire world (§§744–745), but also

⁵¹ Christian Wolff, *Der vernünftigen Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt, anderer Theil, bestehend in ausführlichen Anmerkungen*, reprinted in Abt. 1, Bd. 3 of Christian Wolff’s *Gesammelte Werke* (1724; rpt. Hildesheim: G. Olms Verlag, 1983), §215, p. 369.

⁵² Gerd Fabian, *Beitrag zur Geschichte des Leib-Seele-Problems: Lehre von der prästabilierten Harmonie vom psychophysischen Parallelismus in der Leibniz-Wolffschen Schule* (Langensalza, 1925), p. 40, interprets Wolff’s statement as entailing both dualism and physical influx.

⁵³ Wolff, *Vernünftige Gedancken*: “Ich hatte mir zwar anfangs vorgenommen die Frage von der Gemeinschaft des Leibes mit der Seele und der Seele mit dem Leibe ganz unentschieden zu lassen” (p. vii).

any other changes of its state (§754), just as Leibniz's pre-established harmony dictates. Thus, Wolff clearly accepts pre-established harmony for the mind-body problem as a doctrine that follows from other, more central principles. At the same time, as time goes by, Wolff de-emphasizes the view, repeatedly suggesting that "not much rests on finding this system [pre-established harmony] more probable than the others."⁵⁴ This trend is exactly what one would expect if pre-established harmony were supposed to be a solution to the mind-body problem and nothing more.

In light of this understanding of Wolff's position on pre-established harmony, we can now return to the issue of Wolff's relationship to Leibniz, even if only with respect to issues of metaphysics and its relationship to physics.⁵⁵ In one sense the ways in which Wolff departs from Leibniz's position might seem to be quite minimal. After all, merely being agnostic about idealism rather than being completely committed to it and therefore accepting pre-established harmony only as a solution to the mind-body problem (a solution Leibniz accepted as well) might not seem to amount to very much if one takes into consideration their vast agreement on a host of controversial metaphysical positions. Leibniz and Wolff both emphasize the principles of contradiction and sufficient reason as ultimate principles as well as the complete concept theory of substance and the principle of the identity of indiscernibles; they both argue that bodies are composite entities that owe their reality to simple substances; they both think that simple substances are composed of primitive active and passive forces and that the physical properties of bodies must be explained by means of derivative active and passive forces; they both accept the relational character of space and time as well as the conservation of living forces. In light of this extensive agreement on position, there is ample justification for thinking of Wolff as a Leibnizian or as representing a distinctively Leibnizian point of view in the eighteenth century, even if his access to Leibniz's more detailed "private" views as revealed in correspondence with leading European philosophers in the seventeenth century is limited.

⁵⁴ Wolff, *Der vernünftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen Überhaupt, Anderer Theil, Bestehend in Ausführlichen Anmerkungen*: "[G]ar nichts daran gelegen ist, daß man dieses Systemata für wahrscheinlicher als ein anderes hält" (487).

⁵⁵ Wolff's relationship to Leibniz on other issues may be quite different. For an account of the reception of Leibniz's theodicy in eighteenth-century Germany, for example, see Stefan Lorenz, *De Mundo Optimo: Studien zu Leibniz' Theodizee und ihrer Rezeption in Deutschland (1710–1791)*, *Studia Leibnitiana Supplementa* 31 (Stuttgart: Franz Steiner Verlag, 1997).

At the same time, there is another sense in which Wolff's agnosticism about idealism at the level of simple substances, along with the fact that some of Leibniz's deepest views that might provide a fuller understanding of the relationship between monads and the properties of bodies were not publicly available in the eighteenth century, changes everything. As we see in the rest of this chapter and in sections of Chapter 2, if a Leibnizian is not committed to idealism and does not think that physical properties must be reduced to (or derivative from) mental properties, then nothing stands in the way of developing what might be called a physical monadology. According to such a view, the simple substances that compose extended bodies are not endowed exclusively with mental properties, but rather are physical points. Such a view makes it much more tempting to think that the physical forces ascribed to bodies can be equated to (or at least associated much more closely with) those of the physical points that compose or result in bodies. Moreover, if one accepts a physical monadology and attempts to explain in a detailed and specific way how physical forces depend on the primitive forces of physical points, then it becomes far from clear that one must be committed to pre-established harmony between such physical points (though, as we see below, it is also far from clear that one must be committed to rejecting pre-established harmony). Along with the general intellectual shift described above, it is this change in philosophical position that gives rise to the most interesting debates about pre-established harmony throughout much of the eighteenth century in Germany.

KNUTZEN'S LEIBNIZIAN ARGUMENTS FOR PHYSICAL INFLUX

While the vitriolic debate that ensued throughout the 1720s and into the 1730s between Wolff and various Pietists was extremely influential during the period, for our purposes much of it can be discussed very briefly, in order to be able to consider at greater length the philosophically significant discussions of causality by Knutzen, Baumgarten, Meier, and Crusius in the rest of this chapter.⁵⁶ For it is only in these later figures that we find (1) a sophisticated defense of physical influx that is based on Leibnizian

⁵⁶ For a detailed discussion of the reception of pre-established harmony in eighteenth-century Germany that gives a fuller sense of what is philosophically relevant during this time period, see my "From Pre-established Harmony to Physical Influx: Leibniz's Reception in Early 18th Century Germany," *Perspectives on Science* 6 (1998): 136–203, special issue: "Leibniz and the Sciences," ed. D. Garber; Fabian, *Beitrag zur Geschichte des Leib-Seele-Problems*; Giorgio Tonelli, "Die Anfänge von Kants Kritik der Kausalbeziehungen und ihre Voraussetzungen im 18. Jahrhundert," *Kant-Studien* 57 (1966): 417–460; and

principles and distinctions (Knutzen), (2) an attempt to revitalize pre-established harmony on the basis of systematic and specific considerations pertaining to cosmological issues (Baumgarten and Meier), and (3) an account of causal powers that forms the basis for elaborate treatments of traditional issues in cosmology, such as the definition of the world, and psychology, such as the mind-body problem (Crusius).

From 1724 to 1730, Wolff and his Pietist opponents were primarily engaged in lengthy diatribes that, in the end, amounted to little of lasting philosophical interest. The Pietists (such as Lange, Johann Franz Budde, and Johann Georg Walch) continued to look for the smoking gun that would publicly prove Wolff's intellectual and moral corruption, while Wolff indefatigably reiterated his position in the hopes that he could clear his name from Pietistic calumny. Since many, though by no means all, of the Pietists' objections to Wolff's position were based on (more often than not intentional) misrepresentations of Wolff's position to which Wolff was both able and willing to respond on his own, little remained for Wolff's orthodox supporters to do.⁵⁷ As a result, his closest first-generational followers, such as Thümmig, Bilfinger, and Baumeister, published textbooks that rarely departed very much from Wolff's, though Bilfinger, clearly the best philosopher of the three, also produced, among other things, a lengthy and fairly sophisticated defense of pre-established harmony that drew heavily on Leibniz.⁵⁸ At the same time, several Wolffians did undertake a genuinely objective attempt at clarifying the underlying

Max Wundt, *Die deutsche Schulphilosophie im Zeitalter der Aufklärung* (1945; rpt. Hildesheim: Georg Olms, 1992).

⁵⁷ One of the Pietists' central objections was to Leibniz's and Wolff's explanations of the kind of spontaneity that was supposed to be exercised in our freedom and it may not be clear that Leibniz and Wolff are beyond reproach on this point.

⁵⁸ Ludwig Philipp Thümmig, *Institutiones philosophiae Wolfianae*, reprinted in Abt. 3, Bd. 19 of Christian Wolff's *Gesammelte Werke* (1725–1726; rpt. Hildesheim: G. Olms Verlag, 1982); Friedrich Christian Baumeister, *Institutiones metaphysicae* (Wittenberg, 1743); and Georg Bernhard Bilfinger, *De Harmonia animi et corporis humani maxime praestabilita ex mente illustris Leibnitii, commentatio hypothetica*, reprinted in Abt. 3, Bd. 21 of Christian Wolff's *Gesammelte Werke* (1723; rpt. Hildesheim: G. Olms Verlag, 1984), and *Dilucidationes philosophicae de Deo, anima humana, mundo, et generalibus rerum affectionibus*, reprinted in Abt. 3, Bd. 18 of Christian Wolff's *Gesammelte Werke* (1725; rpt. Hildesheim: G. Olms Verlag, 1982).

For more on Bilfinger's general influence, see Heinz Liebing, *Zwischen Orthodoxie und Aufklärung* (Tübingen: Mohr, 1961). For discussion of Bilfinger's treatise on pre-established harmony, see Joachim Kinttrup, *Das Leib-Seele-Problem in Georg Bernhard Bilfingers Buch "De harmonia animi et coporis humani, maxime praestabilita, ex mente illustris Leibnitii, commentatio hypothetica (1723)" in der geschichtlichen und philosophischen Zusammenschau* (Münster: Münstersche Beiträge zur Geschichte und Theorie der Medizin, 1974).

issues. For example, Samuel Christian Hollmann, the first professor of philosophy in Göttingen, and Johann Christoph Gottsched, who is now better known for his work in aesthetics, devoted significant attention to a careful investigation of the merits of pre-established harmony.⁵⁹ Still, Hollmann did not go beyond showing that pre-established harmony had not been demonstratively established, while Gottsched's ambitions were limited to elucidating how Leibniz's and Wolff's criticisms of physical influx were not necessarily decisive. Thus neither Hollmann nor Gottsched undertook anything even remotely resembling a full-fledged defense of physical influx of the sort that would present a serious challenge to the orthodox Leibnizian position. Since the Pietists who first engaged Wolff in polemics were not primarily philosophers, they were not in a position to mount such a defense either even if their contributions were based on a coherent philosophical position.⁶⁰

A comprehensive defense of physical influx was, however, undertaken by one of Kant's teachers in Königsberg, Martin Knutzen (1713–1751). Like Friedrich Albert Schulz, who supported him (and Kant as well a few years later), Knutzen was a Pietist who hoped that his religious outlook could somehow be combined with, or articulated within the framework of, Wolffian philosophy. Knutzen became well known in the first half of the eighteenth century for his writings in metaphysics, theology, natural science, and mathematics, having published *Philosophischer Beweis von der Wahrheit der Christlichen Religion* (1740) (which went through five editions prior to 1763), *De immaterialitate animi* (1741) (which was also translated into German in 1744), *Vernünfftige Gedanken von den Cometen* (1744), and *Elementa philosophiae rationalis seu logicae* (1747).⁶¹ Yet for our purposes his most significant publication is his dissertation, which was published in 1735 as an independent monograph (titled *Commentatio*

⁵⁹ Samuel Christian Hollmann, *De Harmonia inter animam et corpus praestabilita* (Wittenberg, 1724), *Observationes elencticae in Controversia Wolffiana, disputatori cuidam Halensi, ad vindicandas suas de harmonia inter animam et corpus praestabilita habitas dissertationes* (Frankfurt, 1724), *Dissertatio epistolica ad virum doctissimum Dn. Georg. Bernhardum Büllfingerum* (n.p., 1726), and *Epistolae amoebae de harmonia praestabilita* (Frankfurt and Leipzig, 1728), and Johann Christoph Gottsched, *Vindiciarum systematis influxus physici* (Leipzig, 1727–1729) and *Erste Gründe der gesammten Weltweisheit*, reprinted in *Ausgewählte Werke*, vol. 5, part I (1733–1734; rpt. Berlin: de Gruyter, 1983).

⁶⁰ See Bianco, "Freiheit gegen Fatalismus: Zu Joachim Langes Kritik an Wolff."

⁶¹ In *Martin Knutzen und seine Zeit: Ein Beitrag zur Geschichte der Wolfischen Schule und insbesondere zur Entwicklungsgeschichte Kants* (Leipzig: Verlag von Leopold Voss, 1876), Benno Erdmann gives a helpful account of Knutzen and his historical setting, including an account of Knutzen's other works. Erdmann notes that Knutzen's *Philosophischer Beweis von der Wahrheit der Christlichen Religion* was even translated into Danish (p. 53).

philosophica de commercio mentis et corporis per influxum physicum explicando, ipsis illustris Leibnitii principiis superstructa) and then reissued in 1745 as *Systema causarum efficientium*. For it is in this work that Knutzen developed a detailed case for physical influx.

What makes Knutzen's *Systema causarum efficientium* especially significant is that its argument for physical influx is based on Leibnizian principles. Thus, unlike Lange, Budde, and Walch, Knutzen is first and foremost a philosopher in the Leibnizian-Wolffian camp (broadly construed), and, unlike Hollmann and Gottsched (who could also be considered Wolffians in a similarly broad sense), Knutzen is suggesting not that the arguments in favor of pre-established harmony are inconclusive, but rather that fundamental Leibnizian principles actually entail physical influx.

Knutzen sets up the issue in the first part of the *Systema causarum* with precise definitions of the three causal theories and a statement of his contention (already granted by Wolff) that direct experience can neither confirm nor refute any of the three causal theories. The second part, which contains Knutzen's extended argument for physical influx, begins by carefully articulating several Leibnizian definitions: human beings (§17), spirit or monad (§18), bodies or composites (§19), action (§21), force (§22), space, place, and position (§23), internal and total motion, and motive, primitive, and derivative forces (§24).⁶² On the basis of these definitions, Knutzen then argues for what he takes to be the Leibnizian position that there must be simple elements or substances that compose bodies⁶³ and are in a place, though they do not fill a space.⁶⁴

⁶² Knutzen's presentation of the metaphysical framework that underlies his positive argument for physical influx does not suggest that his position is particularly close to Locke's (though he may be closer to Locke with respect to epistemological issues). See Manfred Kuehn, *Kant: A Biography* (New York: Cambridge University Press, 2001).

⁶³ Knutzen's argument is quite similar to Leibniz's in the *New System of Nature*. For Knutzen argues as follows:

§20. The existence of monads in bodies or the composition of bodies from monads is shown.

Bodies consist of simples or monads. Bodies are composite entities (§19), and therefore consist of parts (according to the same §). Either these parts are in turn composite, that is, they will have other parts that again have others and so on to infinity, or one must reach at last parts that do not consist of others. If the former, there is an infinite number of parts the existence of which implies a contradiction (as is demonstrated in the diss. *de Aeternitate Mundi impossibili* §21): Therefore, a body is composed of parts that do not have other parts. Therefore, it consists of simples, or monads (§18). (*Systema causarum efficientium* (Leipzig: Langenhemium, 1745), pp. 76–77)

⁶⁴ *Systema*, §27. "It is demonstrated that simple elements are in a place and are moved, although they do not fill a space" (p. 88).

The Argument from *vis motrix*

These definitions and the ontological position that derives from them put Knutzen in a position to present his first argument for physical influx in §28:

§28. *A force [that something has] to move itself involves in reality a force of moving another as well.*

A force of moving that brings it about that any being changes its own proper place cannot be conceived without the force of moving other things that surround it, but rather it is necessary that after positing one the other is given at the same time. For a force of moving that brings it about that a being changes its place does not exist except as a *conatus* for changing its own place (§24), i.e., for occupying a place distinct from the one that it now occupies, yet one that is still continuous to it (§ cit.). But the other coexistent things that surround the movable thing on all sides occupy a place distinct from the place of the movable thing, since two distinct beings cannot be in one place at the same time (§23). Therefore, a being endowed with the force of moving itself strives to push other things away, if they resist. But if they are truly also supposed to yield spontaneously, still what is already participating in progressive motion exerts itself in the way that is required to complete the motion beyond itself or to push other things away, since resistance is only the occasional cause of motion and does not add anything to the intrinsic force. Therefore, a being that moves itself enjoys the effort of changing the place of coexistents or the force of moving other things (§24). Therefore, the force of moving itself cannot be conceived without the force of moving other things, but after the one has been posited, the other is posited at the same time.⁶⁵

This is a provocative argument. The basic idea is that if Leibnizians assume that a being possesses the force to change its place⁶⁶ and that the change of place of one being implies the change of place of another being, then the force a being has to change its own place implies, now *contra* Leibnizians, the force to change the place of another being.⁶⁷ In

⁶⁵ *Ibid.*, pp. 91–92.

⁶⁶ Note that Knutzen formulates this argument quite generally so as to pertain to any being of which a Leibnizian will claim that it can change its own place. Thus, the argument will certainly apply to bodies, to corporeal substances, and perhaps even to monads (since monads are in some sense in a place and responsible for the changes that occur in bodies, so that one could say that they change the place of their body). For a discussion of Leibniz's conception of substance in his middle and later periods, see Robert Sleigh, *Leibniz and Arnauld*, pp. 98–104; Christia Mercer and Robert Sleigh, "Metaphysics: The Early Period to the *Discourse on Metaphysics*"; and Donald Rutherford, "Metaphysics: The Late Period," the latter two articles in *The Cambridge Companion to Leibniz*, ed. N. Jolley (New York: Cambridge University Press, 1995).

⁶⁷ As the argument stands, it is incomplete, since Knutzen simply accepts without argument that every being is surrounded in all directions, that is, that there is no void. However, the

short, intrasubstantial causation involving motion entails intersubstantial causation. This argument can be formulated in two different versions. The first version would claim that insofar as one being moves itself, it *in fact* causes the motion of the other. The second version, to be discussed shortly, would claim that the *mere capacity* a being has to move itself implies that it *could* move the other, which is enough to contradict one version of pre-established harmony, since one might think that physical influx is not just incorrect as a matter of fact, but rather metaphysically impossible.

What might one find objectionable in the first version of this argument? What is unquestionably right in Knutzen's argument is that if there is no void, the fact that one being changes its place implies that another being must change its place too. However, one could object that it is illegitimate to infer from this first, uncontroversial fact to the claim that the motion of the second being must be caused by the first being, since according to pre-established harmony each being would be causally responsible for its own changes. Thus, the fact that the first being has the force to move itself implies only that *either* the first *or* the second being has the force to move the second being. All that the motion of the first being implies is that there be *some* cause of the motion of the second being. On what grounds does Knutzen infer that the first being must be its cause?⁶⁸

However, one who was interested in defending Knutzen's position could concede the force of this objection by granting that the second being has the force to move itself, and still argue that what causes the exercise of the force in the second being must be the force of the first being. In other words, even if one were to admit causal activity in the second being, one need not accept that this causal activity is *sufficient* for the second being's motion. For one could argue that if the first being had not moved itself into the place of the second being, the second being would not have moved itself either. If that counterfactual is correct, then it might seem to be entirely appropriate to say that the first being has the

argument would have appeared plausible to Leibniz because he accepts the principle of sufficient reason, which he then takes to exclude the possibility of a void. (See his *Primary Truths, Philosophical Essays*, p. 33). In fact, Leibniz has additional reasons for rejecting the void (see "Specimen Dynamicum," *Philosophical Essays*, p. 130).

⁶⁸ In fact, an anonymous reviewer of this work makes the same objection to Knutzen's argument. See *Zuverlässige Nachrichten vom dem gegenwärtigen Zustande, Veränderung und Wachstum der Wissenschaften*, Teil 73 (Leipzig, 1746), pp. 48–67, esp. pp. 50–53.

force to move the second being, even if it does so only by virtue of acting on, or stimulating the efficacy of, the latter's force.⁶⁹

One might object to Knutzen's reply by invoking Leibniz's "world apart" doctrine, which states that the appearances would be entirely the same as they are even if nothing other than God and I existed (as, so to speak, a world apart), and the plausibility of this doctrine might appear to be evidence in favor of the sufficiency of internal causes.⁷⁰ Regardless of whether Knutzen accepts the "world apart" doctrine, he could still claim that this doctrine cannot play the role that one would need it to. For the "world apart" doctrine states that if only God and one being (e.g., I) existed, the appearances would be no different from what they are now. However, such a claim is irrelevant to Knutzen's argument. For Knutzen is concerned with the relationship that exists between the beings that *in fact* exist rather than the counterfactual situation in which only two beings (one finite and one infinite) were to exist. That different causal relationships would hold in the counterfactual situation in which only God and one being existed would seem to be a natural possibility that

⁶⁹ Knutzen is quite explicit about this model. In §44, in responding to the charge that physical influx is merely the flowing out and metamorphosis of motion and ideas, Knutzen states:

While the body acts on the mind according to the system of physical influx, it does not pour ideas of external things into the mind, nor the force of representation; but rather it modifies only the force of the mind and its substance in such a grounded way that a representation is caused in the mind. But the mind, when it acts on the body, does not pour a moving force into it, but rather only modifies and directs with its actions those things to the extent that they are present in corporeal elements in such a way that finally motion is produced in the body. For ideas and the force of representation are either accidents or substances. If they are accidents, they cannot be poured into the mind by the body and they cannot be transferred into the mind by a certain local motion from the body. For accidents do not migrate from subject to subject (§. 791. Ontol.). But if you suppose that they are substances, similarly such a transition cannot be granted, because the mind is a simple substance (§. 18.), but such a first substance cannot be the receptacle of a number of other substances. Therefore, neither ideas nor the forces of representation can be poured from the body into the mind. However, because representations of external things appear in the mind through the action of the body (§. 40. not.), nothing remains other than that the body, while it is acting on the mind, modifies its force and substance in such a way that representations of external things in fact appear or are caused in the mind. For a similar reason it can be shown that no moving force can be transferred from the mind to the body, and so through the action of the mind only those forces that the moderns have shown to be present in the elements [of the body] (§. 196. Cosmol.) are modified and directed for a certain reason in such a way that determinate motion is finally produced in the body through the determination of these forces. (pp. 145–147)

⁷⁰ See, e.g., "New System of Nature," *Philosophical Essays*, p. 143.

one would need to exclude by argument.⁷¹ Thus, applying the “world apart” doctrine to our second being, to say that the second being *would be* sufficient for its own motions if nothing else existed is not to say that it *is* sufficient for its own motions, given that other beings do in fact exist. Unless such Leibnizians create a philosophical standoff with the faithful on their side and their enemies and defectors on the other, they would have to admit that the presence of other beings could be causally relevant.

Moreover, not only *could* other beings be causally relevant, but Knutzen has also given us a reason to think that they really are. For the motion of one being necessarily implies the motion of another and it is tempting to think that if something were to cause a motion that necessarily entailed another motion, the cause of the first motion would also be the cause of what it necessarily entails, and hence be the cause of the motion of the second being. It would thus follow that the force a being has to move itself implies the force to move another being. Since it is not possible that the second being not move, given the motion of the first being, it seems that there is a sufficiently strong relationship between the first and second beings that Knutzen is justified in calling the first being the (or at least a) cause of the motion of the second.

But Knutzen can also appeal to a Leibnizian definition of action. In §21 Knutzen states: “A being is said to *be acting* when it contains in itself the reason for the existence (or change) of a certain thing.”⁷² If this definition is accepted, then it seems correct to infer, as Knutzen does in this first argument, that the force a being has to move itself implies that it has the force to move another, since the change of the one would seem to contain in itself the reason (or at least part of the reason) for the other’s motion. A Leibnizian could simply reject this definition of action, but it is far from clear what definition could be of help to a Leibnizian that would not beg the question against Knutzen.⁷³

There is, however, a second reading of Knutzen’s argument, one somewhat weaker and less controversial, but still sufficiently strong to present difficulties for at least one version of pre-established harmony. If a being has the force to move itself, then, even if a being contiguous to it were to move of its own accord, the first being must nonetheless have the force to move it, *in case* it had not moved of its own accord. If the first being

⁷¹ In a sense, no relationship would hold at all between various substances, if one grants that there are no relationships between nonexistent substances.

⁷² §21, p. 78.

⁷³ For Leibniz’s own explanation of ideal action, see his *Theodicy* §66, pp. 158–159.

had the force to move itself only if the second being moved itself, then it is inappropriate to say that the first being really has the force to move itself, since it does not contain the sufficient conditions for its motion in itself. In other words, it must be at least possible that the first being move the second, even if God were to set the world up in such a way that this force need never be exercised. But if it must be possible that the first being move the second, then Knutzen can conclude that pre-established harmony is false insofar as it implies that intersubstantial causation is not even possible (e.g., on the grounds that it is inconceivable). Thus, even this second, weaker version of Knutzen's argument can be employed against pre-established harmony.

However, Knutzen's argument raises an important question about how one is to understand the relationship between monads (or the ultimate constituents of reality) and bodies. For a Leibnizian could attempt to dismiss Knutzen's entire line of argument as irrelevant by first allowing that any being having the power to move itself would have the power to move others as well, but then restricting the scope of this principle to *bodies*. In other words, a Leibnizian could agree that if a body can move itself, then it can move others as well, but then deny that *monads* have the force to move anything. Admitting causation among bodies is not tantamount to admitting *intersubstantial* causation, since bodies are merely well-founded phenomena and not real entities, properly speaking.

At the same time, it is unclear whether a Leibnizian can restrict the principle underlying the argument to bodies in this way. If the motion of a body is caused by a derivative active force, and a derivative force (whether active or passive) is simply a modification of a monad's primitive force, then it is not completely unmotivated to assert that, ultimately, the primitive force of a monad is the cause of the motion of its body. And if that is granted, then the argument applies to monads just as it does to bodies. Whether this should be granted, however, would seem to depend on how to understand the relationship between derivative physical forces and the primitive metaphysical forces that underlie them, a relationship that is simply not stated precisely enough in Leibniz's public assertions.

Moreover, the issue of the precise relationship between primitive and derivative forces is closely related to questions about how we are ultimately to understand monads. Knutzen claims that the simple elements that compose bodies (which he thinks of as corresponding to Leibnizian monads) are spatial to the extent that they are *in a place* even if they are *not extended* (and thus indivisible). That is, Knutzen's simple elements would seem to be not just metaphysical, but also physical points. And if simple substances are to be understood as physical points, then it, again, does

not seem possible to restrict the principle concerning motion to bodies alone, since, for the sake of their distinctness, one point must be able to push another point out of a place into which it is attempting to move.

This line of objection raises a delicate question for Leibniz (and not merely for our eighteenth-century Leibnizians). What exactly is a metaphysical point and does Leibniz have any reason for thinking that it could not be physical as well? It is striking in this context that Leibniz (at least in texts publicly available in early eighteenth-century Germany) never seems to consider the possibility that monads might be physical points in this sense. In the “New System of Nature,” he says that the substantial unities required for being (i.e., monads) can be neither mathematical points (since mathematical points are “merely modalities,” i.e., abstractions from reality rather than realities themselves) nor physical points, but are rather metaphysical points. But his “justification” of the claim in this passage seems to presuppose that physical points are extended organic beings rather than truly indivisible, physical points. For he notes: “But when corporeal substances are contracted, all their organs together constitute only a physical point relative to us. Thus physical points are indivisible only in appearance.”⁷⁴ In other words, he seems to think that physical points are organic beings and, “if contracted,” can appear to be indivisible, but are in fact divisible (as organic beings would be, since they are extended).⁷⁵

But if Leibniz has not in fact considered the possibility that monads are physical points so understood, then he has provided no reason to think that his metaphysical points could not be physical points (especially since he explicitly asserts that mathematical points are the points of view from which the metaphysical points perceive the universe). He may believe that metaphysical points have a different kind of unity, a unity that essentially pertains to consciousness rather than spatial indivisibility, but, as Wolff recognized, Leibniz had not demonstrated that simple substances must have representational forces, and he has also presented no argument against them having a physical unity essentially as well.

In fact, Leibniz’s situation is even more precarious if he asserts that every monad is *necessarily* associated with an organic body. If the necessity were to stem from the fact that my mental unity is possible only if I represent the world from a single, particular location in it, then it would seem

⁷⁴ *Philosophical Essays*, p. 142.

⁷⁵ It is likely that Leibniz is arguing against an Aristotelian position, which reinforces the idea that Leibniz is not, in fact, considering the possibility that bodies are ultimately simply aggregates of unextended physical points.

that Leibniz has even supplied the materials for an argument establishing that these two kinds of unity (physical and metaphysical) necessarily go together. If this is the case, Leibniz cannot simply dismiss Knutzen's argument as being based on a misunderstanding of his position, since it exploits a connection in his own philosophy. Moreover, if Leibniz does not address the issue explicitly, one of his most prominent defenders in the late 1730s, Alexander Baumgarten, does, stating as clearly as one could desire (in §399 of his *Metaphysica*): "If by a PHYSICAL POINT you mean an actual thing that is completely determined beyond its simplicity, then certain monads of this universe, namely those the aggregation of which are an extended thing, are physical points."⁷⁶

The Argument from Impenetrability

In §29 Knutzen provides a second argument for physical influx, which, though distinct from the first argument, clearly stems from similar considerations. He writes:

The same can also be demonstrated another way. Simple elements are impenetrable, according to the opinion of the illustrious Leibniz, who asserts that all finite substances are impenetrable. See his *Letter to Cl. Wagner p. 201. Tom. I. Epist. Edit. Kortholtianae.*⁷⁷ Hence, it cannot be the case that one [simple element or substance] is in the place of another. Therefore, there is something real, by

⁷⁶ Alexander Baumgarten, *Metaphysica*, reprinted in Immanuel Kant's *Gesammelte Schriften*, vols. 15,1 (pp. 5–54) and 17 (pp. 5–226) (1739; rpt. Berlin: de Gruyter, 1902–), 17:110.

⁷⁷ This letter, from Leibniz to Rudolph Christian Wagner, dated June 4, 1710, is reprinted in *Die Philosophischen Schriften*, ed. C. I. Gerhardt (Berlin, 1875–1890), vol. 7, pp. 528–532. Leibniz does not explicitly assert that monads are impenetrable, but he does make a series of remarks that could naturally be interpreted that way. For example, he says: "I respond, secondly, that the resistance of naked [or prime] matter is not an action, but merely a passion, as long, of course, as it has antitypy or impenetrability by which in fact it resists the thing about to penetrate, but it does not make that thing rebound unless an elastic force is added that must be derived from what is moved and therefore from an active force that is superadded to matter." While it is not absolutely necessary to assume that impenetrability is to be identified with the resistance of prime matter, one can certainly excuse Knutzen for reading Leibniz in this way. Moreover, later, Leibniz concludes: "Only God is a substance truly separate from matter, because he is pure act, is not at all endowed with the capability of being acted on, that is everywhere, and constitutes matter. And in fact all created substances have antitypy by means of which it naturally occurs that one is outside [*extra*] another, and therefore penetration is excluded." Knutzen could easily take the "one" referred to in the last sentence as referring to "created substances," which are implicitly characterized in the previous sentence merely as being capable of being acted on. In 1739, Alexander Baumgarten is as explicit as one might like (in §398 of his *Metaphysica*): "Therefore, all substances, hence also all monads, of this and of every composite world are impenetrable" (17:110).

whose force one simple excludes and pushes up against another, lest the other invade its place. Since it is most certain that simples are moved (§27) and that distinct simples are not moved according to an opposite line of direction, it is consequently impossible that they penetrate each other mutually, or rather what we may gather from the conflict of bodies and their collision is that in fact they are carried in a contrary direction mutually away from each other. It follows in this case that one must hold that either simples penetrate each other mutually, which goes against Leibniz's assertions, or if they resist each other mutually, they must act on each other mutually. Q.e.d.⁷⁸

As in the first argument, Knutzen tries to show that a property that a Leibnizian ascribes to finite substances implies intersubstantial causation. The basic idea behind this argument is that impenetrability is intelligible only if one substance is attempting to penetrate into the place occupied by a second substance, where the second substance is said to be impenetrable in virtue of resisting the first substance's efforts. But, the argument continues, how can one substance be said to *resist* another substance if not causally? That is, surely resistance is a causal term, and a substance cannot resist itself, so that if resistance (or impenetrability) is a real property of substances, then there must be interaction between substances. The case of impenetrability and resistance suggests causal interaction among substances even more than does the case of motion.⁷⁹

A Leibnizian would seem to have two lines of response to this kind of argument. First, it could be argued that one should divorce all causal connotations from the concepts of resistance and impenetrability. According to this line of response, a Leibnizian would have to claim that resistance and impenetrability ultimately amount to nothing more than the simple fact that substances cannot be in the same place at the same time. Yet Knutzen seems justified in pressing this point, since the metaphysical fact that substances cannot be in the same place seems quite distinct from the physical question of whether substances are impenetrable and resist each other. In particular, it would be natural to think that resistance is the causal *means* for keeping substances from occupying the same place at the same time. To put the problem with this line of response in words closer to Leibniz's own, according to pre-established harmony God is to have arranged the motions of bodies with such great

⁷⁸ §29, p. 95.

⁷⁹ Knutzen formulates this argument so as to apply to substances that have resistance or impenetrability. Thus, again, Knutzen's argument is so general that it applies to bodies, corporeal substances, and again, though this is more controversial, even (the later Leibniz's) monads (insofar as Leibniz holds that two monads cannot be in the same place).

harmony that they should have no need to resist each other in the first place. As a result, a Leibnizian is forced either to accept the (prima facie implausible) idea that resistance and impenetrability reduce to the fact that substances cannot occupy the same place at the same time or to deny that substances are truly impenetrable.⁸⁰

The second line of response open to a Leibnizian parallels the final response to Knutzen's argument from a substance's power to move itself, discussed above. That is, one could restrict the applicability of Knutzen's argument by asserting that only bodies, not monads, are impenetrable. In support of this restriction, one could point out that in, for example, the "Specimen Dynamicum," Leibniz says that impenetrability is caused by the derivative forces that pertain to bodies rather than by the primitive forces that make up simple substances. As he puts it, the primitive passive forces are "that *by virtue of which* it happens that a body cannot be penetrated by another body."⁸¹ Because this statement does not specify the means by virtue of which the primitive passive force makes a body impenetrable, one could read it as claiming that impenetrability depends either immediately on primitive forces (which is how Knutzen understands it) or on derivative forces that ultimately have their ground in primitive forces. However, given the paucity of detailed public statements about the primitive-derivative force distinction, it is understandable that Knutzen would infer that if derivative forces both explain impenetrability and truly derive (causally) from primitive forces, then impenetrability must ultimately be explicable (causally) by the primitive forces as well. While a Leibnizian would presumably reply that the fact that primitive forces ultimately explain impenetrability does not necessarily imply that the primitive forces act between rather than within a substance, one can understand why Knutzen might think that the causal relations between bodies expressed in the notion of impenetrability would have to be reflected in causal relations between the primitive forces of monads, especially if Leibniz's idealistic metaphysics has been replaced with one that accepts the possibility of physical points.

Knutzen can also point out that whether or not Leibniz thinks of the primitive passive forces of monads as the immediate cause of bodies' impenetrability, he does explicitly ascribe resistance to them, and resistance encounters difficulties that are similar to those faced by impenetrability. First, ascribing resistance to a monad strongly suggests that

⁸⁰ It is worth pointing out that in §39 Knutzen explicitly applies these first two arguments to the mind-body relationship.

⁸¹ *Philosophical Essays*, p. 120 (emphasis added).

it is causally efficacious and does not simply indicate the presence of imperfection in the monad. To put the point the other way around, if resistance ultimately means merely imperfection, then resistance is, at best, metaphorical, and, at worse, highly misleading as a description of the ultimate character of real things. Second, if resistance is to be understood causally, then the question arises as to how this causality should be understood. The most natural interpretation of the idea is to claim that one thing resists another. However, by rejecting causal interaction between substances, Leibniz is committed to the counterintuitive assertion that a monad can resist itself. If one could divide a monad into parts such that one part would be in a position to act on another part, perhaps the counterintuitive aspect of this claim could be dispelled. However, given Leibniz's insistence on the unity and simplicity of substance, that move is not likely to tempt him. Rather, he may ultimately be committed to the idea that a substance could be active and passive toward itself in one and the same respect, an idea that Knutzen could reasonably find unattractive.⁸²

The Argument from the Simplicity of (Divine) Action

Knutzen's first two arguments are supposed to apply to any beings that cause motion or are impenetrable, whether they are bodies or the simple elements that compose bodies. Accordingly, Knutzen starts his discussion by addressing the cosmological question of how the simple elements that compose bodies in the world relate to each other, a question that Wolff took no stand on, in order to then turn to the question Wolff does face, namely the "psychological" question of the mind-body problem. In §§30–32 Knutzen prepares to extend the scope of physical influx from beings *per se* or "simple elements" to the mind-body relationship by explaining what perception is (§30) and by arguing that simple elements perceive (§31).⁸³ In §§33–34, Knutzen then applies to the mind the arguments that he has already constructed for simples.

What is surprising, however, is the specific way in which Knutzen does so. For in §33 he turns to consider the proper definition and nature of absolute perfection. "*Perfection* that implies no limitation *per se* or, alternately,

⁸² Leibniz may find this idea more attractive by considering how it applies to the way in which we think of ourselves and our own mental lives.

⁸³ It is worth noting that Knutzen (like Leibniz before him and Crusius and Kant after him) considers the issue of causality both in its completely general form (i.e., for any substances whatsoever) and in the specific guise of the mind-body problem, whereas many parties to the dispute limit themselves to discussing only the mind-body problem. As a result, Knutzen's position attains a greater degree of sophistication.

that can exist together with any other possible entity (as the Scholastics say) is called *perfection absolutely* or *simpliciter*. However, it is demonstrated in natural theology that *anything truly ascribed to God is a perfection simpliciter, which does not contradict anything, except limitations or imperfections.*⁸⁴ In §34 Knutzen then uses this conception of perfection to establish that physical influx is possible for the mind-body relationship, since it involves no contradiction. His justification is as follows:

Physical influx of the mind on corporeal simples and of those simples in turn on the mind is completed by an act (§32). And so if the possibility of physical influx is to be demonstrated, it must be shown in what way the action of the mind outside itself on other simples does not involve a contradiction (§85 *ontol.*).⁸⁵ So let's investigate in particular whether one can discover anything in the mind that contradicts its actions on external things. We discover in the mind those things that primarily amount to this: That the mind is a simple being, moreover, that it is a perceptive being or is a perceiver, and in a far greater degree of eminence than that of simple elements, since it perceives distinctly or is provided with an intellect or a will (§§17–18). Action on external things cannot contradict the simplicity of the mind not only because God acts outside himself (according to the *Princ. of Nat. Theol.*), but also because the simple elements act on each other mutually (§29). Nor can this action contradict the mind to the extent that it is a being that perceives or is perceiving, because simple elements take pleasure in the perception of external things (§31); yet an action of this kind cannot be denied to these simples. Therefore, nothing remains other than that eminent perfection by which the human mind is separated from simples that have inferior perceptions, and that places the faculties of understanding and willing in the mind, and if the external action of the mind does not contradict this, then the possibility of physical influx will have been established beyond doubt. But action on external things is a simple perfection because it can be ascribed to God (§33) and therefore it cannot be inconsistent except with limitation and imperfection (§ cit.). Therefore, it cannot be inconsistent with intellect and volition, which exceed the mere faculty of perception and at the same time confer a greater perfection on the being, since, as was already shown, this kind of action is consistent with a merely perceptive and more perfect being. Therefore, who could doubt that the mind can act on the body? However, because it was demonstrated above (§29) that action on other simples must be attributed to the simples of which the body

⁸⁴ §33: "*Perfectio, quae nullam per se infert limitationem, seu, quae cum omnitudine possibilitium (ut loquuntur Scholastici) consistere potest, perfectio dicitur absolute talis seu simpliciter simplex.* In Theologia naturali autem demonstratur, *quaecunque in Deo T. O. M. locum inveniunt, perfectiones esse simpliciter simplices, nec, nisi cum limitationibus seu imperfectionibus, ullam inuoluere repugnantiam*" (p. 107). (T. O. M. is an abbreviation that stands for *ter optimus maximus*, or "thrice greatest most powerful." Since such an abbreviation is not currently common, I have omitted it in translation.) That Knutzen cites Bilfinger (and Canz) in the explanatory passage to this definition indicates that he does not take himself to be positing anything that would be controversial to Leibnizians.

⁸⁵ Knutzen is referring here to Wolff's *Philosophia Prime sive Ontologia*, which was originally published in 1729.

consists, and surely the mind, insofar as it is a simple substance, can act on such simples (according to the demonstr.), there will be no reason why anyone should judge the action of the body's simples on the mind to be impossible. Therefore, it is established that physical influx is possible.⁸⁶

The basic idea behind Knutzen's argument in this passage is that the mind's actions on the simples that constitute or result in bodies cannot be contradictory because it has already been established that both God and other simples can act in this way, and there is no reason in the features that characterize our mind that could generate a contradiction. The fact that minds are simple and have perceptions cannot create a problem, since God and other simples have these properties as well. Moreover, what distinguishes minds from simples, namely the fact that they are endowed with intellect and will, cannot be problematic either, since the intellect and will represent a greater perfection than that had by the simples that compose bodies.

Still, one might question whether this argument is really sound. After all, the mere fact that the simple elements that constitute bodies can act on each other does not immediately imply that other kinds of simple elements have the capability of acting on others (whether they be bodies or their simples). Moreover, the inference does not necessarily follow even if the latter are more perfect than the former. But notice the principle Knutzen explicitly endorses. He asserts that the ability to act on others is a perfection, since God can act on others (in creation and miracles) and anything that God does reflects his perfection. So the inference underlying Knutzen's argument is not that a superior being can do whatever an inferior being can do (which one could easily doubt), but rather that acting on others is a perfection, and a perfection can generate a contradiction only if limitations or imperfections come into conflict with that perfection. However, the intellect and will cannot be imperfections or limitations, since they are precisely what elevate minds over the simple elements that constitute bodies. Thus, the way in which Knutzen embeds the argument one would naturally expect into the context of the perfection of God and the imperfections of the simple elements represents an interesting turn of thought.

This argument also brings out a feature of Knutzen's version of physical influx that was only implicit above, namely that the force of acting on others is primitive or basic and is thus incapable of being explained in other, more basic terms. As we saw above, Leibniz's most fundamental objection to physical influx was that it could not explain how one substance

⁸⁶ §34, pp. 108–110.

could act on another, since the migration or transfer of accidents is ontologically implausible. We can now see Knutzen providing a response to this objection by arguing that one cannot go beyond saying that a substance has the capacity, force, or power to act on another substance. Any suggestion that such an action must be explained further in terms of accidents migrating from one substance to another (as Leibniz is wont to do, perhaps unfairly, on behalf of proponents of physical influx) is calling for an inappropriate explanation, namely an explanation of something that is already as simple as it gets.

Is it legitimate, one might ask, to accept action as a primitive in this way? In addition to the fact that Knutzen can appeal to the divine case of creation, which virtually everyone in this context would accept as an instance of one (infinite) substance acting on other (finite) substances, *Leibniz* can hardly object to such a primitive force or action because he himself assumes such a force when he claims that it is through a *conatus* or primitive active force of appetite that a substance strives to change from one state to the next. In other words, by accepting intrasubstantial causation a Leibnizian must be accepting some kind of capacity for a substance to act. So the very idea of such a force or action cannot be objectionable.

Yet there is a clear difference between Leibniz's and Knutzen's conceptions of force. On Leibniz's conception a substance can act only on itself by means of its force, whereas for Knutzen a substance is supposed to be able to act on other substances. Is this difference significant? It might seem to be. The point behind Leibniz's suggestion that one substance can act on another only if an accident were to migrate from the one substance to the other is that a property cannot just pop into existence but rather must come from some source (namely from the substance that is acting). If it comes from the substance that is acting, then for physical influx to occur it must somehow get from the one substance to the other. How could it do that if it did not migrate there?

It is far from clear, however, that this difference can provide any tangible advantage to the Leibnizian. First, this objection still does not clarify what the special difficulty is with one substance acting on another substance as compared with a substance acting on itself. For if it is legitimate to ask about the source or origin of a property, both accounts face the same dilemma. Either the property is already present in the substance, in which case no force would seem to be necessary in order to make it true of the substance, or it is not already present in the substance, and then creation *ex nihilo* will be equally problematic for both inter- and

intrasubstantial causation. If the problem of creation *ex nihilo* is solved by the concept of force invoked in pre-established harmony, then, by reason of parity, that concept should solve it for physical influx, too.⁸⁷

Second, even if there is some special problem with the “inter” part of intersubstantial causation, the Leibnizian has not countenanced Knutzen’s specific version of physical influx. As we saw above, two forces can be involved in the production of a property, one in the substance that is the cause and one in the substance that exemplifies the property being produced.⁸⁸ In the case of motion, if the second substance moves itself, it would be in virtue of the first substance “activating” or triggering its activity (e.g., in a collision). In this way the one substance can be responsible for the fact that a new property has been created, while the other can be responsible for that property being the specific one that it is (e.g., that it be a motion in a particular direction with a particular velocity rather than in some other direction with another velocity). In this way Knutzen could undercut the most obvious Leibnizian grounds for objecting to a substance’s simple force of acting on others.

The Argument from Probability

Knutzen’s final argument for physical influx stands one of Leibniz’s arguments for pre-established harmony on its head. After establishing in §§33–34 that physical influx is possible, in §35 Knutzen turns to show that it is more probable than the other two causal theories by considering various criteria of truth. First, it agrees with experience, a point Wolff explicitly admits. Second, as §34 has shown, it is metaphysically possible. Finally, and perhaps most importantly, Knutzen argues that it agrees with divine wisdom. As he puts it: “For it is established in natural theology that God, in conformity with his greatest and infinite wisdom, chooses the natural, shortest path.”⁸⁹ Knutzen proceeds to explain what the shortest path is, namely that “those things that can come about naturally through a select few will not be completed through many or by the longer path.”⁹⁰ In the explanatory text to this passage, Knutzen emphasizes that it is not

⁸⁷ Leibniz might think that the fact that every predicate is contained in the concept of a thing helps his case here.

⁸⁸ This feature of Knutzen’s account is explicitly stated in the *Systema causarum* at §43: “According to the system of physical influx the human mind is not in every respect nor specifically in thinking to be conceived of as purely passive” (p. 142) and §44.

⁸⁹ §35, p. 113.

⁹⁰ *Ibid.*, p. 113.

enough that God choose the shortest path (which occasionalism might embody), but it must also be the shortest *natural* path.⁹¹ In the next section (§36) Knutzen explicitly argues that pre-established harmony involves “roundabout ways because so many skills are necessary for producing this harmony both in the body and in the mind that they may surpass every form of comprehension.”⁹² In other words, he is objecting to the idea that the harmony in pre-established harmony is guaranteed by the extraordinary magnitude of God’s benevolence and wisdom rather than by readily comprehensible natural powers of finite substances alone, since pre-established harmony’s reliance on God’s benevolence and wisdom extends to resources that lie well beyond our understanding.

Knutzen’s point here is well taken. What he finds objectionable is not that physical influx can, whereas pre-established harmony cannot appeal to natures, since, as we saw above, Leibniz himself appeals to natures in distinguishing his views from Malebranche’s. Rather, Knutzen is objecting to the incredible complexity of the natures that Leibniz invokes. If the harmony that is to be found between substances that do not interact with each other is supposed to be so intricate and impressive as to prove the existence of an infinitely powerful and knowledgeable God, then it is clear that that harmony cannot be intelligible on the basis of simple, readily comprehensible general natures. Since there is no reason that physical influx could not invoke general natures that human beings could understand, physical influx can explain things in a shorter, but still entirely natural way.

Objections and Replies

In the third and final part of the *Systema causarum* Knutzen considers ten objections to physical influx. Of particular note are those that concern the one objection Leibniz had raised that Knutzen had not previously addressed in his arguments in the second part, namely that physical influx violates the law of the conservation of motion. To start off, Knutzen’s formulation of the objection differs from Leibniz’s in that Knutzen does not share Descartes’s view that motion is conserved in the world. Rather, he takes Leibniz’s side in the *vis viva* debate that what Knutzen refers to as living forces (*mv²*) are conserved. He then reformulates the objection accordingly so that it states that physical influx contradicts the law of the

⁹¹ Ibid., p. 116.

⁹² Ibid., p. 117.

conservation of living forces. For if a monad were to act on a body at rest so as to put it in motion, there would be an increase in the total amount of living forces in the world.

In his discussion of Objection 6 (§53), where Knutzen confronts this objection directly, he simply denies that the law of the conservation of living forces holds for mind-body interaction. His justification for this denial is twofold. He first notes that the law has been proven only for elastic bodies, not for inelastic bodies, much less for the mind and the body. This is part of what he means when he emphasizes: “I deny . . . that it follows from physical influx that a certain quantity of living forces is not conserved in the collision of bodies *among each other*.”⁹³ For “as long as it has not yet been shown and cannot be shown that this law of motion about conserving a certain quantity of living forces is not only dictated for bodies acting on each other mutually, but also for the mind acting on the body and vice versa, there is absolutely no objection present that violates physical influx.”⁹⁴ To support this contention, Knutzen provides an explicit reason why this law should not hold for the mind. Since Leibniz derives the law of the conservation of living forces from the law of inertia (“that any body remains in its state of rest or uniform motion in a direction unless it is forced to change its state by an extrinsic cause”) and “it is most evident that the mind does not at all remain in its state of rest and uniform motion in a direction until forced to change its state by an external [cause]” (i.e., the law of inertia does not hold for the mind), there is no reason to think that the conservation law ought to hold for the mind.⁹⁵

Two other objections are relevant in this context as well. In Objection 7 (§54) Knutzen considers Leibniz’s objection that physical influx cannot apply to the mind-body relationship because of a lack of proportionality, while in Objection 8 (§55) he considers the possibility that a problem might arise in applying to mind-body interaction the principle that the “full effect” must be equal to the “full cause.”⁹⁶ When discussing these

⁹³ §53, p. 177.

⁹⁴ *Ibid.*, p. 178.

⁹⁵ *Ibid.*, p. 182.

⁹⁶ §55. **Objection VIII. that the effect is greater than its cause according to physical influx.**

The whole effect is equal to the full cause. But in the system of physical influx what is considered to be the whole effect is not equal to the full cause. Therefore, what is considered to be the effect in the system of physical influx is not the effect. Therefore, there is no influx of the mind on the body nor vice versa. The learned Bilfinger built a proof of the minor premise on this foundation because in physical influx the effect is greater than its cause, because sc. not only is the motion of the

objections, Knutzen considers the relationship between bodies and minds in more detail and provides a fuller explanation of how to respond to the charge that physical influx violates the laws of nature. In §55 Knutzen writes:

One must note the following concerning physical influx: 1) that the impression of motion that is communicated to monads causes perceptions in them, for the communication of motion cannot be completed except through the modification of forces that are present in elements and that Leibniz calls perceptual. And thus beyond the communication of motion nothing else is required in order to cause perceptions; and as it is in the case of the monads of the body, so it is in the soul.⁹⁷

In this passage and others, Knutzen can be understood as using Leibniz's own views against himself in the following way.⁹⁸ As we saw above, Leibniz claimed that a monad is a simple substance that brings about its perceptions and appetitions according to inherent teleological principles, whereas bodies are (nothing more than) well-founded phenomena that are determined by efficient, mechanical causes. By asserting that completely different, albeit harmonious laws govern monads and bodies, Leibniz, in effect, drives a wedge between monads and bodies, a separation reinforced by his repeated talk of the "realm of efficient causality" and the "realm of final causality." While this separation may appear attractive as a way of reconciling the freedom of monads with the determinism of bodies, it carries a price. For Knutzen can use this separation to take the force out of Leibniz's objection that physical influx would violate the conservation of living forces. For Leibniz's objection to hold, one needs to be able to find a *specific* correlation between the action of a monad and the living force(s) of a body (or set of bodies) to be able to claim that the action of the one changes the state in the other in such a way that the living force is either increased or decreased over what it would have been otherwise. However, by separating the realms of final and efficient causality, one has established (or at least opened up the possibility) that there might be no specific correlations between monads and bodies of the sort that would be required for the objection to hold. To make the point more concretely, given that different laws govern monads and bodies, whether a monad acts in one way rather than another is determined by its own

fluid nerves in the brain caused by motion in the sense organs, and they in turn cause the motion of another that then releases the full effect, but besides this effect and the other there is sc. representation in the mind that therefore reflects an effect greater than the cause. (pp. 190–191)

⁹⁷ §55, p. 192.

⁹⁸ See also *ibid.*, p. 195.

laws and does not necessarily entail that the living forces of bodies are different in the one case from how they would have been in the other.

Now one might attempt to respond to this objection by arguing that it overlooks the fact that despite their distinctness, the laws of efficient causality and the laws of final causality are closely related, and if it can be shown that the one set of laws depends on the other set of laws, then it might be possible to establish the kind of specific correlations that the original objection presupposes. Moreover, one could appeal to a variety of dependency claims that are supposed to hold between (the laws of) monads and bodies. In the “Specimen Dynamicum” Leibniz claims both that the laws of mechanics derive from metaphysical laws (which govern monads) and that bodies’ derivative forces are derivative from the primitive forces of monads, while in the “New System of Nature” he argues that the divisibility of bodies qua extended beings requires the unity that monads possess and that monads “result” in bodies (i.e., are in some sense the nonextended components that constitute bodies so that bodies, as composite wholes, depend on their constituent components). Yet it is crucial to notice here that whatever dependency is asserted in order to establish a particular correlation between monads and bodies is irrelevant to the question of causality, that is, the truth of pre-established harmony or physical influx. For, in principle, the correlation could hold whether the one stands in a causal or a merely harmonious relationship to the other. The converse holds as well. Whether there is a harmonious or a causal relationship between monads and bodies does not necessarily immediately determine whether there will or will not be specific correlations between them.

Consider the same point from a slightly different angle. Leibniz’s (re-formulated) objection to physical influx states that if the mind were to cause a body to move in certain ways, this action would add to the total amount of living forces in the world, that is, living forces would not be conserved. The inference to be drawn from this objection is that one should deny that the mind acts on the body. However, as Leibniz’s own position might suggest, one can instead infer that one should distinguish between the realm of bodies and the realm of minds or monads, and claim that the latter founds the former only in some general way. Thus, even if the mind causes a living force in bodies, there is no reason to think that there must be more living forces afterward than beforehand because presumably this action is just part of a general or global founding relationship that exists between the two realms. As a result, the “new” living force would have occurred even if the mind had not chosen to act in the

way it did, and the reason the living force would have occurred is that it is simply obeying the laws that hold for its respective realm. So what is crucial to solving the problem posed by the law of the conservation of living forces is not denying causal interaction between mind and body (as Leibniz does), but rather restricting the law to bodies (as Knutzen explicitly does) and arguing that monads are subject to different laws, even if they “harmonize” with the laws that hold for bodies.

Accordingly, Knutzen has developed a subtle response to a powerful Leibnizian objection. He first notes that the law of the conservation of living forces would hold only for beings that can have living forces, namely bodies interacting with each other, not monads. Even more interestingly, he then argues that Leibniz’s separation of the realms of final and efficient causality drives a wedge between monads and bodies in a way that reveals that the original objection is actually based on being able to make specific correlations between monads and bodies rather than on the presence or absence of causal connections between them. In other words, reconciling the law of the conservation of living forces with the freedom of monads by first separating monads and bodies and then asserting that different laws hold for each type of entity is a move that is just as open to Knutzen as it is to the Leibnizian, since any dependency relationships that might hold between the different laws and the different realms would seem to be just as compatible with the presence of causal bonds as with their absence.

To appreciate the sophistication and power of Knutzen’s overall case for physical influx, recall what he has accomplished in the *Systema causarum efficientium*. He has presented several distinct arguments in favor of physical influx that challenge a Leibnizian system in provocative ways without being based on principles that would be completely foreign to it. He has also responded to two of the most important objections that Leibniz raised against physical influx by noting that the law of the conservation of living forces holds only for bodies, which is a distinct realm from that of monads, and by developing a more detailed model of intersubstantial causation. According to this model, physical influx is not literally a migration or transfer of accidents, as Leibniz suggests, but rather the force or power one substance has to act on another substance. In addition to being distinct from the migration model, this model has the advantage that it appeals to a concept of force or power that a Leibnizian cannot object to, since the Leibnizian must invoke an analogous force to explain the change of state within a substance. Knutzen also allows that the powers of both substances can be involved in physical influx insofar as the power of one substance modifies or activates the power of another, which

then produces the new property. For example, when the body acts on the mind, a corporeal element (or set thereof) is causally efficacious in modifying the substantial representational power of a noncorporeal element in such a way that corresponding new representations are produced out of it. This model allows him to avoid the untoward consequence that a body (something material) is completely sufficient to produce representations (which Knutzen holds to be immaterial), since the noncorporeal element (i.e., its representational power) can be responsible for producing the specifically mental dimensions of the representations, even if it is not sufficient for explaining these representations in their entirety.⁹⁹

The philosophical subtlety and force of Knutzen's work did not go unnoticed at the time. In fact, Knutzen's work represents the crucial turning point in 1735 against pre-established harmony in favor of physical influx. Whereas previous Wolffians had generally followed Wolff on the issue and Gottsched and Hollmann did little to change this fact, after Knutzen's work, in the 1740s and 1750s, physical influx became increasingly acceptable for Wolffians, such as Reusch, Plouquet, and Darjes, and an obvious choice for non-Wolffians, such as Reinbeck, Euler, and Crusius.¹⁰⁰

BAUMGARTEN AND MEIER: A NEW CASE FOR PRE-ESTABLISHED HARMONY

Bolstered by Knutzen's defense, the theory of physical influx represented a popular position during the rest of the 1730s and the 1740s. In 1737 Lange even attempted to have pre-established harmony condemned again by petitioning the king. The king set up an independent committee to decide the matter, appointing Johann Gustav Reinbeck as its head. After much deliberation and despite the fact that Reinbeck himself held physical influx, his report declared that pre-established harmony should not be condemned.¹⁰¹ Alexander Gottlieb Baumgarten (1714–1762),

⁹⁹ If the noncorporeal element were sufficient, then the corporeal element would be superfluous.

¹⁰⁰ Euler considers and rejects pre-established harmony in his *Gedanken von den Elementen der Körper* (Berlin, 1746), *Reflexions sur l'espace et le temps* (Berlin, 1748), and *Lettres à une princesse d'Allemagne sur divers sujets de physique et de philosophie* (St. Petersburg, 1768–1772). For Darjes, see §§54–90 in *Elementa Metaphysices*, vol. 2 (Jena, 1743–1744), and for Plouquet, see §§456–483 in *Principia de Substantiis et Phaenomenis* (Frankfurt, 1753 and 1764).

¹⁰¹ See Johann Gustav Reinbeck, *Erörterung der philosophischen Meynung von der sogenannten Harmonia Praestabilita, worinnen gezeigt wird, 1. was diese Hypothesis eigentlich sagen wolle,*

professor of philosophy at Frankfurt on the Oder, and one of his students, Georg Friedrich Meier (1718–1777), who eventually became a professor of philosophy at Halle, attempted to take advantage of the space created by Reinbeck's decision by making a new case for pre-established harmony. Thus, physical influx was by no means unopposed in Germany in the late 1730s and 1740s.

Baumgarten and the Perfection of the World

In 1739 Baumgarten published *Metaphysica*, a metaphysics textbook whose relevance to our current context is underscored by the fact that Kant used it in the classroom for virtually his entire career.¹⁰² In this work Baumgarten departs from Wolff not only by considering the issue of causality in its more general, cosmological form (such that he can then apply it to the mind-body problem in his treatment of psychology), but also by setting up the debate differently. His discussion of cosmology, in Part II, is divided into three chapters: on the concept of the world, on its parts, and on its perfection. In the third chapter, after arguing in the first section (esp. §441) that the most perfect world will have the greatest possible universal connection, harmony, and agreement, he turns, in the second section (§§448–465), to the “interaction of mundane substances.” He begins (§448) by noting that everything in the world is in

und warum sie der menschlichen Freyheit nicht nachtheilig sey. 2. Was dieselbe vor dem Systemate influxus für einen Vorzug habe, und 3. Warum der Autor nichts destoweniger derselben nicht beypflichtete, aus Liebe zur Wahrheit und zur Verhütung fernerer verworrenen Streitigkeiten, nebst einem Vorbericht herausgegeben (Berlin, 1737). Lange's attempt thus backfired, and pre-established harmony and physical influx received renewed attention. See, e.g., Johann Friedrich Bertram's *Beleuchtung der Neu-getünchten Meynung von der Harmonia Praestabilita durch Veranlassung der jüngst-edirten Reinbeckischen Erörterung* (Bremen, 1737), a work by an anonymous “*theophili sincer*” entitled *Sendschreiben an Alethophilum, darin deutlich wird, daß der Herr Probst Reinbeck die Wolffische Meynung von der Harmoniae Praestabilita in der That angenommen habe, es nur nicht Wort haben wolle*, as well as *Acht neue merckwürdige Schriften, die in der Wolffischen Philosophie von neuem erregte Streitigkeiten betreffend*, both published in 1737 (n.p.). For further details, see Fabian *Beitrag zur Geschichte des Leib-Seele-Problem*, pp. 113–114.

¹⁰² For more on Kant's use of Baumgarten's *Metaphysica*, see Karl Ameriks, “The Critique of Metaphysics: Kant and Traditional Ontology,” in *The Cambridge Companion to Kant*, ed. P. Guyer (New York: Cambridge University Press, 1992), pp. 249–279. For an account of the various textbooks that Kant used throughout his career, see Emil Arnoldt's “Möglichst vollständiges Verzeichnis aller von Kant gehaltenen oder auch nur angekündigten Vorlesungen nebst darauf bezüglichen Notizen und Bemerkungen,” in *Gesammelte Schriften*, ed. Otto Schöndörffer (Berlin: Verlag von Bruno Cassirer, 1909), pp. 173–344.

universal harmony and mutual influx. He then notes (in §§449–450 and §452) that the causal theories of pre-established harmony, physical influx, and occasionalism are all ways of explaining this universal harmony and mutual influx. Accordingly, Baumgarten has shifted the terms of the debate by using “universal harmony” and “mutual influx” to indicate the phenomena that the three causal theories are to explain rather than expressions that denote any of the three causal theories themselves.¹⁰³

Further, Baumgarten’s primary argument for pre-established harmony (and against physical influx and occasionalism) is based on the perfection of the world by maintaining that pre-established harmony would allow for a more perfect world than either physical influx or occasionalism. His reasoning is that since perfection consists in the greatest harmony or agreement of things according to their grounds and since pre-established harmony in its Leibnizian form (with each monad representing every state of every other monad) posits that every finite monad contains the sufficient ground of every change in every other finite monad, a world governed by pre-established harmony would contain the greatest amount of perfection. In particular, it would, he thinks, contain more perfection than worlds in which only some finite substances contained the ground for some states of some others, as would be the case for physical influx, or worlds in which no finite substance contained a ground or reason for any of the states of any others, as would be the case for occasionalism (which would therefore possess the least amount of perfection).

What is striking about this argument is that it exploits the idea that one monad would contain the sufficient ground for every state of every other monad (as well as the idea that perfection is to be measured by this type of grounding relationship). While it is one thing to say that every Leibnizian monad represents every other one, it would seem to be quite another to say that every such monad contains or is the ground (and, in fact, the sufficient ground) of the state of every other. After all, on Leibniz’s description of pre-established harmony, it might seem that a substance could contain the grounds only of its own states, since it is causally responsible only for its own states, and not for the states of other substances, even if it does represent them.

The key to understanding this idea properly lies in the way that Baumgarten modifies a distinction that Leibniz draws, for example, in §66 of the *Theodicy*, between what is called real and ideal influence or dependency. Leibniz’s idea is that even if the mind and body do not act on

¹⁰³ Baumgarten’s shift in terminology was influential in some quarters.

each other causally, he still wants to claim that there is a *kind* of mutual dependence between their states. The mind and the body depend on each other

ideally, insofar as the reason of that which is done in the one can be furnished by that which is in the other. This had already happened when God ordered beforehand the harmony that there would be between them. Even so would that automaton, that should fulfil the servant's function, depend on me *ideally*, in virtue of the knowledge of him who, foreseeing my future orders, would have rendered it capable of serving me at the right moment all through the morrow. . . . For insofar as the soul has perfection and distinct thoughts, God has accommodated the body to the soul, and has arranged beforehand that the body is impelled to execute its orders. And insofar as the soul is imperfect and as its perceptions are confused, God has accommodated the soul to the body. . . . This produces the same effect and the same appearance as if the one depended immediately upon the other, and by the agency of a physical influence. . . . Each one is assumed to act on the other in proportion to its perfection, although this be only ideally, and in the reasons of things, as God in the beginning ordered one substance to accord with another.¹⁰⁴

Since the "dependence" of mind and body on each other is not strictly causal, but rather depends on God's knowledge, Leibniz calls it ideal.

Baumgarten modifies the underlying conceptions of real and ideal influence accordingly as follows (§212): "If the passivity of a substance that another influences is at the same time an action of the one being acted upon [i.e., of the patient], then the PASSIVITY and INFLUENCE are said to be IDEAL. If, however, the passivity is not an action of the patient, then the PASSIVITY and INFLUENCE are said to be REAL."¹⁰⁵ Accordingly, if two substances are in harmony and mutually influence each other (which they must if they are to belong to the same world), there are two possibilities. Either the substance undergoing the change does so by means of its own action, in which case Baumgarten wants to call that influence ideal, or it does so without any action of its own, in which case the action of another is required and the influence is real. Thus, for Baumgarten, if two substances are related by ideal influence, pre-established harmony holds, whereas if they are related by real influence, physical influx (or influence) holds.

Baumgarten's notion of an ideal influence thus represents an attempt at defining certain primitive metaphysical concepts in such a way that he can highlight what he takes to be the advantages of pre-established harmony. On the one hand, because the relationship between finite

¹⁰⁴ Leibniz, *Theodicy*, §66, pp. 158–159.

¹⁰⁵ Baumgarten *Metaphysica*, §212, 17:71.

substances is merely ideal, he can still draw a contrast between pre-established harmony and physical influx. On the other hand, because an ideal influence is still an influence in the sense that it tracks a relation between, for example, the mind and the body, it allows him to posit a connection between substances that supports the claim that the one contains the sufficient ground of the state of the other (even if it is an ideal rather than real ground), thereby increasing the amount of perfection that the world has, according to pre-established harmony.

In his discussion of the three causal theories, Baumgarten then turns what might appear to be a purely terminological shift into an explicit objection to physical influx. For he argues:

Hence according to the universal system of physical influx no substance that is part of the world can act by its own force in any of its harmonious changes (§448). Now all changes of a substance in the world can in fact be sufficiently cognized from the force of any other monad that belongs to this world (§354, §400). Therefore, all changes are harmonious (§448), and hence no mundane substance acts in any of its changes, according to the universal system of physical influx, but is rather really acted upon by other substances of the world that, however, do not, on account of this reason, act at all (§210), and hence are not forces [in the proper sense] (§197).¹⁰⁶

The idea behind this objection seems to be that if substances were to act only on others and not on themselves, then they would not truly be substances or forces in the proper sense, since such substances or forces contain the reasons for their own states in themselves. Regardless of whether this objection begs the question by assuming a certain conception of substance or force, it clearly rests on an uncharitable reading of what universal physical influx is committed to by assuming (1) that a substance cannot act on itself at the same time that another substance acts on it and (2) that a substance cannot be responsible for any of its own states. As we saw above, Knutzen rejects the first assumption, and the second assumption is warranted only for the most extreme version of physical influx, *universal* physical influx in the sense that *all states* of every substance must be caused by every other substance, and thus does not address more modest versions, like Knutzen's, which would hold that only some states of one substance are caused by another.¹⁰⁷

¹⁰⁶ Ibid., §451, 17:120–121.

¹⁰⁷ In his *Beweis der vorherbestimmten Übereinstimmung* (Halle, 1743), Georg Meier presents the same objection as follows:

All harmonious changes of the substances of this world can be known from other finite substances. Thus, they have their reason in other finite substances that are present beyond those whose state is changed in a harmonious way (§9). That change of state

Nevertheless, this objection raises an important issue. If the changes in one substance are in some sense contained or reflected in the other and, moreover, if one substance changes “because” of changes in the other, there is an important sense in which the one could be called the cause of the other, even if this sense is weaker than what proponents of physical influx endorse. But if one substance contains the reason for the state of another, Baumgarten (or Leibniz) might ask what more one could wish for. Our everyday language suggests that there is causal interaction between objects, but it might seem to be unclear what further content is to be associated with “real” causal interaction and, accordingly, that the cost of denying real interaction is minimal as long as one can still say that one object contains the reason for the state of the other. At the same time, since Leibniz and Baumgarten need the notion of a real causal influence in order to distinguish their own position from occasionalism, they cannot press too hard on this point.

Meier

While Baumgarten presents his views on causality as one topic among many within the context of his general metaphysics textbook, Meier devotes an entire treatise, *Beweis der vorherbestimmten Übereinstimmung* (*Proof of Pre-established Harmony*), to the issue in 1743. Although Meier follows Baumgarten’s main line, he feels justified in advancing pre-established harmony as a theorem rather than as a hypothesis that would be merely probable, because he both raises objections to physical influx and occasionalism and presents arguments in favor of pre-established harmony. While his positive argument for pre-established harmony and one of his main objections to physical influx are not significantly different from Baumgarten’s, Meier does develop several novel objections to physical influx. First, Meier presents the following objection:

When a finite substance acts, its inner state is thereby always changed. Or, whenever a finite substance acts, an inner state is produced in it by this action that was not to be met with in it before it acted. Assume the opposite. A finite substance is

that has its reason in another thing is called a passion, and thus all harmonious changes in this world are passions. These passions have their reason in other finite substances. All harmonious changes in the world are accordingly natural passions (§14). A universal influxionist considers all natural passions to be real passions (§14). Consequently, a universal influxionist must consider all changes in this world that are produced naturally to be real passions. Accordingly, not a single change of state can be produced through a substance’s own force in which this change is effected, thus it behaves merely passively in this case (§11). . . . But what influxionist would admit this? (p. 72)

supposed to act, but not produce through this action any determination in itself. A determination is produced by every action (§46). Consequently the determination brought about by this substance would have to be produced outside of it. Accordingly, one can view this substance in a two-fold state, prior to its action and while it acts. Accordingly, if it were not itself changed by its action, it would, when it started the action, remain completely the same substance, without the smallest change, that it was before it acted. Now because the consequences remain the same when the grounds are the same (*posita eadem ratione ponitur idem rationatum*), the assumed substance, or force, would bring about the same effects in this, its two-fold state. Before the action it did not cause the determination that was supposed to be produced, otherwise it would not need to act (§46). Consequently, it will also produce no determination in the other state when it is in the process of action, which contradicts not only the concept of action, but also what would have to be assumed if one wanted to deny my theorem.¹⁰⁸

Meier's argument seems to depend on the same terminological shift that we saw in Baumgarten, namely one according to which intrasubstantial causation is equated with pre-established harmony and intersubstantial causation that excludes intrasubstantial causation is identified with physical influx, an alignment that neglects the possibility that *both* substances could be involved in the production of any determination or property.

However, Meier's argument focuses on the "mechanics" of causation in a way that Baumgarten's did not. In particular, he objects that physical influx cannot explain how it is that the grounds of one substance could be responsible for changing properties in another substance. For insofar as one substance changes the state of another substance it must thereby also change itself. For if it did not change itself, then it could not have changed the other substance, given that the same grounds that were originally present in the first substance continue to be present and one set of grounds cannot be responsible for first one set of determinations and then a contrary set.¹⁰⁹ But if it changes itself, then the proponent of physical influx is committed to intrasubstantial causation as well, which Meier (question-beggingly) equates with pre-established harmony to the exclusion of physical influx.

Meier presents a second objection to physical influx that runs as follows:

I first want to consider the smallest substance or force and show that it cannot act physically on another. It is self-evident that this substance is finite, because

¹⁰⁸ Meier, *Beweis*, pp. 88–90.

¹⁰⁹ This principle shall be important in considering Kant's principle of succession below (in Chapter 2).

God possesses the largest force. If one wanted to object that no substance is really present that would be the smallest, I will admit this for the sake of the general harmony in the world. Instead, let one assume a greater substance that, however, uses its force only to the smallest degree. One will surely admit this case. Cannot all larger actions and forces rightly be viewed as a summation of the smallest actions and forces? Thus posit the smallest force, or a greater substance that, however, uses its force only in the smallest degree. This substance can undertake only the smallest action (§47). But can this action be real influx? Let us assume it for a while. Through this smallest action a determination in the smallest substance itself is produced (§49, §50). This determination cannot, however, be smaller than the smallest. The smallest is precisely the one that, if it were smaller, would contain a contradiction. Now if a real passion is supposed to be brought about in another substance by this smallest action – for it is assumed that this smallest action is a physical influence (§13, §11) – by this action alone a further determination would have to be effected that would have to be at least the smallest. Thus, if the smallest action of a finite substance could be real influx, then, through the smallest action, at least two smallest determinations would have to be effected, which is impossible (§47).¹¹⁰

Meier's idea here is that if it is possible for a substance to bring about a smallest effect, then physical influx is impossible. Why? Because if a substance acts on another substance, it brings about two effects, one in itself and one in the other substance. However, by producing two effects, it cannot be producing the smallest possible effect, since two such effects are necessarily greater than the effect that would be produced by a substance that brings about a change only in itself. Accordingly, physical influx is incompatible with the possibility that a substance could bring about a smallest effect. Though Meier's objection is obviously specious, it is nonetheless relevant to see how he is focusing on metaphysical details of causality rather than any alleged ethical, political, or religious implications of pre-established harmony.

By focusing on detailed metaphysical issues concerning what perfection is, how a substance can cause an accident, and whether one can distinguish between causal and quasi-causal influences, Baumgarten and Meier thus show that pre-established harmony continues to represent a significant philosophical option, indeed, an option for which new arguments can be adduced and from whose standpoint novel objections to physical influx and occasionalism can still be raised. Moreover, pre-established harmony continues to find important supporters throughout the 1740s, such as Stiebritz and Baumeister, and, as late as 1755, one of

¹¹⁰ Meier, *Beweis*, pp. 92–93.

Prussia's most popular and influential proponents of the Enlightenment, Moses Mendelssohn, still adhered to this doctrine.¹¹¹

CRUSIUS AND FUNDAMENTAL POWERS

Although pre-established harmony continued to represent the orthodox Wolffian position in the 1740s and 1750s – in spite of the notable defection of Knutzen and several others – the case for physical influx found significant support with a leading Pietist, Christian August Crusius (1715–1775). Crusius was born and raised in Leipzig and taught by a staunch Pietist, Adolf Friedrich Hoffmann, before becoming a professor there, first in philosophy, then in theology – a move that was not at all uncommon at the time, though in his case it truly did correspond to a shift in his interests as well. His publications in philosophy, all in the 1740s, display sophistication, subtlety, and originality in mounting a full-scale attack on Wolff's position that would radically alter the philosophical landscape that Kant would encounter.

Part of what makes Crusius's work so important at a general level is the novel way in which he combines elements from a variety of distinct traditions. For example, his Pietism finds expression in his claim that “free and rational spirits” are God's ultimate purpose in creation, and his sympathy with empiricism is clear from his view that all of our general representations (including even our most basic concepts of ontology) must be abstracted from actual objects that are presented to us in sensation. At the same time, his conception of metaphysics draws on Leibniz's and Wolff's in several fundamental respects. He holds that metaphysics is about the necessary truths of reason, which he discusses in his most important philosophical work, *Sketch of the Necessary Truths of Reason* (*Entwurf der nothwendigen Vernunft-Wahrheiten*), published in Leipzig in 1745; he follows Wolff's division of metaphysics into ontology, cosmology, psychology, and natural theology (even if he thinks they should be treated in a different order from Wolff's); and his focus on what he calls “metaphysical essences” as what distinguishes one thing from another is similar in certain respects to Leibniz's doctrine of complete concepts. Despite these similarities, however, he is highly critical of Leibniz's and Wolff's

¹¹¹ See Johann Friedrich Stiebritz, *Erläuterungen der Wolffischen vernünftigen Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt* (Halle, 1742–1743); Friedrich Christian Baumeister, *Institutiones metaphysicae* (Wittenberg, 1743); and Moses Mendelssohn, *Philosophische Gespräche* (Berlin, 1755).

positions insofar as he accepts libertarian freedom for finite rational creatures, a voluntaristic conception of God, and substantive principles that cannot be derived from the principles of contradiction and sufficient reason (much less, from the former by itself, as Wolff had maintained).

Yet it is also clear that even if Crusius does draw on previous philosophers in a somewhat eclectic way, he is an insightful and original thinker in his own right. For example, he claims that substantive principles that extend beyond the principles of contradiction and sufficient reason can be known on the basis of “the essence of our understanding,” which suggests that he has abandoned, at least implicitly, any traditional conception of the understanding. He also argues that space and time are fundamental principles of the existence of things, since we cannot, he believes, even think of a complete thing without thinking of it as existing somewhere and at some time.

Crusius’s originality is also evident with respect to the specific issue of causality. First, he bases his ultimate metaphysics on his account of causality insofar as he (i) devotes more attention to what a fundamental power is than to any other metaphysical topic and (ii) views it as the core notion on which others (such as that of freedom) depend. Second, he develops the causal notion of a real ground to explain what a world is in the cosmology section of his metaphysics and to argue against pre-established harmony and for physical influx. Finally, he provides a detailed description of how the mind and body, as two substances that stand in a real, that is, causal connection in the world, can and do act on each other. Thus, Crusius’s reflections on causality represent an original contribution in this context.

Basic Ontological Concepts: Powers, Grounds, and the Possibility of Freedom

In the ontology section of his *Sketch of the Necessary Truths of Reason*, Crusius begins his analysis by distinguishing between the essence and the existence of a thing. Although Crusius’s provocative reflections on space and time (which are directly relevant to Kant’s views in the *Transcendental Aesthetic*) pertain to the existence of things, the vast majority of his attention is devoted to discussing the essence of a thing. In the course of his discussion of metaphysical essences (in chapter 3), Crusius introduces the notion of a power (§29) as the most basic notion that one might use to characterize a metaphysical essence: “The existence of a thing cannot be viewed as equal to its non-existence. Accordingly, through

every thing something else must become possible or actual, whether it be made possible or actual through itself alone or by adding several things. The possibility of one thing, B, which is connected to another thing, A, is called a **power** in the broadest sense in thing A. Consequently, every thing has several powers, but at least one power.”¹¹² After linking the concepts of power and causality, Crusius turns (§34) to consider what a ground or cause is in the broad sense:

Everything that brings about something else either in part or in whole and insofar as it is viewed as such is called a **ground** or **cause in the broad sense** (*principium, ratio*). For that reason efficacious causes are one kind of ground, whose necessity is clear from the preceding (§15, §29). But they are not the only kind. For that reason we must also consider here the remaining kinds of grounds. Namely, what one calls grounded and whose production one attributes to another is either cognition in the understanding or it is the thing itself, outside of our thoughts. For that reason a ground is either a ground of cognition, which can also be called an ideal ground (*principium cognoscendi*), or a real ground (*principium essendi vel fiendi*). A ground of cognition is one that brings about cognition of a matter with conviction and is viewed as such. A real ground is one that brings about or makes possible, either in part or in whole, the thing itself, outside of our thoughts.¹¹³

Given this distinction between real and ideal grounds – which is different from Leibniz’s and Baumgarten’s distinctions between ideal and real influences and the kinds of grounds that might underlie them – Crusius then draws a further distinction among real grounds, a distinction that introduces a further novel element into his metaphysics, both generally and at the level of particular issues (such as cosmology and the mind-body relationship):

§36. *Further division of real grounds into efficacious causes and inefficacious real grounds or existential grounds.* When a **real ground** brings about or makes possible a thing outside of thought, it does so either by means of an efficacious cause and, in that case, is called an **efficacious cause**. Or the laws of truth in general do not allow anything else other than that after certain things or certain of its properties have already been posited, something else is now possible or impossible, or must be possible in this way and not otherwise. This kind of ground I wish to call an **inefficacious real ground** or also an **existential ground** (*principium existentialiter determinans*). Accordingly, an **existential ground** is one that makes something else possible or necessary through its mere existence due to the laws of truth. E.g., the three sides of a triangle and their relations to each other constitute a real

¹¹² Christian August Crusius, *Entwurf der nothwendigen Vernunft-Wahrheiten*, reprinted in Christian August Crusius, *Die philosophischen Hauptwerke*, ed. G. Tonelli (1745; rpt. Hildesheim: Georg Olms Verlag, 1964), vol. 2, pp. 45–46.

¹¹³ *Entwurf*, pp. 52–53.

ground of the size of its angle, but only an inefficacious or existential ground. By contrast, fire is an efficacious cause of warmth.¹¹⁴

Later, in chapter 5, which focuses on causality, Crusius elaborates on this distinction as follows:

§79. *Division of power in general into the inefficacious capacity of an existential ground and an active power.* Whatever a cause contributes to the production of an effect, it accomplishes either 1) **through its mere existence because through it the existence, or a certain manner of existing, of another thing is made possible, impossible, or necessary.** For this reason, in that case, its power is also nothing other than the possibility, impossibility, or determination of another thing that is connected to the mere existence of a thing by means of the laws of truth. Above, we called such causes **existential grounds** (§36). The power thereof can be called the **inefficacious capacity of an existential ground** (*facultas existentialis*). E.g., a wedge or a lever has the power to create some easing by overcoming a resistance. But this occurs by means of the mere existence of their shape and structure. For this reason they are merely existential grounds, and their power consists in an inefficacious capacity. This is how it is with all mechanical causes, that is, with all those substances that, and to the extent to which, they have an influence on the determination of their effect by the shape and position of the parts of a composite thing. Or 2) **the cause acts due to an inner property of its essence, which is now directed toward the production of this effect.** One thus attributes an **activity or self-activity** to it. It is called an **active cause** and its power an **active power** (*Facultas activa*). Thus, an **active power** is a property connected to a substance belonging to its inner essence due to which something else is actual through it or comes to be, without it being merely a conclusion that one would immediately have to concede according to the principle of contradiction after positing existential circumstances. Of such a sort are the active powers of the elements, thinking, and desiring. **Both of these can even coincide in a single cause.** E.g., when a body acts, an inefficacious capacity by which its effect is to a certain extent determined lies in its shape and in the position of its parts. However, one should not forget the active power of the elements or the active power of other things external to it that move it, without which the effect could not occur. E.g., the axe splits the wood according to the laws of the wedge. To that extent, it is an existential ground. But the active power is to be sought in the elements it consists in, in the combination of the bodies among each other in the world, and in what guides them.¹¹⁵

In addition to the causal activities of substances endowed with the power to think and desire – activities that Leibniz might understand as merely nominally distinct from his own notions of perception and appetite – Crusius adopts a *nonactive* causal principle that pertains at least to matter and a mechanical cause that generates motion, calling such a cause an

¹¹⁴ Ibid., pp. 34–35.

¹¹⁵ Ibid., pp. 135–137.

existential ground, since it brings about its effect by means of its mere existence. Further, Crusius implicitly suggests (both due to the geometrical example he gives and by contrast with his account of active grounds) that an existential ground brings about its effect by means of the principle of contradiction. In connection with this point, in an earlier passage (§59) Crusius seems to suggest that space and time contain existential grounds insofar as he calls them “inefficacious possibilities” and sees them as “distinct from the power of the efficacious causes” that are located in space and time. Accordingly, existential grounds are closely connected to space and time, which are primitive principles of the existence of things.

While Crusius’s introduction of existential grounds is thus significant in its own right, the context in which he does so is noteworthy as well. For Crusius’s main purpose in chapter 5 is to differentiate power in the broad and narrow senses, the latter of which he calls “fundamental power,” describing, at length, its eight distinguishing features. He also discusses the related notion of a fundamental activity, providing another lengthy description of its distinctive features. The ultimate point of these detailed discussions is that they allow him (1) to explain how it is that (libertarian) free actions are possible and (2) to establish (§84) the principles of sufficient and determining reason. The principle of sufficient reason holds even for fundamental activities of freedom, since “nothing is lacking that is necessary for causality,” whereas the principle of determining reason holds for all fundamental activities that are not free actions and maintains that after something’s determining ground is posited, it “cannot be or occur otherwise.” In this way, Crusius places causal notions at the core of an account that is to have the resources necessary to reconcile “the first actions of freedom” (§83) with deterministic causation of the sort found in nature.

The “Real” World

Crusius’s discussion of the world in the cosmology section of the *Sketch of the Necessary Truths of Reason* draws on the account of causality developed in the ontology section. For example, he defines the world (§350) as “a real connection of finite things that are not in turn themselves a part of another to which they belong by means of a real connection. Or: **a world** is a system of finite and really connected things that is not in turn itself contained in another system.”¹¹⁶ Wolff had defined the world

¹¹⁶ *Ibid.*, §350, p. 657.

similarly in terms of (spatio-temporal) connections between things, but then explained that these connections were what Baumgarten and Leibniz both called ideal.¹¹⁷ Thus, what distinguishes Crusius's definition of the world from Wolff's is primarily the addition of the word "real" or "really" (*realiter*).¹¹⁸ In §359 he explicitly incorporates his earlier distinction between active and existential grounds into this definition:

Because the world is a system of things whose parts have a real connection even outside of thought (§350), **things in the world must be able to act on each other** so that the one, as an efficacious cause (§36), can change the state of the other (§94). Now this relation can be mutual or not. For this reason two kinds of things are possible in the world: First, **active things**, which can act on others just as others can act on them, and **merely passive things**, which do not have an active power, but rather make something else possible, impossible, or necessary by means of its very existence (§36). Now because the possibility of causing something else is called a power, two different kinds of powers are also possible in the world, namely active powers and those that are only inefficacious capacities of an existential ground (§79).¹¹⁹

In short, both active and existential grounds can form a real connection between things in order for them to constitute a world.

Accordingly, the account of causality that Crusius developed in the ontology section of the *Sketch* provides the central notion of his cosmology insofar as the world is defined in terms of the causal notions articulated in that account. Moreover, his distinctive combination of ontology and cosmology reveals a commitment to physical influx. For he explicitly asserts that the connection between things that are to form a world must be real and cannot be merely ideal. If the connection between such things were ideal, that is, consisted in our thoughts alone, then there would be only one possible world, since every possible object of thought could be connected ideally, that is, in our thoughts. But to assert that there is necessarily only one possible world undermines one of the main points of invoking possible worlds, since the introduction of possible worlds supplies us with a means of talking clearly about various different ways in which the world could have been. Crusius is thus suggesting that "real" causal connections, that is, connections that exist outside our thoughts,

¹¹⁷ See Wolff, *Vernünfftige Gedanken*, §§540–548.

¹¹⁸ Another important difference is that Crusius attempts to derive (§§351–354) the following features of the world from its definition: its finitude, creation, and hence beginning in time, its conservation at every moment by God, and that creation must include rational and free creatures in it.

¹¹⁹ *Entwurf*, §359, p. 677.

are needed to connect things into a single world as distinct from other sets of causally connected things.

Crusius had paved the way for this point earlier in the ontology section (in the chapter on identity and distinctness):

Any connection of finite things that is to be a real *unio existentialis* outside thought must rest on a causal connection of things due to which at least one must act on the other, but also both can act on each other reciprocally as well as be passive with respect to each other. For there is otherwise nothing else outside thought that can provide a ground of connection between complete things. But as soon as one takes this away, then one must connect them only in a concept in the understanding, i.e., the things thus have either no connection or one that is merely ideal.¹²⁰

Crusius immediately uses this point to introduce a new criticism of pre-established harmony:

Consequently, I cannot, e.g., admit that those who believe in pre-established harmony leave a real connection between body and soul. . . . Their connection is only ideal even with respect to God. One cannot even say that they are connected by the intervention of God. For then at least the arrangement of the essences of the body and soul would have to be attributed to God. But the defenders of pre-established harmony can never say this in the Leibnizian sense because they do not leave God any honor beyond bringing the essences of substances into existence, rather than arranging them, since all essences are supposed to be eternal. Thus, a mere correspondence rather than a real connection remains.¹²¹

Crusius's objection here starts with the familiar idea that a Leibnizian explanation of the union of the mind and the body is insufficient, because the connection between the mind and the body, lacking causal relations, is only ideal and not real. But Crusius takes the objection further by responding to the rejoinder that a Leibnizian might make, namely that God could take it upon himself to establish a special union between mind and body.¹²² Crusius objects that such a rejoinder is not available to a Leibnizian, because on the Leibnizian account, at least as Crusius understands it, God can establish only a *correspondence* between substances and not a *real connection* or dependence.¹²³ For although God brings

¹²⁰ Ibid., §94, pp. 160.

¹²¹ Ibid., p. 161.

¹²² It may be that Crusius is insinuating other arguments here as well. For example, his allusion to the eternity of substances (and thus the world) is a consequence he clearly believes to contradict the Principle of Contingency (§33), which he states as follows: "that whose non-being can be thought really did not exist at one time, which one calls the Principle of Contingency" (ibid., p. 51).

¹²³ As we shall see in Chapter 2, Kant reiterates Crusius's criticisms (28:215).

beings into existence, God is not responsible for their essences, since their essences are given necessarily in the divine understanding and cannot be altered by the divine will. Crusius is thus claiming that, according to pre-established harmony, although God can bring into existence substances that correspond to each other, God cannot bring them into existence with a real but non-causal connection. The best Leibniz can do, so Crusius thinks, would be to “fabricate a special class of connection which one would call the metaphysical [class], and which would consist in the one thing representing the other, but then the original concept of real connection is abandoned.”¹²⁴ In short, either the connection between them is causal, but then pre-established harmony has been abandoned, or it is not, in which case one is left with “mere correspondence” and an inappropriate (he thinks) restriction on God’s powers.

It is clear that part of the source of Crusius’s objection here stems from his voluntaristic conception of God, according to which God ought to have control over the relations between substances, in contradistinction to his understanding of the Leibnizian position, according to which God would not have any control over what connections could exist between substances. Yet the cause of the disagreement may also stem from differences in what a “real” or fully causal connection between two substances would amount to, in contrast to the kind of harmony-based correlations with which a Leibnizian might be satisfied. As we saw in our discussion of Baumgarten, there is a serious question about what the content of a *real* causal connection is beyond mere correlations (or regularities). While Crusius does not clearly articulate which of these two issues underlies his objection, he is aware of a significant systematic issue when he notes:

[To proving the existence of God] it belongs in part that God brought about the simple substances and their fundamental powers (§144, §145), which can be called **creation in the narrow sense**; in part that he connected them with each other in a certain order, which, viewed generally, is an efficacy that is not proper to God per se as the former was. Finite substances, too, can cause certain connections between things that are already there, but only in those things that are subject to their powers (§145). However, the connection of things in the entire world surpasses their capacity.¹²⁵

This passage suggests (without giving either a full explanation or a proper justification) that God must create not only finite substances, but also their general cosmological relation(s), since that sort of relation exceeds

¹²⁴ *Entwurf*, §94, p. 161.

¹²⁵ *Ibid.*, §327, p. 609.

the powers of the finite substances. This would presumably contrast with the Leibnizian view, according to which God creates (ideal) relations between substances in the very act that brings them into existence.¹²⁶

In his cosmology section, Crusius also presents an argument in favor of physical influx that is based on similar considerations, even if it occurs in a different context and has limited force against proponents of pre-established harmony. In §363 Crusius claims that spirits and matter must be able to act on each other. “For, according to its essence, matter is not God’s ultimate purpose, but rather a means (§354). Consequently, it must be created for the sake of a real connection that it has either mediately or immediately with those creatures that are God’s ultimate purpose, namely rational and free spirits. But it could not have this if spirits and matter could not act on each other.”¹²⁷ If matter did not act on minds, there would be no point in God creating matter, since it would contribute nothing to God’s purpose. Since matter does exist, it must act on minds. In addition to the fact that this argument has no force against idealists, it presupposes that only real rather than ideal connections could be relevant to God’s purposes in creation, an assumption that someone such as Baumgarten would obviously reject.

Mind-Body Interaction and the Power to Move

If Crusius has thus established that things must stand in real, causal connections in order to belong to the same world, what do these cosmological views imply for his account of the mind and the body in the chapter on psychology? Even if Crusius simply assumes that minds and bodies, as the kinds of things that inhabit this world, must be able to act on each other, it still leaves open the question of how they can do so. Crusius takes up this question in this chapter by responding to two of the objections typically raised against physical influx and by providing a positive description of his model of causal interaction between mind and body.

Both of the objections that Crusius considers here stem from the difficulty that arises from the radical distinctness of the mind and the body.¹²⁸

¹²⁶ That this issue is an important one for the pre-Critical Kant will become apparent in Chapter 2.

¹²⁷ *Entwurf*, p. 683.

¹²⁸ In the cosmology section, Crusius also considers the traditional objection that physical influx would violate the law of the conservation of motion (and living forces). However, he considers the issue not as an objection to his view, but rather in the context of discussing the laws of motion. In considering whether it is a law of nature that motion

The objections can be seen as two horns of a dilemma. Either the mind and the body are radically distinct, in which case they cannot act on each other (given that they have nothing in common), or they are not radically distinct, in which case one has turned mind into matter (or vice versa). The basis for Crusius's responses lies in his discussion of the world. At §362, Crusius explains how activity between substances can occur as follows:

Finite things in the world can act on each other only through motion. No finite thing can act on another except through motion (§145). . . . Either they thereby only move each other, or inner active powers are awakened. But **when one substance acts on another through motion, either it moves the other from its space either only due to the impenetrability of both, or the motion caused in the one, or its effort to move, becomes, according to a rule, a condition under which a certain active power acts on the other substance or is awakened** (§74). By contrast, not all inner activity of substances is motion, but rather it can also be a thinking or desiring. Accordingly, we find here a reason for distinguishing **two highest main classes** of substances in a world. Either they have no power other than the capacity to move, and then we want to call it **matter**. Or they have another power that is different from the capacity to move . . . , in which case we want to call them **spirits in the broadest sense**.¹²⁹

While this passage thus suggests that mind and matter are distinct, shortly thereafter, in §364, Crusius minimizes the distinction between matter and spirits by arguing that spirits necessarily have the power to move:

The capacity to move can be understood from the essence of a finite spirit. Further, one asserts without reason that the capacity to move cannot be understood from the essence of a spirit. It can in fact be understood from the essence of a spirit, not from the differential essence, but rather from the general essence of a finite spirit, which belongs to it just as much. For no substance can be understood otherwise than as impenetrable. On the basis of counter-arguments, we have already established (§51, §58, §250) that this proposition is subject to an exception in the case of the infinite substance of God, that all creatures are penetrated by him, and that he is with them at the same time in the same place as they

be conserved in the world, he argues that it cannot, since if it were, the absurd result would follow that minds could not cause any motion and that matter would not be able to fulfill the purpose for which God intended them, namely as a means for rational and free beings. In short, Crusius turns the argument around and challenges the conservation laws. §419: "For if this were the case [i.e., if this law of the conservation of motion were true], minds could cause no motion; and if said, that a single sum of motion remains in the material world constantly, then no part of the motion of matter could be used for the motion of the substance of minds. But then the material world would be of no use to minds, and it would have been created completely without a purpose." See §420 for his discussion of the law of the conservation of living forces.

¹²⁹ *Ibid.*, pp. 680–681.

are. . . . Now since finite spirits are thus also impenetrable, just as matter is, matter must yield when there is a sufficiently strong effort in spirits to occupy the place of matter. Similarly, spirits must yield and thus be moved if a sufficiently strong effort is present in matter to move to the place at which they are currently found. Consequently, the capacity to move is comprehensible from the essence of every finite substance.¹³⁰

Crusius's strategy in replying to the dilemma (and thus to the two objections based on it) is to attempt to draw new distinctions within the essence of the mind and the body. On the one hand, for reasons reminiscent of one of Knutzen's arguments for physical influx, he denies that mind and matter are radically distinct, since they both have the power to move (in fact, necessarily) as part of their general essence.¹³¹ As a result of this common power, there is no special difficulty in explaining how they can act on each other, since each one can cause the other to move. On the other hand, he also denies that mind and matter sharing the power to move would entail materialism by turning mind into matter, since minds still have the distinctive capacities of thinking and desiring, active capacities that matter as such does not have, thereby revealing that the mind and the body have specific differences.

Is it possible, however, to attribute both a power to move and a power to think or desire to a simple substance such as the soul? Wolff explicitly denies that the soul could have more than one power (§745) on the grounds that each distinct power would require a distinct substance, and a plurality of powers would thus compromise the simplicity of the soul. Crusius, however, denies that distinct powers would require distinct substances. He argues (§73) that a fundamental power can have only one kind of effect and, in the section on rational psychology, is quite pessimistic even about being able to reduce all of the soul's various effects to a single power of representation (§444), as is necessary for Wolff.¹³² Accordingly, not only does he feel the need to posit distinct powers of motion and thinking, but he also wants to separate the power of thinking from the power of willing. Since Crusius thinks that the soul must have

¹³⁰ Ibid., pp. 685–686.

¹³¹ In this context, Crusius adds that impenetrability is included in the general essence of the mind without noting that impenetrability and the capacity to move another body are not necessarily identical. However, in §402 of the *Entwurf* he does suggest that impenetrability is the ground of moving another substance.

¹³² For discussion of this topic in Wolff and Crusius, see Stefan Heßbrüggen-Walter, "Die Seele und ihre Vermögen: Kants Metaphysik des Mentalen in der 'Kritik der reinen Vernunft,'" Ph.D. diss., Universität Münster, 2000), chapter 3, "Vermögenspsychologie im 18. Jahrhundert."

several powers in order to bring about its various kinds of representations, he faces no principled objection in attributing both the power to move and the power to think or desire to a single substance.

In his *Weg zur Gewißheit und Zuverlässigkeit der menschlichen Erkenntnis* (*Path to the Certainty and Reliability of Human Cognition*), published in 1747, two years after the *Sketch of the Necessary Truths of Reason*, Crusius develops further objections to pre-established harmony and provides a more detailed explanation of how mind and body can interact.¹³³ Like Knutzen, he rejects the idea that accidents could be transferred from the mind to the body, since “ideas are mental activities, which are neither motions nor possible through motions.”¹³⁴ He also notes that ideas are not “a special class of things that would be an intermediary between a substance and an accident.”¹³⁵ Rather, Crusius explains:

A real action cannot, however, consist in the fact that motions in the body preceding the sensation bring about the sensation-idea. For that would be contradictory because there would be more in the effect than in the cause. . . . For just that reason an idea cannot be either the proximate or the sufficient cause of a motion. For it conflicts with our understanding of an idea, because an idea is only an activity through which something is represented in the understanding, but not something from which an effect of the soul outside itself is considered possible. Consequently, either a motion must be only a condition upon whose presence an idea arises by means of a mental force whose efficacy, however, is tied to the motion, or the motion must arise as a byproduct from the efficacy of such a mental force that is awakened at the same time through the liveliness of another mental force as through its condition.¹³⁶

Crusius explains his model of mind-body interaction in more detail as follows: “The motion in the instruments of the external senses causes a motion of the substance of the soul. And this motion of the substance of the soul has been made by God, by means of certain laws of actions in nature, into the condition under which certain mental forces, which are the true efficacious causes of representations, become lively and efficacious.”¹³⁷ Crusius sees the need for a motion of the substance of the soul

¹³³ Christian August Crusius, *Weg zur Gewißheit und Zuverlässigkeit der menschlichen Erkenntnis*, reprinted in Crusius, *Die philosophischen Hauptwerke*, ed. G. Tonelli, vol. 3 (1747; rpt. Hildesheim: Georg Olms Verlag, 1964). At §§71–76, pp. 124–136, Crusius develops a series of further objections to pre-established harmony showing both a priori and a posteriori that pre-established harmony is not true, nor probable, nor possible.

¹³⁴ *Ibid.*, §77, p. 140.

¹³⁵ *Ibid.*, pp. 140–141.

¹³⁶ *Ibid.*, §79, pp. 144–145.

¹³⁷ *Ibid.*, §80, p. 145.

because “a motion can bring about nothing other than another motion. Consequently, in this way nothing other than a motion of the substance of the soul can be caused which can thus be only the condition of the sensation-idea that arises.”¹³⁸ Further, “with this motion [of the substance of the soul] the matter that immediately surrounds the soul, which are presumably the life spirits, must be able to yield easily and move out of the way.”¹³⁹ In short, motions can cause only motions so that if the body is to act on the soul, then the soul must be in motion. The soul’s motion is then in some divinely inspired way a necessary condition for its mental activity. In this way Crusius develops an account of interaction between the mind and the body that is consistent with the distinctive principles of his larger metaphysical system.¹⁴⁰

But note, in conclusion, how Crusius’s position compares with Knutzen’s. On the one hand, there is a fundamental and overwhelming difference between the two that can hardly be overlooked. Knutzen’s arguments for physical influx are developed on the basis of principles that a Leibnizian might accept (since Knutzen believes that ultimately only a physical monadology is capable of explaining the properties of bodies), whereas Crusius, as a sophisticated and creative Pietist, intentionally develops his basic position in fundamental opposition to Leibnizian-Wolffian philosophy (and, in the course of doing so, expresses no sympathy with any sort of monadology, whether metaphysical or physical). On the other hand, there is an important resemblance between Knutzen’s and Crusius’s cases for physical influx. Despite all of their differences, Crusius and Knutzen both ultimately think that finite substances must be able to cause changes in motion and agree that such a causal power effectively refutes pre-established harmony and establishes physical influx.

CONCLUSION

To make sense of Kant’s own reflections on a given issue throughout his career, it is clear that one must also understand the background to his views. To this end, we initially raised the following questions: What issues were considered important at the time? What philosophical options

¹³⁸ *Ibid.*, p. 146.

¹³⁹ *Ibid.*

¹⁴⁰ This solution may have significant costs. For instance, the soul would seem to be material in some sense, since it moves, and the problem of mind-body interaction is simply pushed back into the depths of the soul, because it remains unclear how the motions of the substance of the soul can be related to the soul’s mental activities.

were thought to be open on those issues? What arguments or objections would one have to contend with? What would count as an important contribution to the debate?

To see what form these questions take in the case of causality, we had to begin with Leibniz's doctrine of pre-established harmony and the arguments and objections he developed in support of it. Part of Leibniz's motivation to accept and develop pre-established harmony stemmed from difficulties that he saw with physical influx and occasionalism. Leibniz argued that physical influx encountered problems explaining (1) how the mind and the body could interact given their lack of homogeneity, (2) how it did not violate relevant conservation laws, and (3) how any two finite substances could act on each other if it was agreed that their accidents could not migrate from one substance to another. He argued that occasionalism was ultimately in no better shape, insofar as (1) it could not provide identity conditions for substances (given that God's action leaves no traces in things and there are no activities in substances that could explain how they could exist both as distinct from God and as numerically identical over time), (2) it was incompatible with human freedom (which Leibniz thought presupposed the activity or spontaneity of substances), and (3) it ran the danger of being committed to perpetual miracles by invoking God as the cause of everything.

Yet Leibniz's acceptance of pre-established harmony also derived from his own metaphysical commitments. Perhaps his main positive argument for pre-established harmony was based on the idea that substances are, by definition, self-sufficient, which he takes to mean that they should suffice causally for all of their own states. But he supplemented this argument with his complete concept theory of substance and the idea that it is more probable than the other two theories (e.g., because it highlights our freedom and can provide a new argument for God's existence). Finally, the doctrine fit well with his other metaphysical commitments, in particular, his idealism (according to which all simple substances are minds) and his distinction between primitive and derivative forces. As a result, Leibniz thought of pre-established harmony as a doctrine that was both well supported and central to his philosophical position.

While German philosophy during the first half of the eighteenth century was clearly concerned with Leibniz, a number of factors complicated his influence. First, many of Leibniz's most detailed statements of his position occurred in private correspondence with leading European intellectuals rather than in publicly accessible academic journals. Second, Germans became much more internally focussed as they grew in

importance and prosperity and became embroiled in a bitter and drawn out battle between proponents of the Enlightenment and Pietists. Third, Leibniz never wrote a systematic philosophy textbook in German that could be used to disseminate his views (a task that was effectively taken over by Christian Wolff). The result of this complicated set of historical factors is that despite widespread agreement (in many circles) about a whole host of doctrines (including, e.g., the principle of contradiction and sufficient reason and the necessity of simple substances that underlie bodies), there were certain issues in Leibniz's "public" view that were left unspecified and that could thus be exploited in different ways by different figures.

The first and arguably most important figure in German philosophy in the three decades after Leibniz's death is Christian Wolff. In addition to the fact that Wolff's stress on the importance of reason and his leadership in the Enlightenment movement led to a central role in the battle with the Pietists and his acceptance of pre-established harmony resulted in his expulsion from Prussia, he articulated a systematic philosophical framework and methodology that was accepted for the greater part of the century. Metaphysics, which he divided into ontology, cosmology, psychology, and theology, was to provide the foundational principles for all remaining disciplines (from ethics to economics to applied mechanics), and the method common to these disciplines dictated that one start with the definition of a certain concept (whether it be *being* for ontology, *the world* for cosmology, *the soul* or psychology, or *God* for theology), and then derive further principles from it by means of reason alone.

Although Wolff accepted pre-established harmony, the systematic framework that he developed allowed him to de-emphasize it by restricting its scope to the mind-body relationship and hence the domain of rational psychology, perhaps rightly anticipating that the doctrine could come under attack by the Pietists. But if one is committed to pre-established harmony only for the mind-body relationship, then there is less reason to accept Leibniz's idealism, the idea that all simple substances are minds, endowed with powers of appetite and perception, and not bodies, endowed with physical powers. Moreover, while Wolff had no objection to Leibniz's monads, he also did not see that Leibniz had in fact given a proof that *all* simple substances were minds, and that at least some simple substances could not be endowed with some other kind of nonmental power, even if one could not specify what it might be. Thus, in some sense, Wolff pursued what one might think of as a conservative strategy by accepting much of Leibniz's basic position in the context of his attempt at laying out

a systematic position, while restricting the scope of one central doctrine and being more cautious about another that might not have adequate proof on its side. However, as we have seen, such a “conservative” strategy ended up allowing for the possibility of what one might think of as radical changes (e.g., the replacement of Leibniz’s trademark doctrine, namely pre-established harmony, with physical influx) and, perhaps more important, a new set of issues and questions.

Martin Knutzen, one of Kant’s teachers in Königsberg, starts with what he takes to be Leibnizian principles, namely that a simple substance that underlies a body both has the power to move itself and is impenetrable. He then argues that the relational character of such physical properties requires (or at least is more plausibly understood in terms of) physical influx rather than pre-established harmony. If physical influx holds for the simple substances that underlie bodies, it would be natural to infer that it would also hold for monads, that is, for simple substances that are endowed with will and intellect. Yet Knutzen provides explicit support for this inference. Since the power to act on others is a perfection – something that is clear from the fact that God has that power and God has only perfections – there is no reason to think that such a power would be incompatible with minds. After all, Knutzen had already argued that this power is not incompatible with the simple substances that underlie bodies, which are less perfect than minds. Moreover, since what distinguishes minds from the simple substances that underlie bodies, namely the intellect and will, are precisely the perfections that elevate the mind above the simple substances that underlie bodies, it is clear that they cannot be what generates an incompatibility with the mind’s ability to act on others.

Knutzen continues his case for physical influx with a move reminiscent of one of Leibniz’s own arguments for pre-established harmony. He argues that physical influx is not only metaphysically possible – the notion of force he invokes to explain its possibility is fundamentally similar to what Leibniz invokes in his own account – but actually more probable than pre-established harmony, since it allows God to set up substances such that they act in the shortest, most natural way (rather than having to arrange independent substances in such complicated ways that they merely appear to interact). Knutzen’s case for physical influx is complete with his responses to Leibniz’s objections to that doctrine. Knutzen can explain the metaphysical possibility of physical influx without having to suggest that accidents might have to migrate, and he can save the law of the conservation of living forces by restricting it to the realm of bodies,

which is all one should expect insofar as Leibniz's derivation of the law depends on the law of inertia, a law that does not hold for minds.

If Knutzen starts with Leibnizian assumptions, how could he arrive at results so contrary to Leibniz's actual position? There are two crucial issues here that play a significant role throughout this period, even if in different ways for different figures. First, Leibniz explicitly asserts that derivative forces, which are directly responsible for the physical properties of bodies (such as their motions and the spaces they occupy), are derivative from the primitive forces of monads. Now, if a derivative force causes the motion of one body and the motion of one body entails the motion of another body, then it follows that this derivative force acts both on itself (or its own body) and on another. But if a primitive force is the cause of this derivative force, then it is understandable that one would infer that a primitive force acts both on itself (or its own body) and on another. To the extent that the "public" Leibniz did not clarify the distinction between primitive and derivative forces in sufficient detail (or in a way that prevented different interpretations of it), one can understand why Knutzen would be tempted to draw such an inference.

Second, though the "public" Leibniz described simple substances as metaphysical points and distinguished them from physical points understood as organic bodies, he did not explain why they could not be thought of as physical points. Even those who thought of themselves as orthodox Leibnizians (e.g., Baumgarten) explicitly endorsed the idea that monads are physical points. After all, insofar as Leibniz emphasized that simple substances must have points of view on the world, it is a small step to move from monads necessarily *having* physical bodies to them *being* such physical entities. But as soon as simple substances are physical points and are therefore in a position to be the seats of physical forces, it is straightforward to infer, as Knutzen did, that substances can act on each other, contrary to what Leibniz held. Thus, Knutzen's arguments draw on ambiguities about the primitive-derivative force distinction and the possibility of physical points, and Knutzen was not alone in attempting to sort out these doctrines in ways that might not agree with Leibniz's own position.

Like Knutzen, neither Baumgarten nor Meier follows Wolff in restricting the issue of pre-established harmony to rational psychology. In fact, Baumgarten's main argument for pre-established harmony turns on the idea that God would create the world that had the greatest amount of harmony among all its substances. Such a world, he thinks, would be governed by pre-established harmony, since according to pre-established harmony, every state of every substance harmonizes with everything else.

But the case of Baumgarten and Meier also illustrates that even those who saw themselves as defending Leibniz could appropriate his views in ways that might not meet with his approval. For example, their understanding of the distinction between real and ideal influence – a distinction that plays a crucial role in one of their new arguments for pre-established harmony – is drawn in terms of whether a substance acts on itself (ideal influence) or not (real influence), which stands in contrast to Leibniz's view, according to which either there is merely a correlation between the states of substances (ideal influence) or also a causal connection (real influence). In fact, Meier goes even farther than Baumgarten insofar as he constructs several arguments in support of pre-established harmony that are based on the "mechanics" of action in ways that go beyond what Leibniz has to say. For his arguments rely on principles such as "When a finite substance acts, its inner state is thereby changed," "a smallest effect must be possible," and "the consequences remain the same when the grounds are the same," principles that Leibniz never explicitly endorses.

While Wolff, Knutzen, Baumgarten, and Meier all work within a broadly Leibnizian framework, exploring different ways in which Leibniz's system could be presented, articulated, defended, and modified, while still remaining true to its spirit, Crusius does not, and the unique viewpoint he advances in subtle and interesting ways from within a broadly Pietistic framework opens up a whole new set of philosophical issues and options in the mid-1740s. For example, Crusius is especially interested in what is required for substances to form a single world. If, following Leibniz, every substance were a world apart from all others, then only the fact that God thinks of them as belonging to a single world makes them belong to the same world. Crusius provides a striking contrast to this picture by arguing that substances must stand in real causal connections in order to have the kind of unity that is required for them to form a single world. And the ground for this connection cannot lie exclusively in God's understanding, as Leibniz would have it, but rather must depend, at least in part, on God's will. Moreover, Crusius thinks that substances belonging to a single world must be connected either by active grounds or by what he introduces under the term "existential grounds." While existential grounds can be understood in terms of the principle of contradiction, active grounds are based on principles that extend beyond the principle of contradiction. In particular, Crusius distinguishes between the principle of sufficient reason and the principle of determining reason in order to allow for a distinction between libertarian free actions and events that are determined in the ordinary course of nature.

In addition to his innovations at the level of general ontological and cosmological issues, Crusius also brings new thoughts to bear on the mind-body problem. On this point, there are certain similarities between Crusius's and Knutzen's positions, since they both restrict the law of the conservation of living forces to bodies and argue for physical influx from the power a substance has to move. However, Crusius draws crucial new distinctions between the general and differential essences of the mind and body such that one can attribute the power to move to a substance in virtue of its general essence, without thereby denying any differences between the differential essences of mind and body, and he also makes clear that in order to make these distinctions one must reject the Leibnizian idea that a substance can have only one fundamental power. Rather, the soul in particular must have both the power to move and the power to desire and perceive. Given this new conceptual machinery, he then presents a detailed description of the mechanics of the causal interaction between mind and body.

By the time Kant finished his studies at the university and began his independent career, it is clear that causality continued to represent a prominent issue in metaphysics and that both pre-established harmony and physical influx were live options for Leibnizians and non-Leibnizians alike. On the one hand, Leibniz's central objections to physical influx had been, or at least could be, resolved, since physical influx could appeal to a notion of force that is similar in significant respects to Leibniz's own and the law of the conservation of motion or living force could apparently be restricted to bodies. On the other hand, the single most significant shift in philosophical position between Leibniz's own view and the set of views that dominated the first half of the eighteenth century in Germany took the form of a rejection of idealism in favor of views that either were simply agnostic about idealism (Wolff) or emphasized the reality of physical properties (Knutzen and Crusius). Since this shift – along with other changes during the period, such as Baumgarten's and Meier's attempts at finding new lines of support for Leibniz's traditional position – drew attention to both the fundamental nature of substances (whether purely mental or also physical) and issues surrounding the explanation of physical properties, questions such as the following came to appear more pressing: What kind of relation must substances have to each other in order to belong to a single world (a real or an ideal relation)? What is the proper distinction between real and ideal relations? What is God's role in the relations between substances that form a single world? How should the primitive-derivative forces distinction be understood?

What kind of causal explanation can be given for the physical properties of bodies, such as motion and impenetrability? How are grounds to be understood, for example, as existential and based on the principle of contradiction or exclusively as active and based on the principle of determining or sufficient reason? It is these questions we must face as we turn to consider Kant's pre-Critical views on causality in Chapter 2.

Kant's Pre-Critical Theory of Causality

INTRODUCTION

From the late 1740s through the 1770s, debate in Germany about pre-established harmony and physical influx lost much of its political significance, in large part because the intellectual tolerance that came with the inauguration of the “philosopher king,” Frederick the Great, in 1740 considerably improved conditions for free and open public debate, so much so that Wolff could be enticed to return to Halle from his refuge in Marburg. But the issue of causality also attracted somewhat less philosophical attention than it had received earlier in the century, either because physical influx had won the debate or because the sides had become so entrenched that it was clear to many that neither side would budge from its antecedently held views. Thus, most German philosophers during this period were not thinking about causality nearly as much as they had been in the 1720s and 1730s.

However, Kant was a prominent exception to the trend, since he focused on the issue of causality throughout his entire pre-Critical period (1746–1770). For some time after leaving the university, at least from around 1746 to 1758, Kant’s intellectual activities were devoted exclusively either to answering scientific questions or to explaining the metaphysical underpinnings of the basic properties of bodies. He published a series of short scientific essays in the mid-1750s, on fire, on the age and rotation of the earth, on earthquakes, and on winds, in addition to his lengthy *Allgemeine Naturgeschichte und Theorie des Himmels* (*Universal Natural History and Theory of the Heavens*) of 1755. And metaphysical principles that would support these scientific views are discussed in his

Gedanken von der wahren Schätzung der lebendigen Kräfte (*Thoughts on the True Estimation of Living Forces*) in 1746–1747, *Principiorum primorum cognitionis metaphysicae nova dilucidatio* (*A New Elucidation of the First Principles of Metaphysical Cognition*, for short, *Nova dilucidatio*) in 1755, and *Metaphysicae cum geometria iunctae usus in philosophia naturali, cuius specimen I. continet monadologiam physicam* (*The Employment in Natural Philosophy of Metaphysics Combined with Geometry, of which Sample I Contains the Physical Monadology*) in 1756.

What unifies these early pre-Critical publications is Kant's interest in developing a sophisticated metaphysical account of causality that would explain various properties of bodies.¹ Thus, Kant attempts to solve the *vis viva* debate in the *True Estimation* by introducing a "metaphysical concept of force" and by characterizing force not in terms of motion, but rather abstractly in terms of an ability to act on other substances, without specifying the particular kinds of effects it might have. Although Kant does not present any independent argument for this concept or for the theory of causality in which it is embedded, he boldly states that it can solve the mind-body problem and "make the triumph of physical influx over pre-established harmony complete" (1:21). The *Physical Monadology* likewise attempts to explain physical properties on the basis of metaphysical concepts, though in this case it is the idea that a force's activity must be understood as relational if one is to reconcile the infinite divisibility of space with the unity of monads, since the intrinsic activity of a monad that is governed by pre-established harmony does not admit of division in the way that space does.

Without downplaying the distinctive contributions made by the *True Estimation* and the *Physical Monadology*, it is clear that Kant's most focused and developed thoughts on causality in the early years of his pre-Critical period are to be found in the *Nova dilucidatio*. In the *Nova dilucidatio* Kant "corrects" Wolff's principles of contradiction and sufficient reason and then supplements them with two principles of his own. In the case of the first principle, the principle of succession, Kant provides an explicit argument against pre-established harmony and for physical influx where the *True Estimation* had simply presupposed the truth of the latter. This argument clarifies what his notion of a ground is and how he may be justified

¹ I am thus in substantial agreement with the general theses of Martin Schönfeld, *The Philosophy of the Young Kant: The Precritical Project* (New York: Oxford University Press, 2000), and Alison Laywine, *Kant's Early Metaphysics and the Origins of the Critical Philosophy*, North American Kant Society Studies in Philosophy, vol. 3 (Atascadero, Calif.: Ridgeview, 1993).

in understanding grounds in this way, given Wolff's, Baumgarten's, and Meier's treatments of that notion, even if it diverges from Leibniz's own conception in significant respects. Additionally, by keeping in mind that Kant is thinking primarily of the case of bodies in motion, his positive model of causality can be reconstructed on the basis of his explicit arguments. In the case of the *Nova dilucidatio's* second principle, the principle of coexistence, Kant presents an argument against Crusius's notion of existential grounds by examining what does and does not follow from the "mere existence" of a substance. It is in this context that Kant provides a detailed analysis of God's role in the causal relations between substances.

In his later pre-Critical publications, *Reflexionen*, and lecture transcripts – from 1762 to 1770 – we see Kant undertaking significant revisions of different aspects of this early theory. Specifically, it is against the account of causality that he had developed in the *Nova dilucidatio* that we can understand how Hume's position became important to him. For what Kant found striking about Hume's position was not that Hume adopted a skeptical position on the issue of causality. After all, Hume's skepticism was based on exclusively empiricist principles that Kant was never even tempted to accept. Rather, Kant saw that Hume's argument raised a fundamental challenge to the theory of causality he had been developing earlier in his pre-Critical period. Specifically, Hume helped Kant to see that, as a proponent of physical influx, he could not understand grounds as purely logical (as Wolff and Baumgarten had). As a result, he introduced the notion of a "real ground" and attempted to work out its consequences in *The Only Possible Argument in Support of a Demonstration of the Existence of God* (1763), the *Negative Magnitudes* (1763), and various *Reflexionen* of the period. In his Inaugural Dissertation in 1770, Kant revised his account of causality further by adding sophistication to his arguments for physical influx and by transforming it into a fundamental principle of his philosophical system, which gave it an even greater systematic prominence within his overall position. As a result, Kant was consistently working on formulating an adequate account of causality throughout his pre-Critical period.

The first three sections of this chapter thus treat Kant's early pre-Critical views on causality. In a first section, we investigate the metaphysical conception of causality that is either implicitly or explicitly involved in Kant's discussion of scientific issues in the *True Estimation* and the *Physical Monadology*. In the next two sections we turn to a detailed discussion of the *Nova dilucidatio's* principles of succession and coexistence. The last two sections of this chapter then consider the importance of Hume to

Kant in the early 1760s and new developments toward the end of Kant's pre-Critical period in his Inaugural Dissertation. This will, I believe, make it clear how Kant is reacting to the views of his predecessors as described in Chapter 1, and also what the backdrop is against which Kant's account of causality in the Critical period can be understood.

KANT'S CONCEPT OF FORCE IN THE *TRUE ESTIMATION* AND
PHYSICAL MONADOLGY

Kant begins the *True Estimation* with reflections on the "metaphysical concept" (1:17) of force and its relationship to motion. In §§1–2, Kant agrees with Leibniz that a body has an essential force that can be characterized as active and that inheres in it even prior to extension, criticizing those – "Wolffians" are explicitly mentioned – who would characterize force exclusively as the cause of motion. For explaining motion by means of a *vis motrix*, or moving force, is, Kant thinks, vacuous in just the way that Scholastics invoking *vis calorifica* to explain the presence of heat is. Kant attempts to loosen the connection between force and motion further by arguing (§3) that even bodies at rest can be active insofar as they are merely attempting to move (e.g., a ball resting on a table is acting on the table even if it is not moving). But if force is not to be understood in terms of what may be its most common effect, namely motion, how should it be characterized? Kant suggests that force be understood more generally as essentially active, that is, as that which acts on substances without specifying either the nature of the substances acted on or the kinds of effects that might be brought about in them.²

Despite the fact that the connection between force and motion is not so close that one can immediately (and vacuously) deduce motion (as an effect) from a force's action (as its cause), Kant nonetheless argues (in §4) that "nothing is easier" (1:19) than to explain motion with his concept of an essential active force.

Substance A, whose force is determined to act externally (that is, to change the internal state of other substances), either immediately encounters an object which receives its entire force at the first moment of its endeavor, or it does not encounter such an object. If the former took place with all substances, then we would not become acquainted with any motion whatsoever, nor, in consequence,

² Although Kant does not explicitly note the point in these initial sections (§§1–2), it is clear that this fundamental concept of force commits him to physical influx, even if he has not yet given any direct argument for it.

would we name the force of bodies after it. But if substance A cannot exert its entire force at the moment of its endeavor, then it will exert only part of it. But the substance cannot remain inactive with the remaining part of its force. Rather, it must act with its entire force, for otherwise it would cease to go by the name of force when not exerted in its entirety. Because the consequences of this exertion cannot be found in the coexistent state of the world, one must therefore locate them in the world's second dimension, namely, in the succession of things. That is why the body will not exert its force all at once, but will do so only gradually. However, in the succeeding moments it cannot act on the very same substances on which it acted right at the start, for these receive only the first part of its force and are incapable of receiving the rest. Thus, body A gradually acts on ever different substances. Substance C, however, on which A acts at the second moment, must have an entirely different relation of location and position with respect to A than B does, the substance on which A acted initially. For otherwise there would be no reason why A should not initially have acted all at once on both substance C and substance B. In the same way, each of the substances on which A acts in subsequent moments has a different position with respect to the initial location of body A. That is, A changes its location in acting successively. (1:19)

Kant's basic idea here is that if one substance were to act (or exert its force) on another all at once, then there would be no motion at all.³ Since there is motion, it is clear that a substance must act on other substances in succession, that is, it must exert only part of its force on any given substance at any given moment, leaving another part of its force for another substance at a later moment. It is not enough, however, for one substance to act on others in succession. The other substances must have changed their positions with respect to that substance, for, so Kant thinks, if the substances had not changed their position, there would be no reason why the first substance did not act on them with all its force in the first instance.

This argument is problematic in several respects. For example, even if one grants that substance A must exert its force on another substance and over time, it is unclear why it must act on different substances and cannot simply act on the same substance at different times. His explicit justification is that the second substance, substance B, would be "incapable of receiving the rest" of substance A's force, but no reason is given for this claim. The fact that B was incapable of receiving it at the first moment in time does not obviously and immediately imply that it would be

³ As we see below, in the *Nova dilucidatio* Kant provides the following explicit justification for this claim: If a substance exerted its force all at once, it would thereby have posited all of the properties that it would ever be capable of positing in other substances and thus be unable to posit any new ones, which, however, is what is required in order to cause a change in location.

incapable of receiving it at a later moment in time. Also, it is unclear why substances B and C must change their positions for substance A to act on them. Kant attempts to support the claim by arguing that if B and C had not changed their positions, then A could have acted on both of them at the first moment of time. But it is difficult to see that this point is at all relevant to his argument insofar as he does not explicitly state that the activity of a substance is subject to any location conditions.⁴

Whether or not the argument can be defended on its own terms, it is helpful to understand it in terms of motion that is initiated in collisions between impenetrable bodies (although it must also be applicable to the other kind of motion Kant countenances in the *True Estimation*, namely bodies moving of their own accord). Accordingly, if substance A is an impenetrable substance, it would first exert part of its force on substance B by pushing it away, and then exert another part of its force by pushing C away, thereby generating motion. Because A pushes B away, it would be incapable of acting on it any more and would thus have to act on C in order to continue acting (as it must if active force is to be essential to it). This would solve the first problem. Moreover, because impenetrability is a contact force, it becomes more plausible to think that B and C would have to change their locations in order for A to act on them successively, which solves the second difficulty. By interpreting the argument as applying to impenetrable bodies, the argument thus appears somewhat more intelligible. At the same time, this argument still makes a series of controversial assumptions about forces, grounds, and substances that Kant does not make explicit. Since a closely related argument is developed at greater length in the *Nova dilucidatio*, we have reason to return to it in the next section.

In the *True Estimation*, Kant next turns (in §§5–6) to the mind-body problem, suggesting that his new conception of force can be incorporated into a version of physical influx that would explain how matter produces representations in the mind and vice versa. As he puts it, his new metaphysical concept of force allows the mind-body problem to

disappear . . . and more than a little light is shed upon physical influx, when the force of matter is ascribed not to motion, but rather to its actions upon other substances that need not be defined further. For the question of whether the soul can cause motions – that is, whether it has motive force – is transformed into

⁴ There are many other worries that one might have about this argument. For example, is it appropriate to think of bodies as having forces that they can use up (as is implied by expressions such as “receiving its entire force”)? Also, how does the argument establish that substance A must move, rather than that substances B and C must move?

the question of whether its essential force is directed to act externally, that is, whether it is capable of acting outside itself on other entities and of producing changes. One can answer this question quite decisively by saying that the soul must be able to act externally by reason of the fact that it is in a specific location. For when we analyze the concept of what we call location, we find that it suggests the actions of substances upon each other. All that kept a certain acute author from making the triumph of physical influx over pre-established harmony complete was nothing more than this little confusion of concepts, a confusion that is easily overcome as soon as one's attention is turned to it. (1:20–21)

Kant thus thinks that the mind-body problem can be solved by accepting the notion of active force that he has just introduced and by making the further assumption that having a spatial location is derivative from the interaction between substances endowed with such forces (an assumption he supports with arguments in §7 and §9 of the *True Estimation*). As we saw in Chapter 1, the problem Kant is addressing here is that the mind and the body would appear to be so heterogeneous that the kind of intelligible connection purportedly required by causal interaction is impossible for them. The primary basis for the heterogeneity claim lies in the difference between the mind's and the body's powers. The mind has the power to think, whereas the body has the power to move. Kant holds that characterizing force more abstractly as active rather than in terms of motion solves the problem because it shows how to understand force in such a way that there is no heterogeneity between the mind and the body at the relevant level. The soul must exercise its active power because it is in a location and locations are possible only due to the interaction of forces that creates space in the first place. Likewise, the body can act on the mind insofar as it "acts on everything spatially connected with it, and hence also on the soul; that is, it changes the internal state of the soul insofar as this state is related to what is external to it . . . [which] goes by the name of *status representativus universi*" (1:21). As a result, the body can act on the soul so as to change its internal (representational) state just as the soul can act on the body, and therefore both are able to act externally on each other.⁵

Kant's solution to the problem of the heterogeneity of mind and body is thus interestingly different from both Crusius's and Knutzen's views. As we saw at the end of Chapter 1, Crusius solves the heterogeneity problem

⁵ For an interesting discussion of the context of Kant's *True Estimation*, see Manfred Kuehn, "Kant and His Teachers in the Exact Sciences," in *Kant and the Sciences*, ed. E. Watkins (New York: Oxford University Press, 2001), esp. pp. 21–27. Kuehn argues that while Kant is referring to Knutzen in this passage, the reference is to be understood as a sarcastic put-down, not as a compliment.

by identifying a specific force that both minds and bodies have, namely the force to move. Similarly, Knutzen starts with the idea that a simple substance, such as the soul, must have the power to move itself (or its body) and then argues that it must therefore also have the power to move other simple substances (or the bodies they compose or result in), though he also expresses a certain degree of skepticism about our ability to understand the specific causal mechanisms that pertain to mind-body interaction in particular and thus about a *detailed* solution to the heterogeneity problem.⁶ Kant, by contrast, defines force more generally so that he need not attribute the same particular forces to the mind and the body. Rather, it is enough if whichever particular forces that the mind and the body might have are still forces in the general sense of being able to act on others. As a result, Kant can “solve” the mind-body problem by being committed merely to souls having the power to interact with other (bodily) substances, which he thinks is justified insofar as that is required for the soul to *have a location*, even if it may still fall short of the power to *move* bodies.⁷

Kant devotes much of the rest of the *True Estimation* to presenting detailed criticisms of Leibnizian and Cartesian positions on the measurement of force in the hopes of putting himself in a position to chart a middle course between what he perceives to be their one-sided views on what quantity is conserved in nature. Abstracting from all the details of his argument, however, there is one fundamental issue in this work that is directly relevant to Kant’s views on causality. To split the difference between the Leibnizian and Cartesian positions on what quantity of force is conserved in the world, Kant distinguishes between motions that require a constant external cause (dead force or pressure) and those that derive from within a substance (living forces). According to this account, the motion of a projectile (such as that of a bullet fired from a gun) is caused by a living force within the bullet, which is to be measured by mv^2 . In such cases, Kant thinks that a substance has the power to change its own state, without the influence of any other substance, and thereby implicitly

⁶ In §42 of the *Systema causarum* Knutzen argues for the following claim: “*The very specific mode of physical influx or action by which the mind influences its body and the body in turn influences the mind, cannot be distinctly understood by us on account of the nature of the thing*” (pp. 139–140).

⁷ See Friedhelm Nierhaus, “Das Problem des psychophysischen Kommerziums in der Entwicklung der Kantischen Philosophie,” Ph.D. diss., Universität Köln, 1962, and Andrew Carpenter, “Kant’s First Solution to the Mind/Body Problem,” in *Kant und die Berliner Aufklärung*, ed. V. Gerhardt, R. Horstmann, and R. Schumacher (Berlin: De Gruyter, 2001), vol. 2, pp. 3–12, for discussions of the issue.

rejects Newton's force of inertia. However, sometime prior to 1755, Kant changes his mind on this crucial point. At least for matter, he comes to accept Newtonian forces of attraction and repulsion and explicitly argues against the possibility that a substance could change its own state without being acted on by another.⁸

Despite the fact that Kant thus replaces a quasi-Leibnizian with a quasi-Newtonian physics in 1755, he continues to adhere to and refine his notion of an active force. Thus in the *Physical Monadology* (1756) his main aim is to reconcile the unity demanded in metaphysics with the infinite divisibility of space required by geometry. Like Knutzen, he starts with physical monads, that is, with simple substances that are in a place and compose extended bodies without being themselves extended, and then proceeds to argue for the necessity of Newtonian forces of attraction and repulsion (in addition to a force of inertia that he identifies with mass) in order to make sense of various properties of bodies (e.g., contact, density, determinate volume, and elasticity). The issue Kant perceives to be most pressing, however, lies in understanding how extended bodies could be composed of unextended monads without the divisibility of the former threatening the destruction of the latter's essential unity. Kant's solution invokes a metaphysical concept of active force that is continuous with the one he had introduced earlier in the *True Estimation*. In particular, he argues that a monad fills a determinate space due to "the sphere of the activity by means of which it hinders the things that are external to it and present to it on both sides from drawing any closer to each other" (1:480). One substance can thus act on another that is external to it by resisting it, or by keeping it from coming any closer to the substance that is on its other side. Accordingly, although a physical monad is an unextended point in space, through its activity it can be present in an extended region by keeping out other monads (or the presence they have by means of their sphere of activity).

How is it that this notion of a sphere of activity can allow for the divisibility of space while retaining the indivisibility of the point that is present throughout a given space? Kant explicitly addresses this question as follows:

But, you say, substance is to be found in this little space and is everywhere present within it; so, if one divides space, does not one divide substance? I answer: This

⁸ For a discussion of many details of Kant's *True Estimation*, as well as of his conversion to Newtonianism, see Martin Schönfeld, *The Philosophy of the Young Kant: The Precritical Project*, chaps. 1–3.

space itself is the orbit of the external presence of its element. Accordingly, if one divides space, one divides the extensive quantity of its presence. But, in addition to external presence, that is to say, in addition to the relational determinations of substance, there are other, internal determinations. If the latter did not exist, the former would have no subject in which to inhere. But the internal determinations are not in space, precisely because they are internal. Accordingly, they are not themselves divided by the division of the external determinations. (1:481)

Perhaps surprisingly, what is crucial to Kant's official answer is not so much his notion of activity per se, but rather the way in which it is combined with a claim about the connection between intrinsic and relational properties. According to Kant, the division of space does not imply the division of the monad that fills that space, because the division of space is the division of a relational property, which need not affect the intrinsic properties of a monad. For at this point in his career Kant holds that even if relational properties require intrinsic properties (because relational properties stand in need of subjects to serve as their relata), changes in relational properties do not necessarily imply any changes in a monad's intrinsic properties.⁹ In this way, the relational properties that constitute the "external sphere" of a monad's activity can be divided without the monad's intrinsic properties being changed in any way.

There is, however, more to Kant's official solution than meets the eye. In the course of his discussion, Kant raises an objection that "derives from the positing outside each other of the determinations of one and the same substance. For the action of the monad which is in space BCD [one half of the sphere carved out by a monad's activity] is external to the action which is in space BDA [the other half of its sphere of activity]. They thus seem to be really different from each other and to be found outside the substance" (1:482). Kant's succinct reply – "But relations are always both outside each other and outside substance" – merely reiterates his original solution and thus fails to address what really motivates the objection.

To understand the real point of the objection and what aspect of Kant's position is crucial to addressing it, recall Wolff's and Baumgarten's position on substance. They explicitly equate substance with activity, so that if one can distinguish two activities at a time, then there must be two substances, not one. But since one can divide the space that a physical monad is supposed to be present in by virtue of its sphere of activity, it would

⁹ Rae Langton, *Kantian Humility* (New York: Oxford University Press, 1998), p. 102, makes this point as well.

seem that two separate activities must be ascribed to a single monad, with one activity per region of space. The fact that Kant distinguishes between relational and internal determinations and requires the latter for the former cannot be used to reply to the objection, because the activities or determinations that are internal to substances and thus indivisible are irrelevant as long as the relational determinations that generate space are divisible. Finally, it is really beside the point for Kant to insist that relations are outside substances. Wolff and Baumgarten can agree with this claim. The real disagreement, which Kant does not explicitly acknowledge here as such, concerns whether the activity that is to account for the spatial presence of a monad should be understood in terms of extrinsic relations or internal/intrinsic properties. Wolff and Baumgarten, as proponents of pre-established harmony, must hold that (causal) activities are intrinsic, whereas Kant, as a proponent of physical influx, thinks that they must be extrinsic relations.

Accordingly, one can see the point of the *Physical Monadology* as follows. In attempting to explain how monads constitute, contribute to, or result in the properties of extended bodies, Kant follows Knutzen's general strategy of arguing that accepting physical influx makes this task much easier (or perhaps possible in the first place). For if one accepts pre-established harmony and thus the claim that a monad's activities are internal/intrinsic rather than external/extrinsic (or relational), then one cannot straightforwardly appeal to these activities to explain the spatiality of bodies. For spatiality implies divisibility, and if the activities constituted spatiality, then they would be divisible too, which is impossible if the activities are intrinsic to the substance. As a result, the point of the *Physical Monadology* is to establish that one must accept causal relations between substances, that is, physical influx, in order to explain in any detail the relationship between monads (which, according to agreed-on metaphysical principles, possess unity essentially) and the spatiality of bodies (which necessarily entail infinite divisibility, according to the laws of geometry) in terms of activities.

A Leibnizian is not, however, without possible resources here. Perhaps a distinction could be drawn between a monad's activities and the effects of those activities such that the effects of such activities would be spatial and hence divisible even if the activities themselves were not. Or perhaps spatiality (along with the properties that derive from it, such as divisibility) could be explained on the basis of the confusion inherent in the representations of finite monads. Accordingly, Kant's argument in the *Physical Monadology* is not necessarily decisive on this issue.

Even so, it is important to see what Kant is attempting to do. Like Knutzen, Crusius, and others at the time, he is trying to articulate in a detailed way how the spatial and physical properties of bodies are supposed to be derived from (metaphysical) forces. In particular, his account goes beyond merely claiming that the one is derivative from, or is a limitation of, the other. Instead, he appeals to the (metaphysical) notion of force that he had introduced in the *True Estimation* and expanded on in the *Physical Monadology*, according to which force is to be understood not in terms of motion, but rather in terms of a substance's ability to act on other substances. Accordingly, while Kant discusses questions in the *True Estimation* and the *Physical Monadology* that grow naturally out of the historical context described in Chapter 1 – his focus is on how to understand an account of the physical properties of (physical) monads – what is distinctive about Kant's answers, that is, what he and others at the time would have seen as his unique contribution to these issues, is the way in which he attempts to set straight certain details in metaphysics, namely, how forces should be understood in terms of relational activities that have a sphere of influence. Kant's initial (though by no means exclusive) focus in his pre-Critical period is thus on metaphysical aspects of causality.

THE NOVA DILUCIDATIO AND THE PRINCIPLE OF SUCCESSION

Kant presents his most detailed metaphysical account of causality in the pre-Critical period in the *Nova dilucidatio* (in 1755, one year prior to the *Physical Monadology*). For one, he provides an explicit argument for physical influx where the *True Estimation* had simply presupposed its truth in the form of his novel conception of force. Yet he also develops more clearly the metaphysical framework that underlies his model of causality by explaining how change is possible, by exploiting a certain notion of what a ground is and how a ground posits determinations, by examining what does and does not follow from the "mere existence" of a substance, and by providing a detailed analysis of God's role in the causal relations between substances. Accordingly, the *Nova dilucidatio* deserves sustained attention.

While the bulk of Kant's argument in the *Nova dilucidatio* is devoted to clarifying the status and various consequences of the principles of contradiction and sufficient reason or, as he prefers, "determining ground," in the third section Kant turns to presenting two causal principles that are "extremely rich in consequences and derive from the

principle of the determining ground" (1:410), namely the principles of succession and coexistence.¹⁰ After first stating the principle of succession, Kant provides three separate proofs of the principle and offers an elucidation of these proofs as well as four of its applications, before concluding with a scholium.

Is Pre-established Harmony Consistent with Change?

The principle of succession explicitly aims to establish physical influx by refuting pre-established harmony. It states: "Substances can change only insofar as they are connected with other substances; their reciprocal dependence determines the mutual change of state" (1:410). Kant provides three arguments for this principle, all based on the idea that the kind of causally isolated substances invoked in pre-established harmony are incapable of undergoing change given the way in which determinations are posited by grounds in a substance.¹¹

¹⁰ In the first section of the *Nova dilucidatio*, Kant argues that what is commonly referred to as the principle of identity is not an absolutely first principle, because it is based on two more primitive principles, namely that whatever is, and whatever is not is not. These two principles are more properly thought to be fundamental because they employ the simplest and most general terms out of which more complex and determinate principles can arise. Moreover, Kant criticizes the principle of contradiction (when formulated as "it is impossible that the same thing should simultaneously be and not be") on the grounds (1) that it would seem to be merely a definition of impossibility, which lacks any proof that all truths must be established by reference to this definition rather than to some other principle and (2) that it is "considerably worse even than a paradox" (1:391) to make a negative rather than a positive proposition the fundamental principle of all truths.

In the second section of the *Nova dilucidatio* Kant provides a detailed discussion of his understanding of his concept of "ground." He begins by distinguishing between "antecedently determining" and "consequentially determining" grounds, which is a version of the distinction between metaphysical and epistemological grounds (the ground of becoming or being vs. the ground of knowing). He then criticizes Wolff's definition (according to which a ground is that by reference to which it is possible to understand why something should rather be than not be) on grounds of circularity (since any explanation of the phrase "why" in the definition would tacitly involve the concept of a ground (1:393)) and agrees with Crusius that "determining ground" is preferable to "sufficient ground," given that it can be unclear how much is sufficient in any given case. He also puts his concept of ground to significant use by demonstrating the principle of sufficient reason, proving God's existence, establishing the compatibility of divine foreknowledge with freedom, solving the problem of freedom and determinism, and refuting the principle of the identity of indiscernibles.

¹¹ Kant's reason for rejecting Leibniz's position is thus not what *he* asserts is the typical one, namely its use of final causes (which Kant agrees are unfitting for God and generally of little help), but rather its "internal impossibility," that is, the fact that change would be impossible according to pre-established harmony if the arguments to be presented above

Three Arguments

Kant states his first argument as follows:

Suppose that some simple substance . . . were to exist in isolation. I maintain that it could undergo no change of its inner state. The inner determinations, which already belong to the substance, are posited in virtue of inner grounds, which exclude the opposite. Accordingly, if you want another determination to follow, you must also posit another ground. But since . . . no external ground is added to it, it is patently obvious that the new determination cannot be introduced. (1:410)

The main thrust of this argument is that any causally isolated substance cannot change because change would require a new determination and thus a new ground, but such a new ground is nowhere to be found, given that the isolation of the substance rules out external grounds and all of its internal grounds have already been posited.¹² Kant's argument proceeds from what he takes to be an analytic claim about change along with the principle of determining reason. For a substance to change, it must lose one of its determinations and gain another that is incompatible with the first determination. That is the analytic truth about change.¹³ If the principle of determining reason states that there must be a reason or ground for positing any determination, then it follows that there must be one ground for the initial determination of a substance and another ground for its later determination. The ground for the initial determination must be internal to the substance, since positing a substance entails positing the grounds that compose it. That is, one cannot posit a substance without also positing the grounds that are essential to it. What, then, is the ground of the later determination? Once again, it cannot be found in

are sound. It is also worth stressing just how strong Kant's claim is. He is claiming not merely that Leibniz's position is wrong, as a matter of fact, but rather that it is incoherent, unintelligible, or impossible. The burden of proof on Kant is thus considerable.

¹² While a ground in general establishes "a connection and a conjunction between the subject and some predicate or other" (1:392), Kant's basic definition of a determining ground is as follows: "A ground, therefore, converts things which are indeterminate into things which are determinate" (1:392). Thus, a determining ground posits that a determination, predicate, or property inheres in a subject or substance.

¹³ The argument presupposes the traditional idea that change cannot be understood as simply the addition or subtraction of a determination, but rather must be conceived of as the replacement of one determination by a contradictory determination. Such a presupposition is entailed by Leibniz's complete concept theory of substance, since it requires an opposition between the existing determination and the new determination and an opposition can be guaranteed only if one of every pair of contradictory determinations must belong to the substance. As a result, since Kant's argument is directed against a Leibnizian position, the presupposition seems warranted.

any other substance, since the substance is causally isolated. Nor can it be found within that very substance, since that substance already contains the opposite ground (i.e., a ground that posits the predicate that excludes this determination). Thus, there can be no new ground and thus no new determination, which implies that no change is possible.¹⁴

Kant develops a second argument as follows:

It is necessary that whatever is posited by a determining ground be posited simultaneously with that determining ground. For, having posited the determining ground, it would be absurd if that which was determined by the determining ground were not posited as well. Thus, whatever determining factors exist in some state of a simple substance, it is necessary that all factors whatever which are determined should exist simultaneously with those determining factors. But since a change is the succession of determinations . . . , it follows that the change cannot take place by means of those factors that are to be found within the substance. If, therefore, a change occurs, it must be the case that it arises from an external connection. (1:411)

This argument attempts to refute the possibility that change in a causally isolated substance's determinations could be due to a change in its grounds. What motivates the possibility Kant is attempting to refute is the idea that one must be able to apply the principle of determining reason to each and every determination that a substance has. Accordingly, if a substance has determination F at t_1 and determination G at t_2 , then there must be a ground, a , for F as well as a ground, b , for G . Moreover, because Kant explicitly asserts that grounds and the determinations they posit must be simultaneous (on pain of absurdity), grounds a and b are such that at t_1 a grounds F and at t_2 b grounds G . If one accepts the

¹⁴ Langton, *Kantian Humility*, discusses only the first argument and suggests that it "presents many causes for philosophical disquiet" (p. 105). In particular, she objects: "The assumption that for every intrinsic property there is another intrinsic property that is the 'reason' for the first seems to imply an infinite regress. And what of the apparent counterexamples? An alarm clock can be set (at its creation if need be) to ring at six, with no outside interference" (ibid.). The question of counterexamples is not easily settled, as Langton herself notes, because of ambiguity about what a single substance would be. However, her first objection is beside the point, since Kant nowhere assumes that a ground is an intrinsic property, and thus no infinite regress looms. Thus, neither of Langton's objections to the argument is as serious as she seems to think. At the same time, Langton intentionally does not focus on the argument ("I leave these causes for disquiet, for the chief interest of the argument lies elsewhere" (ibid.)). Yet given that Langton wants to focus on what follows from the principle (namely that knowledge requires receptivity), it is difficult to see that the principle and the arguments Kant develops to support it would not be crucial to her own interests.

principle of determining reason in this form, then it seems to follow that a change in determinations would have to be due to a change in grounds.

However, Kant thinks that explaining a change in determinations by means of a change in grounds is inadequate for two reasons. First, Kant identifies the essence of a thing with its necessary grounds. Accordingly, if the essence of a thing is immutable, then its grounds will be immutable as well. But if a ground is necessarily simultaneous with its determinations, then whatever determinations follow from the immutable grounds of a thing's essence will be immutable as well. As a result, if a causally isolated substance exists, it must do so with the immutable grounds that constitute its essence, but they, in turn, must simultaneously posit their determinations, which precludes the possibility that those determinations could change. While one might try to distinguish between essential and inessential grounds such that the latter might change while the former remain constant, Kant thinks that no explanation of the addition of inessential grounds to the essential grounds of the substance will be available. What could the source of such grounds be? The essential grounds are immutable and thus would be incapable of adding anything that would not be immutable, and, given the isolation dictated by pre-established harmony, no external substance could add any grounds. There is thus nothing else to the substance beyond its essential grounds and the determinations that follow from them that could explain change.

Second, Kant could reject the idea that changing determinations must be explained on the basis of a change in grounds for the simple reason that insofar as he is attempting to explain change, appealing to changing grounds would simply push the problem back one stage. For if grounds were to change and if the principle of determining reason were applied to grounds and not merely to determinations, one would want to know what the reason was for the change in grounds and one would face the very same challenge there as one faced in explaining a change in determinations. For these reasons, Kant rejects the idea that change could be explained on the basis of changing grounds.

Kant's third argument for the principle of succession asserts that it is equally absurd to think that *unchanging* grounds within a causally isolated substance could be responsible for change. As Kant explains: "Suppose that a change takes place under the conditions specified. Since . . . no grounds, apart from those which are internal, are supposed to be involved in determining the substance from any other source, it follows that the same grounds, by which the substance is supposed to be determined in a certain way, will determine it to the opposite, and that is absurd" (1:411).

In short, Kant's idea is that one and the same set of grounds cannot posit first one and then a contrary set of determinations.

Consider two possible ways of attempting to explain changing determinations within an isolated substance on the basis of one and the same set of grounds. One might first suggest that ground *a* posits first determination *F* and then determination *G* (which is incompatible with *F*). However, it is unclear in this scenario that an intelligible explanation can be provided for the change in determinations. After all, no change has occurred in ground *a* that would explain the change in determinations, and any change external to the substance is necessarily irrelevant, given that the substance in question is supposed to be causally isolated. Moreover, if one accepts (as Kant clearly does) the principle that a ground is simultaneous with whatever determinations it posits, then ground *a* would have to posit *F* and *G* simultaneously, which is clearly contradictory.

Alternately, one might suggest that there are two grounds, ground *a* and ground *b*, such that ground *a* posits determination *F* and then ground *b* posits determination *G*. However, Kant can argue, once again, that no intelligible explanation of the change has been provided, because the change in determinations would be due not to changing grounds, but rather to the fact that first one ground posits one determination and then a second ground posits another, and no explanation of why the grounds alternate in being effective is available. External grounds are irrelevant by stipulation, and, as we just saw, positing an internal ground to explain the change in the efficacy of grounds simply pushes the problem back one stage. For in that case one is faced with explaining a change (now in the efficacy of grounds rather than grounds or determinations) on the basis of either an unchanging ground (which was the task originally posed) or changing grounds (which would require further explanation). Moreover, if grounds are simultaneous with their determinations, it would be impossible for the second ground to "wait" until the first ground is done in order to posit its "new" determination. Since neither of these ways of attempting to explain a change of properties within an isolated substance on the basis of one and the same set of grounds seems tenable to Kant, he infers that change within a causally isolated substance is impossible, that is, that pre-established harmony is false and that there must be causal connections between substances (physical influx) in order for change to be possible.

After presenting these arguments, Kant makes several remarks that are most naturally understood as rejoinders to replies that one might make to his arguments. "This truth [the principle of succession] depends on an

easily understood and infallible chain of grounds. Nonetheless, those who give to the Wolffian philosophy its renown have paid so little attention to this truth that they maintain, on the contrary, that a simple substance is subject to constant change in virtue of an inner principle of activity" (1:411). His response: "Although I for my part am thoroughly familiar with their arguments, I am, nonetheless, convinced of their sterility [*ficulnea*]" (1:411). Unfortunately, Kant does not describe the grounds of his conviction, nor, for that matter, even hint at why such arguments are supposed to be sterile as a fig tree is in isolation. Moreover, as we saw above, Kant requires activity in his own account of causality in the *True Estimation*. So his objection cannot be based on the fact that substances are active. As a result, taken in isolation, these remarks turn out to be puzzling, rather than helpful in coming to understand Kant's ultimate view.

The only other passage in the *Nova dilucidatio* that is directly relevant here is the following: "Once they [i.e., supporters of Wolff] have constructed an arbitrary definition of force so that it means that which contains the ground of *changes*, when one ought to declare that it contains the ground of *determinations*, they were bound to fall headlong into error" (1:411). Again, Kant does not explicitly state what "error" Wolffians thus fall headlong into or what is inappropriate about grounds of change as opposed to grounds of determinations. Given the scant immediate textual evidence, the idea could be the very simple one that grounds of determination are prior (perhaps both conceptually and ontologically) to grounds of change. For example, a ground of change from *F* at t_1 to *G* at t_2 necessarily involves grounds of determination, namely one that posits *F* at t_1 and a second that posits *G* at t_2 , whereas a ground of a determination does not necessarily involve a ground of change (insofar as we can, e.g., imagine a world in which there is a single ground of determination and nothing else). At the same time, it is hard to know whether this kind of argument can actually carry the heavy burden that Kant would require of it.

Kant and Wolff on the Nature of Grounds

To make progress in sorting out Kant's ultimate reasons for asserting the impossibility of change within a causally isolated substance, we must understand more clearly both how he thinks of grounds and why he might have come to think of them in this way. Let us begin by considering three distinctive aspects of grounds as they were understood in early to mid-eighteenth-century Germany. First, Wolff, Baumgarten, and Meier all seem to think of grounds in primarily (even if not exclusively) logical

terms. If the principle of contradiction is essentially a logical principle and if one derives the principle of sufficient reason from it, as Wolff does, then it would be natural to infer that the principle of sufficient reason is a logical principle as well.¹⁵ Moreover, Wolff directly links causality to the principle of sufficient reason by defining causality as a principle that contains the (sufficient) reason of the existence or actuality of another.¹⁶ Accordingly, if the principle of sufficient reason is a logical principle, then it would seem that causality is a logical principle as well. This interpretation would also accord with the idea that an effect would follow with (logical) necessity from its cause. In this context, it is striking to take note of Baumgarten and Meier's ubiquitous use of the notion of positing, which has clear logical connotations as well.¹⁷

Second, both Wolff and Meier explicitly assert that a determination must be posited simultaneously with its ground. Wolff, for example, asserts that "what is grounded in another subsists as long as its ground subsists and for that reason cannot be changed as long as its ground is not changed."¹⁸ In short, if ground *a* exists at t_1 , then the determination it posits, namely *F*, must also exist at t_1 . While one might attempt to support the simultaneity of grounds and determinations directly by means of the logical interpretation of the principle of sufficient reason, it can also be motivated by purely ontological considerations. For example, just as objects that never existed cannot bring anything about now, so too, one might think, objects that no longer exist cannot bring anything about now either. For, so the reasoning might go, the existence of something in the past does not necessarily entail anything about the future, since it is both logically and metaphysically possible that the world cease to exist after what would be the cause goes out of existence but before its putative effect begins to exist. To guarantee that causes or grounds necessarily bring about their effects or determinations, they cannot stand at a temporal distance from each other, that is, they must be simultaneous. Alternately, one might attempt to develop an argument based on an explanation of the temporal gap between a ground and its determination.

¹⁵ See §70 of Wolff's *Prima Philosophiae sive Ontologia* (Frankfurt and Leipzig, 1729), for his most explicit formulation of his derivation of the principle of sufficient reason from the principle of contradiction.

¹⁶ See §§66 to §81 of *Prima Philosophiae sive Ontologia*.

¹⁷ While Wolff normally uses the terms *setzen* and *poni* (e.g., in §878 of *Prima Philosophiae sive Ontologia*), he also employs a variety of other terms.

¹⁸ See Wolff's *Vernünfftige Gedanken*, §176 (which is discussed in more detail below), as well as pp. 88–90 of Meier's *Beweis der vorherbestimmten Übereinstimmung*.

That is, if a ground is posited (say, at t_1) and its determination is not posited at the same time, then, so the argument might go, something else in addition to (the positing of) the ground must occur for its determination to arise and, in particular, for it to arise at precisely the time it does (namely at t_2 and not at t_3). But if something other than the ground is necessary for its determination to occur (at that time rather than any other), then one might call into question whether the ground is truly a sufficient ground.

Third, Wolff's conception of what an essence is entails that the grounds that constitute an essence must be necessary and thus immutable. In §32 of his *Rational Thoughts*, Wolff argues that every thing must have a necessary ground or set of necessary grounds that are responsible for its determinations.

If one can distinguish various [determinations] in a thing from each other, one of them must contain the ground within itself for why the rest are attributed to it, and because this cannot in turn have its ground for why they are attributed to it in one of the rest, as can easily be comprehended through the principle of contradiction (§10), they must necessarily be attributed to it. For what necessarily exists in this way requires no further ground for why it exists in this way.¹⁹

In other words, to stop a regress of grounds within a thing Wolff posits a necessary ground or set of necessary grounds within that thing and identifies it with the thing's essence. In §42, Wolff then infers the immutability of a thing's essence from its necessity: "Since the essence of a thing is thus necessary (§38), it is also immutable."²⁰ In §176 of his *Rational Thoughts* Wolff even extends the connection between necessity and immutability from essential grounds to determinations by means of the simultaneity of determinations and grounds:

Since the essence of a thing is necessary (§38), everything that is grounded in it alone must also be necessary. For what is grounded in another subsists as long as its ground subsists and for that reason cannot be changed as long as its ground is not changed. Now the essence of things is immutable (§42). Therefore, whatever is grounded in the essence of things alone must also be immutable.²¹

¹⁹ Wolff, *Vernünftige Gedancken*, p. 18.

²⁰ *Ibid.*, p. 22.

²¹ *Ibid.*, §176, p. 95. §§38 and 41 provide arguments for the main premises of this argument. §38 states: "*The essence of things is necessary.* What is possible cannot be impossible at the same time (§10), and if something is possible in a certain way, it cannot be impossible in that way at the same time, and is therefore necessarily possible (§36). Now since possibility is intrinsically [*an sich*] something necessary, but the essence of a thing consists in it being possible in a certain way (§35), its essence is necessary" (p. 21). §41 states: "Whatever is

Accordingly, Wolff thinks that since the essential grounds of a substance are necessary and immutable, the determinations they posit must be so as well in light of the simultaneity of grounds and their determinations.

However, Wolff thinks of these eternal grounds as grounding merely the *possibilities* that are inherent in things. He thus recognizes the need for something in addition to essential grounds that would make the states of things actual.²² Later in his *Rational Thoughts* (§628) Wolff distinguishes between a thing's essence and its nature by means of the notion of an active force: "And insofar as it is an active thing, one attributes a nature to it: accordingly, nothing else is understood by **nature** than an active force insofar as it is determined in its mode [*Art*] by the essence of a thing."²³ A nature is thus an active force, that is, a force that acts in accordance with the thing's essence. Accordingly, Wolff introduces the notion of an active force as what makes a thing (or its states) actual. By appealing to active force in this way, he can be read as giving content to his earlier description of force or power as "the source of changes" (§115), since active force is what brings about the successive states of a thing. In his lengthy Latin treaty on ontology, Wolff then describes force as "that which contains in itself the sufficient reason of the actuality of an action."²⁴ In short, Wolff ultimately accepts two kinds of grounds: (1) immutable grounds that explain the possibility of all states of a substance and (2) an active force that functions as the sufficient reason of the existence or actuality of these states.

If one assumes that Kant adopts such a conception of essential grounds in the *Nova dilucidatio*, one can see quite clearly both what the fundamental premises of his three arguments for the principle of succession are and why he would think that they represent a serious challenge to the Wolfian position. Concerning the first argument, if (1) all of a substance's

necessary is also immutable. For if it could be changed, then it could also not be, which runs contrary to its necessity (§36)" (*ibid.*).

²² In his *Prima Philosophiae sive Ontologia*, Wolff modifies his position as follows. First, he distinguishes between the reason for the possibility of a thing and the reason for its actuality (§874), calling the former a *principium essendi* and the latter a *principium fiendi*. He then suggests (§875) that the essential grounds of a thing are the essential grounds of the essence of the modes, whereas the grounds of the actuality of these modes are either in antecedent modes of that thing or in other things (or in some combination thereof). Baumgarten's take on the issue in his *Metaphysica* is similar. For while he agrees with Wolff (in §§106 and 132) that essences are necessary and immutable, he argues (§§104–105, 108, and 133) that modes are contingent and can therefore change.

²³ Wolff, *Vernünfftige Gedanken*, p. 384.

²⁴ §722, p. 542, though he immediately admits that this is only a nominal definition.

grounds must be posited in order for it to exist and if (2) grounds must posit their determinations as soon as they exist, then it follows that all of a substance's determinations are posited as soon as it exists. While the second assumption is identical to one of the three aspects of Wolff's notion of ground described above, the first assumption might appear to be original, since it seems to rest on the idea that all of a substance's grounds must be posited at the same time in order for it to exist. Although this idea is not exactly identical to any of the three features of Wolffian grounds described above, it nonetheless follows from the familiar idea that a substance's grounds are immutable components of its essence. For insofar as positing the existence of a substance implies positing its essence, positing the essence of a substance entails that all of its immutable grounds are posited as well, since these grounds constitute that essence. The point is not simply that *all* of its grounds must be posited. Rather, it is that they must all be posited *at the same time*, since they are immutable and thus could not come about at a certain point in the course of a substance's history. As Wolff himself recognizes in the *Rational Thoughts* (§42): "But if I can think of a possible change in the essence of a thing, the essence of the thing is not thereby changed; rather, by cognizing it I have attained cognition of the essence of another thing."²⁵ Accordingly, both assumptions of Kant's first argument for the principle of succession derive either directly or indirectly from Wolff's conception of essential grounds.

Kant's second argument for the principle of succession depends even more clearly on Wolff's conception of grounds. The argument is that a change in a causally isolated substance's determinations requires a change in its grounds, but changing grounds within a substance are impossible. Since Wolff admits that grounds and any determinations they might posit must be simultaneous, he is committed to the idea that a change in determinations requires a change in grounds. However, he also admits that grounds cannot change within a substance, since the grounds that make up the essence of a substance are immutable. Therefore, a causally isolated substance cannot undergo change. Moreover, Wolff's statement at §176 of the *Rational Thoughts* quoted above develops precisely this argument (though he does not explicitly draw this conclusion).

Kant's third argument attacks the possibility that unchanging grounds within a causally isolated substance could be responsible for change. The main point that Kant wants to press against this possibility is that

²⁵ Wolff, *Vernünfftige Gedancken*, p. 22.

regardless of how it is cashed out, it violates Wolff's principle that determinations must be simultaneous with the grounds responsible for them. If a single ground is supposed to be responsible for two mutually incompatible, successive determinations, F and G , then F and G would not always be simultaneous with their ground, since G would not be actual when the ground was positing F and F would no longer be actual after the ground had posited G . If two separate grounds are supposed to be responsible for the two successive determinations that constitute change, then insofar as the grounds are immutable, the changing determinations would not always be simultaneous with the grounds that posit them, since the ground that posits determination G could not be positing it when the ground that posits determination F was positing it (on pain of contradiction). Again, Kant's third argument for the principle of succession directly draws on Wolff's conception of essential grounds.

At the same time, one could object that since all three of Kant's arguments are based on Wolff's conception of *essential* grounds, they misrepresent the resources that he has to explain change. For if Wolff's active forces are invoked instead as the grounds that posit the actuality of the states that are otherwise merely possible, it is no longer immediately clear that Kant's arguments are as compelling. Kant's response to this objection is to deny that active forces can be identified with (nonessential) grounds of change. As we saw above, after presenting his three arguments Kant emphatically objects to the idea that grounds could be understood as grounds of *changes* rather than as grounds of *determinations*. Thus, the point of Kant's objection to defining grounds as grounds of changes is not primarily to emphasize the conceptual or ontological priority of determinations over changes, but rather to reject the idea that active forces could be understood as grounds of changes.

But why might Kant reject such grounds? If an essence is a set of merely possible grounds, or a merely possible being, is Wolff not right to think that some further ground must be added for that essence to exist or be actualized? What Kant wants to reject is not the idea that there is an important difference between actuality and possibility, but rather the idea that the ground of actuality or existence could lie within the substance whose states are to become actual. Stimulated into thinking about the nature of existence by Crusius's novel arguments in the *Sketch of the Necessary Truths of Reason*, Kant discusses the topic in several of his early pre-Critical works. In *The Only Possible Argument*, for example, Kant is explicitly critical of both Wolff's and Baumgarten's accounts of existence (2:76) and develops his own account in terms of a distinction

between *what* is posited and *how* it is posited, arguing that existence is not the relative positing of one determination with respect to another – for example, omnipotence is posited relative to God as an infinitely perfect being – but rather an absolute positing.²⁶ Accordingly, it is a mistake, Kant holds, to think that one kind of ground within a substance posits its merely potential determinations with a second kind of ground positing existence as a predicate. Rather, however the essential grounds are posited, that is, whether posited as merely possible or as fully actual, they must posit their determinations accordingly.

If this critique of Wolff's conception of existence were not enough, in a passage in the *Nova dilucidatio* that precedes the principle of succession, Kant also explicitly argues that it is absurd "to say that something has the ground of its existence within itself" (1:394). Kant's argument proceeds as a *reductio*. Assume that something contains the ground of its existence within itself, that is, is the cause of itself. Since the concept of a cause is prior to the concept of its effect, something would be both prior and posterior to itself, which is absurd. In this context, Kant is objecting to Wolff's characterization of what it is for something to be a necessary being, arguing that one should not say of God that he contains within himself (or within his essence) the sufficient reason of his existence.²⁷ But Kant's argument can be generalized to hold for any being; in no case should one say that something contains the sufficient reason of its existence within itself. Yet this is precisely what Wolff is asserting when he claims that a substance is to have an active force or ground that actualizes or brings into existence its successive determinations. Kant thus rejects the notion of an active force or ground as what brings the successive states of a substance into existence, and therefore the Wolffian cannot legitimately appeal to such a notion to explain change. But since the Wolffian concedes that essential grounds cannot explain

²⁶ "Existence is the absolute positing of a thing. Existence is thereby also distinguished from any predicate; the latter is, as such, always posited only relative to some other thing" (2:73).

²⁷ In §308 of *Prima Philosophiae sive Ontologia*, Wolff explicitly asserts: "A being exists necessarily if it contains the sufficient reason of its existence in its essence" (p. 244). One might suggest that a being could contain the sufficient reason of its existence within itself, but not by containing it within its essence. However, to develop such a suggestion into a cogent argument one would need to explain both which part of a being contains the sufficient reason of its existence (if not its essence) and how this part relates to its essence. As we see below, Kant develops certain distinctions concerning existence that would complicate such an attempt.

change, Kant's arguments for the principle of succession seem to be effective.

Kant and Leibniz on Grounds and Explaining Change

One can thus see how Kant's argument for the principle of succession can be understood against the background of how grounds were conceived in Germany at the time by Wolff, Baumgarten, and Meier – that is, as based on a conception of immutable grounds that posit their determinations with logical necessity as soon as they exist, and on the rejection of active forces as grounds of change, since they involve in an illegitimate way grounds of the existence of things that would have to be internal to those very things, something that is impossible even for a necessary being. However, one can still wonder how his argument relates to Leibniz's view. To address this issue, it is necessary to consider first the differences between Leibniz's conception of ground and the Wolffian conception that Kant is, in certain respects, adopting at this point.

One immediate difference is that Leibniz rejects the idea that a ground must be simultaneous with its determination, since he thinks that the determinate state of a substance at one moment is the cause of its state at the next moment in time. While such a difference about whether or not grounds must be simultaneous with their determinations might seem to be a minor, technical issue, it actually reveals significant differences between Leibniz's position and that of Wolff and his followers. For one, Leibniz's denial that grounds must be simultaneous with their determinations makes it possible for him to explain change at the level of derivative rather than primitive forces. Thus, in a sense one can say that for Leibniz, explanation of change runs "horizontally" from one determinate state to another, rather than "vertically" from one determinate state down to the primitive force that constitutes a monad. It is true that derivative force depends on primitive force for Leibniz, but the dependence relationship is not, Leibniz thinks, immediately relevant to explaining any specific features of the change. Rather, primitive forces, which endure throughout change, are required primarily as a metaphysical unity against which change takes place; they do not explain why one state is determinate in the way that it is rather than in some other way or at that time rather than at some other. Because Wolff and his followers, including Kant in this case, are committed to the simultaneity of grounds and determinations, they cannot explain change in terms of derivative forces, but rather must

do so in terms of the primitive forces or essential grounds that constitute a substance.

Since these differences between Leibniz's conception of ground and explanation of change, on the one hand, and Wolff's and Kant's, on the other, are significant, Kant's explicit arguments for the principle of succession are not aimed directly at Leibniz and might seem to be moot.²⁸ However, even if Kant's explicit arguments do not directly target Leibniz's position, it is still possible to see how their views could engage, rather than simply talk past, each other. The issue could be decided by settling such thorny issues as whether a cause or ground must be simultaneous with its effects or determinations in order to produce them. While Kant accepts such a simultaneity principle (both in his pre-Critical period and, as we see below in Chapter 4, in a somewhat modified way in his Critical period), he does not present any explicit arguments that Leibniz might find decisive. Rather than tackling particular philosophical claims directly, it might be more profitable first to consider differences in what Leibniz and Kant aim to explain and then what instances of change their accounts are designed to render intelligible in the first place.

From this perspective, there is one general feature of Leibniz's view that Kant would find clearly lacking. By explaining change at the level of derivative forces, there is a genuine sense (or set of senses) in which Leibniz fails to deliver (or even expect as possible) an ultimate explanation of change. One way of putting this point is to note that Leibniz himself concedes that one should not invoke primitive forces in explaining particular states.²⁹ Leibniz invokes primitive forces not to explain any specific features of change, but rather to satisfy metaphysical demands (such as that only what is truly one is real, that substantiality necessarily requires activity, or that a derivative force must be a determination

²⁸ This would be the case especially if Leibniz were to admit that monads are not ultimately temporal. While it is unlikely that such an admission reflects Leibniz's considered view, both textual and philosophical motivation can be found for it. Leibniz seems to develop an argument quite similar to Kant's in a letter to Foucher from 1675. As Leibniz puts it: "Now, this variety [in our thoughts] cannot come from that which thinks, since a single thing by itself cannot be the cause of the changes in itself. For everything would remain in the state in which it is, if there is nothing that changes it; and since it did not determine itself to have these changes rather than others, one cannot begin to attribute any variety to it without saying something which, we must admit, has no reason – which is absurd" (G. W. Leibniz, *Philosophical Essays* (Indianapolis: Hackett, 1989), p. 3). See Robert McRae's "Time and the Monad," *Nature and System* (1979): 103–109, for a rational reconstruction of an argument, based on principles that Leibniz explicitly accepts, showing that temporal notions and thus change are not real.

²⁹ See, e.g., his "A New System of Nature," *Philosophical Essays*, p. 139.

of a more fundamental, determinable entity). What, one might ask, has Leibniz omitted in explaining change at the level of derivative rather than primitive forces? Kant would contend that explanations of change in terms of derivative forces cannot be satisfying insofar as derivative forces themselves stand in need of explanation and not just in terms of other derivative forces, but rather in terms of primitive forces. While this objection can seem weak, especially when stated in such abstract terms, there is a real point to it, since it is based on the idea that any satisfactory explanation of change ought to involve in specific and intelligible ways not just the changing circumstances in which the change occurs (its derivative forces), but also the nature of the thing that is changed (its primitive forces). Without a direct connection between the ultimate metaphysical nature of a thing and its particular states, there is only a nominal sense in which change is change of the thing and, what's more, the benefit of positing primitive forces at all becomes attenuated. For example, if Leibniz refuses to draw a specific connection between the unity and activity of primitive forces and particular determinations of bodies, then he cannot easily rule out the possibility that there is a single soul for the entire corporeal world. If extension requires unity and activity and nothing more, then a single monad would suffice.

Another way of pointing out that any explanation of change that invokes derivative forces alone might seem to fall short is to note first that even the most successful explanation of change in terms of derivative forces is necessarily interminable, that is, leads to an infinite regress, and second that such an interminable explanation or infinite regress can hardly be desirable at a purely metaphysical level. It is clear enough that explanations of change in terms of derivative forces will be interminable, since the state of a substance is causally connected to all of its infinitely many past and future states. While one might think (with Hume) that explaining each and every state by a previous state leaves nothing unexplained, later in the first *Critique* Kant argues (not in his own Critical voice, but rather in a voice that is consistent with his pre-Critical position) that reason cannot accept such an idea. In the Thesis of the Third Antinomy he argues that explaining change in terms of previous change, "when taken in its unlimited universality, contradicts itself," since "nothing happens without a cause sufficiently determined" and "at every time there is only a subordinate but never a first beginning, and thus no completeness of the series on the side of the causes descending from one another" (A446/B474). In short, the sufficiency that is required for the actuality of a change conflicts with the incompleteness generated by an infinite

causal regress. As Kant points out in his “Critical Decision of the Cosmological Conflict of Reason with Itself” (A497/B525), when it comes to how things really are (as opposed to how things appear to us), reason cannot be satisfied with anything less than the unconditioned condition of all conditioned states, which is precisely what is absent from explanations of change in terms of derivative forces.³⁰

Another aspect that is relevant to assessing the comparative strengths of Leibniz’s and Kant’s conceptions of ground and their corresponding explanations of change concerns what they take as the primary instance of change that stands in need of explanation. Leibniz, as an idealist, thinks that the apparent interaction among bodies need not be taken to imply actual interaction between substances (at least not at the ultimate level of metaphysics), since bodies are merely well-founded phenomena, where apparent interaction can be seen to be just that, namely apparent. What does require serious explanation are perceptions (as mental states), and his conception of grounds, according to which one of my perceptions leads to the next, can provide a much more intuitive explanation of such changes than would be given by claiming that some other mind acts on me so as to cause my current perception. But if Leibniz’s conception of grounds and explanation stands and falls with his idealism, then Kant’s account could easily appear more attractive. For if Kant is not committed to Leibniz’s idealism, but is rather open to a physical monadology (as he obviously is, given his explicit rejection of Leibnizian idealism in the *Nova dilucidatio* in favor of “the real existence of bodies” (1:411) as well as his position in the *Physical Monadology*), then it can be tempting to assert that reality is ultimately constituted not by metaphysical points alone (which Leibniz identifies with minds), but also by physical points endowed with physical forces that generate causal interaction between each other so as

³⁰ The Critical Kant will concede that not all of reason’s demands can be satisfied in the world of appearances, but such a concession is of no use to Leibniz at the level of metaphysics, since even the Critical Kant continues to hold that an infinite regress is unacceptable to reason. “If the conditioned as well as its condition are things in themselves, then when the first is given not only is the regress to the second **given as a problem**, but the latter is thereby already **given** along with it; and because this holds for all members of the series, then the complete series of conditions, and hence the unconditioned is thereby simultaneously given” (A498/B526).

One might object, on Leibniz’s behalf, that the unconditioned condition of all conditioned states could simply be an infinite series, but if the series is incomplete, it is difficult to see that one has truly attained anything unconditioned (unless one takes the series as a whole to be more than all its members, an assumption Kant could resist).

to explain the physical states of bodies. That is, physical points could be the seats of primitive, physical forces.

However, for physical forces to be primitive forces, they can no longer be identified with particular, observable, and determinate bodily states that might change when bodies communicate their motion to each other in collisions. Rather, physical forces would need to be understood as enduring, essential aspects of substances that are not only not directly observable (insofar as everything we observe undergoes change), but also both underlie and explain the determinate states we do see. That Kant makes this connection is clear from his position in the *Physical Monadology*, where attractive and repulsive forces are used to explain the particular bodily states we observe and are understood in terms of the unobservable activities that occur in accordance with the unchanging masses of bodies. Accordingly, accepting this model puts Kant in a position to provide a stronger, more satisfying explanation of (changes in) determinate states, an explanation that provides more than what Leibniz thinks can be had.³¹

In sum, given that Kant's arguments for the principle of succession are based only on the Wolffian conception of grounds and not on Leibniz's version as well, and given that he has not provided a non-question-begging refutation of Leibniz's idealism, Kant is overstating his case in claiming that change within a causally isolated substance is metaphysically impossible. In light of the important differences between Leibniz and Wolff, Kant's position must be more nuanced. Certain conceptions of change within a causally isolated substance (e.g., Wolff's) can be shown to be metaphysically impossible, while other conceptions of change (Leibniz's) may ultimately be unsatisfying or at least less attractive, all things considered.

Kant's Positive Account of Change

If these considerations render intelligible Kant's multifaceted reasons for rejecting pre-established harmony, that is, if they show why he thinks that change cannot occur in a causally isolated substance, one must still ask how he thinks change does occur. Obviously enough, there must be

³¹ It should be noted that I am not claiming that adopting a *physical* (rather than idealistic) monadology immediately entails the impossibility of pre-established harmony. One could accept physical points that God has set up with such great harmony that their self-caused changes coincide. (I thank Bob Adams for discussion of this kind of possibility.) Rather, the point is simply that, barring strong philosophical reasons to the contrary, it would be more in line with common sense to assume causal interaction between bodies.

a causal connection between substances.³² Still, a number of questions arise about the nature of this causal connection. How could causality *between* substances explain change if causality *within* a substance cannot? Could God be the substance that causes substances to change, which would reopen the door to occasionalism and pre-established harmony? Could the causal connection be merely one-directional or must it be reciprocal, that is, mutual interaction? Could a substance change its own state at the same time that it changes the state of another? In attempting to address these questions, it is important to keep in mind that the case of bodies in motion continues to be Kant's primary model of change and that changes in mental states (in the soul) are derivative from bodily changes. Although he presents no explicit justification, he asserts that "motion is the appearance of a changed connection" (1:410), and, in the scholium to the principle of succession, that "the soul's state would be

³² Even if two substances must be connected in order to explain change, one might still ask why the connection must be causal, rather than spatial, temporal, or logical. Of course, if the arguments that Kant presents for the principle of succession turn on the idea that *causally isolated* substances could not change, then the connection would necessarily be causal. However, one might think (as Langton seems to, *Kantian Humility*, p. 118) that the arguments are ultimately based on what follows from the mere *existence* of a substance, that is, from the existence of a substance with its intrinsic, necessary properties or determinations. If the argument is based on this, less robust notion, then the mere fact that a connection is required does not immediately imply that the connection between the substances must be causal, since the substances could stand in noncausal relations. If one substance exists and another substance comes to exist at a later time, then the first substance has changed at least some of its determinations without any overtly causal connection arising between them. For example, the first substance comes to have the determination of not being the only substance to exist.

Although Kant does not explicitly consider the possibility of noncausal connections in this context, he clearly thinks that determining grounds are causal in nature and that they must be causal to explain the fact that a substance comes to have a determination that it previously lacked. As with any version of the principle of sufficient reason that might appear at all plausible, one must divide all determinations (properties or states of affairs) into "real," "positive," "primitive," ones, which stand in need of a reason, and "unreal" (or "ideal"), "negative" (or "logical"), "derivative" ones, which do not require separate reasons of their own. Accordingly, if the determination is real, then there must be a ground for it, that is, it must have a cause. All other determinations – for example, determinations that exist if two substances are brought into existence in such a way that they stand in some relationship to each other without the determination depending on either a real determination of either substance – do not require a separate ground or cause, since they are to be explained derivatively, that is, by means of the grounds that explain the real determinations on which the ideal determinations depend. If such a distinction can be upheld, then Kant's assertion that the connection between substances required for change must be causal is justified by the principle of sufficient reason or, to put it in his own preferred way, the principle of determining ground, which was demonstrated in the second section of the *Nova dilucidatio*.

immutable and constantly like itself if it were completely released from external connection" (1:412).

Causal Connections plus Changing Relations

How is change to be explained on Kant's account? It would be natural to think that introducing a causal connection between two substances would suffice to explain the occurrence of a change (in at least one of them). The one substance exists with its determinations and then the second substance causes a change in the first substance by bringing about a new determination in it. Although the first substance cannot bring about a new determination in itself, it might seem that the second substance could do so by means of its internal grounds, for the grounds that are internal to it are external to the first substance.

However, Kant recognizes that simply introducing a causal connection between two substances will not suffice to explain change. In explicating the principle of succession, he states: "Furthermore, even were this simple substance to be included in a connection with other substances, if this relation did not change, no change could occur in it, not even a change of its inner state" (1:410). That is, Kant infers that change in the intrinsic determinations of a substance requires not just another substance, but also changing relations between these substances. On the one hand, this claim is not particularly surprising insofar as it is understood as applying to bodies and motion. If one body is sitting at rest, positing a second body at rest will not necessarily cause any change in the intrinsic determinations of the first body. Something more, such as the motion of the second body toward the first, is needed for a change to occur. On the other hand, at the metaphysical level it represents a striking reversal of Leibniz's position, according to which relations between substances could change only if their intrinsic or monadic properties changed.³³

What is Kant's justification for this important claim? The idea is already implicit in his arguments for the principle of succession. Just as there is no element within a substance that could explain change in that substance, introducing a second substance does not immediately add a suitable element, either. It, too, has grounds that, once they exist, posit their determinations immediately. Thus, if God creates two substances with all of their grounds, all of the determinations of both substances will

³³ Langton (*Kantian Humility*, p. 106 n. 14) rightly notes this reversal, though she seems to think that it is Kant's rejection of Leibniz's doctrine of the reducibility of relations that is ultimately responsible for this reversal, whereas the account developed above suggests that it is due to Kant's conception of what a ground is.

be posited immediately, either by means of each one's internal grounds or by means of the external grounds in the other substance.

However, if the substances are changing with respect to each other, then the effects that their grounds might have on each other could be changing too. In other words, if they change their relations to each other, then which grounds are efficacious could be changing as well. Thus, changing relations between two substances could suffice to explain changes within those substances if they are causally connected. At the same time, if the substances were *not* causally connected, then the justification for thinking that changing relations implies changing intrinsic determinations would disappear, since relations are external to the substances and their grounds and would therefore have no effect on them. Thus, Kant's positive explanation of change invokes both a causal connection between substances and changes in their relations so that the changing causal connection between the substances can be responsible for changes in their internal determinations.

At this point, two obvious difficulties arise for Kant's view. First, if Kant argues that changing determinations require changing grounds (due to the principle of determining ground), but that grounds cannot change within an isolated substance, then it might seem equally unacceptable to claim that grounds (and thus the determinations they posit) could change for connected substances. Even adding the idea that substances change their relations to each other might seem to be irrelevant, since all of the grounds of a substance are immediately posited along with its very existence. In other words, changing the relations between substances cannot change the grounds that are given along with their existences and it is thus not immediately apparent how changing relations would be of any help in explaining change.

Second, Kant's explanation of change might seem to be vulnerable to a charge of vacuity. For it might seem that explaining the change of internal determinations on the basis of changing external grounds is vacuous insofar as one is still invoking change in order to explain change. Moreover, Kant might seem to be avoiding these difficulties when he concludes his elucidation of his arguments for the principle of succession as follows:

Again, suppose that someone wished to know how, in the final analysis, the alterations, of which the succession is apparent in the universe, take place, granted that they do not issue from the internal factors of a substance considered in isolation. I would have that person turn his attention to things that follow as a consequence in virtue of the connection of things, in other words, in virtue of

the reciprocal dependence of their determinations. For the rest, to offer a more detailed explanation of these matters here would take us rather beyond the limits of our treatise. Accordingly, our demonstration establishing that the matter certainly could not be otherwise will have to suffice. (1:411)

Despite these seemingly evasive remarks, Kant may have the resources necessary to respond to these two charges. First, while Kant clearly must maintain the principle of determining ground and thus the idea that a change in internal determinations requires a change in grounds, he can deny that the changing relations between substances are irrelevant to their causal connection, as the first objection maintained. For even if the change in relation between the connected substances cannot change their grounds, it can cause different grounds to be effective. Consider, once again, the case of a body in motion. If one body moves closer to another, the second body might resist more strongly the attempt of the first to penetrate it, that is, first the ground that determines a push of strength x is effective since the circumstances necessary for it being efficacious are given, and then the ground that determines a push of strength y is effective since the circumstances, which have changed, now satisfy the conditions for it being efficacious. In this way, Kant need not hold that grounds literally change. Rather, he can say that the change in relations between bodily substances changes which grounds in each substance bring about which determinations in the other. Accordingly, Kant can explain how changing relations might explain changing determinations without falling prey to his own objection to Leibniz.

Second, in response to the charge of vacuity or circularity, while it is true that Kant's own account explains change in virtue of change, there is an important difference in the kind of change that explains the change in determinations. For Wolff, change in determinations is to be explained in terms of an internal ground (or set thereof), and Kant's objection to Wolff is that an internal ground (or set thereof) is immutable if contained in the substance's essence. Because he hopes to explain change not exclusively in terms of (either internal or external) grounds that might be immutable, but rather by also invoking contingent relations between substances that could constantly change, the structure of his explanation is importantly different. Since Kant's explanation involves (minimally) two substances, their relations to each other can change, which generates change in their determinations. That is, Kant has a natural place for the requisite element of change, namely in the relations between substances, which, as we see below, are not immediately and completely governed

by their internal grounds. By contrast, according to pre-established harmony, since all external substances are, by definition, causally irrelevant, the only resources that one could draw on to explain change would have to be internal to each substance. However, so Kant's critique goes, no such internal resources would seem to be available, given that each substance is self-sufficient by containing all of its grounds and thus all of its internal determinations within itself as soon as it exists at all.

Kant's response to these objections thus reveals that his general strategy for explaining change is radically different from, and, he thinks, superior to, the explanations that Leibniz and Wolff adopt. As we saw above, Leibniz explains change in determinations in terms of change of grounds, making it unclear that change has really been given an ultimate explanation (insofar as change is explained entirely in terms of changing determinations and tied in no specific way to any essential aspect of substance). Wolff avoids this objection by invoking immutable grounds that are directly linked to the essence of a substance to explain the determinations of that substance, but the link between immutable grounds and determinations ends up being so strong that change becomes impossible, a cost that Kant thinks is too high. As we now see, Kant, by contrast, attempts to explain change in determinations by appealing to the contingent and mutable relations between substances endowed with immutable grounds. While it is true that Kant's account is, to some extent, vacuous insofar as it still explains change (in determinations) in terms of change (in relations) – a kind of circularity that may, to some degree, be inevitable – it can avoid the problem that Leibniz's account faces (namely that there is not even a partial ultimate explanation of change in terms of the essential features of a substance). For specific features of change are at least partially explained by essential features (i.e., immutable grounds) of the causally interacting substances. At the same time, despite asserting a connection between determinations and their essential grounds, Kant can still account for the contingency inherent in change because the determinations depend in part on the contingent relations between the causally interacting substances. In this way, Kant's own positive explanation of change can, he believes, attain advantages over both Leibniz's and Wolff's positions.

Can God Cause Change?

However, if God is ultimately responsible for change creating substances that are changing their relations to each other, then one might think that pre-established harmony or occasionalism could be saved by noting that,

technically speaking, neither one is in fact incompatible with the principle of succession as long as God is the substance responsible for changes in finite substances. That is, the principle of succession requires that one substance be causally connected with another substance for change to occur. Why not suggest, as occasionalists do, that God is the cause of the changes that occur in finite substances or, as Leibnizians would maintain, that God and finite substances concur so that finite substances can bring about their own changes? In that way, changes are brought about by means of a causal connection (namely creation) between two substances (one finite, the other infinite), just as the principle of succession requires.³⁴

Kant would reject attempting to save pre-established harmony or occasionalism in this way because, so he would think, God cannot both bring about the existence of a substance and “then” change it by changing his relation to the substance. Change occurs when an already existing substance loses one determination and gains another. If the principle of succession is correct, the substance changes in this way not only because the substance is causally connected to a second substance but also because of a change in their noncausal relations to each other. The suggestion currently under consideration is that God could bring about change in a finite substance considered in isolation from all other finite substances. Given the principle of succession, God could change this substance only due to a change in his relation to it. However, God cannot change relations to the substance in this way.

If we consider the case of motion, Kant's primary instance of change, we see how implausible it is to think that a substance could change its own state due to a change in its spatial relation to God. God is not spatial and thus there is no spatial relation between him and the substance in question that might change, and, in the *True Estimation*, Kant holds that a causally isolated substance would not exist anywhere in space insofar as it is only through the causal interaction of substances that space emerges. As a result, if the case of motion is being considered, it is implausible to think that pre-established harmony could be saved in this way.

However, Kant even has resources to respond to such a Leibnizian attempt at a more general level. First, Kant explicitly argues that since God

³⁴ This is not to say that pre-established harmony necessarily collapses into occasionalism. For one need not claim that *only* God's causality is required; one could hold that the internal grounds within a substance could still be necessary for explaining at least some of its properties.

stands in a merely one-way causal bond with finite substances, he must be immutable. And if God is immutable, he cannot be changing his relations to his creatures. Second, in the *True Estimation*, Kant explains the origin of motion in the world as follows: “The very first motions in this universe were not produced by the force of a matter in motion; for otherwise they would not have been the first. But neither had they been caused by the direct power of God, or any intelligent being, as long as it was still possible that they could arise through the action of a matter at rest; for God spares himself as many operations as he can without adversely affecting the mechanical structure of the world, while making nature as active and efficacious as is possible” (1:62). Accordingly, Kant is suggesting here that God would not be the direct cause of motion if some mundane cause, for example, the action of matter at rest, could be identified.³⁵ Finally, one might think that such a change is not possible because God is causally responsible for the existence of the substance in the first place and thus cannot be changing with respect to what he is creating at that very moment. In other words, one might doubt that God could bring about the existence of a substance with all of its determinations at one moment in time and change its relationship to that substance at the very same moment in time, thereby creating a new determination in the substance. God can create the substance so that the substance is changing. In fact, as we saw earlier, Kant’s own position requires that God create substances whose relations to each other are changing. However, one might think that God cannot create the substance so that it is changing *because* of the way in which his relation to that substance is changing. It must be changing rather because of a second substance God creates such that it is causally connected with, and changing with respect to, the first. Thus, neither pre-established harmony nor occasionalism can be saved in this way.

Must the Connection between Substances Be Mutual Interaction?

If causal interaction between substances is thus required for change, it still remains to be seen exactly how change occurs, that is, by what causal mechanisms change can be produced. In particular, one striking feature

³⁵ In his *Universal Natural History and Theory of the Heavens*, Kant continues to think that it is the attractive forces of matter that generate motion from rest. As he puts it: “In space that is filled in this way universal rest lasts only a moment. The elements have essential forces for setting each other in motion and are themselves a source of life” (1:264).

of Kant's principle of succession is that it requires that the causal connection between substances be reciprocal, that is, amount to mutual interaction, rather than something weaker, for example, a one-directional causal bond.³⁶ What justifies this stronger claim?

It is, again, helpful to start with the case of the collision of two bodies. In such a case, the two bodies first come close enough to each other that they exercise their repulsive forces and then, as a result of that exercise, they stop moving closer and perhaps rebound away from each other. In such a case, we do not hesitate to say that mutual interaction has occurred since, in light of Newton's law of the equality of action and reaction, the action of the repulsive force of the one corresponds to the reaction of the repulsive force of the other. Moreover, what the action and reaction explain is both (1) a reciprocal change of the substances insofar as both bodies move differently with respect to each other after the collision than before it and (2) a change in their own intrinsic states insofar as each one has a motive force afterward different from that beforehand.

If Kant is thinking primarily of the case of bodies in developing his metaphysical account of causality, then the principle of succession would seem to be simply a metaphysical version of this idea. The motion of two bodies toward each other allows for a change in the efficacy of the grounds of each substance so that a new determination can result. After the two bodies have collided, reciprocally exercising their repulsive forces and thereby both stopping their mutual approach and initiating their mutual withdrawal, the change that has occurred is properly described as reciprocal given that motion is necessarily reciprocal.³⁷ Moreover, insofar as each body has a different force after the collision, the reciprocal change of the two substances also entails changes in their intrinsic states. Given this model, we can better understand why Kant asserts that "their reciprocal dependency on each other determines their reciprocal changes of state" (1:410).

It is worth highlighting the fact that Kant thinks both that the causal interaction of two substances brings about a reciprocal change in them and that the intrinsic change of each of the interacting substances depends

³⁶ In fact, implicit in the previous attempt at saving either pre-established harmony or occasionalism was the idea that a unidirectional bond from God to finite substances might suffice to explain change. Even if a unidirectional bond from God to finite substances is not feasible, that does not immediately rule out such bonds between finite substances.

³⁷ Kant also emphasizes the reciprocal nature of effects when he introduces the concept of negative magnitudes (2:174).

on this reciprocal change. Without these assumptions, the door is opened to questions about whether the causal interaction must be reciprocal. Assume, for example, that change need not be reciprocal. If substances A and B change their relationship to each other so that the efficacy of the grounds contained in them could change, but, in fact, only A changes, then it may not be immediately obvious why one would need to suppose the efficacy of grounds in both A and B in order to explain the change in A. After all, why is the simplest and most natural explanation of the change in A not that it was caused by grounds only in B?³⁸

Can a Substance Cause Changes in Itself?

If Kant must claim that mutual interaction (rather than any weaker, one-directional causal bond) is required to explain change, a greater burden must be placed on his argument for this claim. In particular, one might ask what argument Kant has for excluding the possibility that a substance could cause a change in both another substance and itself at the same time. *Prima facie*, it might seem that one substance could change itself by changing another substance. Could a substance not change its own state if the other substance is involved merely as a passive recipient of a change in its state and not also as a cause of a change in the first substance? For example, if a person causes students to learn, it would be natural to say that the act of teaching makes that person a teacher at the same time that it brings about a change in the students' knowledge or abilities.³⁹ Moreover, the case of matter in motion might seem to support this possibility insofar as one body might exercise its attractive force on

³⁸ In a slightly later context (1762–1764), to which we have reason to return below, Kant provides an argument that would show that grounds in both A and B would be required for change even in a single substance as follows:

If a substance suffers, then it must contain in *itself* by its own power the ground of the inherence of the accident, because the accident otherwise would not inhere in it. But the ground of this must also be *in the efficient power of the substance*, because it otherwise would not have an effect. . .

An accident thus inheres by its own power, which contains the *sufficient inner ground* of it yet also by *alien* power, as by *an outer ground* of the inherence without which it would not have inhered. (28:51–52)

While the argument Kant develops here is interesting, especially insofar as it addresses the question of why causal relations must be mutual and not merely one-directional, it also clearly goes beyond the considerations he employs in the context of the principle of succession.

³⁹ For a similar example, see 28:26, though Kant uses his example to show that a teacher can put students in a passive state of attentiveness.

another body, thereby pulling it closer, which immediately entails that it has changed its own state (given that it has caused itself to be closer to the body it is attracting).

It might seem that Kant's arguments for the principle of succession do not exclude such a position, given that they explicitly target only the possibility that a substance would change its own state all by itself (i.e., without being connected to other substances). However, in the case envisioned the substance would not be changing its state all by itself, since the substance could change its own state only by changing the state of another. It is true that such a possibility would require that a single ground be able to bring about two determinations at one and the same time in two different substances. Yet it is not immediately obvious on the basis of what Kant says in the *Nova dilucidatio* why "complex" grounds of that sort should not be possible.

One might think that the difficulty with a substance changing its own state in this way lies not so much in the idea that a single ground could bring about two determinations, but in the dependency that exists between the two determinations that it posits. For the determination being posited in the substance that is acting depends on the determination that is posited in the second substance. If the second substance does not exist or if it exists but is, for whatever reason, incapable of accepting the determination the first substance is attempting to posit in it, then the first substance cannot posit the determination in itself. But if the ground in question can posit both determinations *only* if the second substance exists and acts in a particular way, then it would be natural to interpret the situation as one in which the second substance does in fact contain a further ground for the existence of the determination in the first substance and the connection would, contrary to initial appearances, be reciprocal. The intuition here is not necessarily specious insofar as "teachers" whose students do not learn may not be teachers in any nonderivative sense.

One might respond to this objection by distinguishing between active and passive grounds such that the first substance is an active ground insofar as it actively brings about changes, whereas the second substance is merely passive. For, on the model being proposed, all that is required for the changes to occur in addition to the activity of the first substance is the passive receptivity of the second substance. Thus, the second substance does contain a ground for its own changes, namely the passive ground by means of which it can receive the determination posited actively by the first substance, but it does not contain an active ground that would cause a change in the first substance. Accordingly, the principle of succession

might seem to be consistent with the idea that a substance could bring about changes in itself at the same time that it brings about changes in another substance. The second substance would have to have a ground, but as a merely passive ground it might not necessarily entail mutual interaction *per se*. Therefore, it might seem at this point that a substance could act on itself so as to change itself, which goes against what seems to be the main thrust of the principle of succession.

It is important to note, however, that this counterexample depends on several points that are hardly trivial. First, even if one grants a distinction between active and passive grounds, it is not a trivial undertaking to establish that a “merely” passive ground in one substance in conjunction with an active ground in another substance would not be an instance of mutual interaction. Second, if explaining change requires both an active and a passive ground, then the change that occurs in the substance that is changing both itself and another substance would have to have both an active and a passive ground within itself. Yet one might argue (as Kant himself seems to at B153) that a substance cannot be active and passive toward itself in one and the same respect.⁴⁰ It is thus far from clear that Kant would accept this rather speculative counterexample to the principle of succession.

THE NOVA DILUCIDATIO AND THE PRINCIPLE OF COEXISTENCE

While the principle of succession states merely *that* substances must be causally connected for change to be possible, Kant devotes a second principle, the principle of coexistence, to explaining *how* such a connection can come about. It states: “Finite substances stand in no relationship to each other through their mere existence and have no community except insofar as they are conserved in the form of reciprocal relations by the common ground of their existence, namely the divine understanding” (1:412–413). In short, this principle claims that only the divine understanding can enable the reciprocal interaction asserted by the principle of succession. Kant demonstrates the principle in two steps. He first argues for the negative thesis that substances do not stand in relations to each other by means of their mere existence. He then argues positively

⁴⁰ In the second edition *Transcendental Deduction* Kant avoids the contradiction by distinguishing between inner sense and apperception (and thus by drawing a contrast between the senses in which the self is active and passive at the same time).

that a common cause, namely God, is required to establish such causal relations between substances.

Leibniz, Crusius, and the “Mere Existence” of Substance

Kant argues for the negative thesis as follows:

Individual substances, of which none is the cause of the existence of another, have a separate existence, that is to say, an existence which can be completely understood independently of all other substances. If, therefore, the existence of some substance or other is posited simply, there is nothing inhering in it which proves the existence of other substances distinct from itself. But since a relation is a relative determination, that is to say, a determination that cannot be understood in a being considered absolutely, it follows that a relation and its determining ground can neither of them be understood in terms of the existence of a substance, when that existence is posited in itself. If, therefore, nothing further than this were admitted, no substance would stand in relation to any other substance, and there would be no interaction at all between substances. (1:413)

To understand Kant's argument for the principle of coexistence, it is crucial to understand what position Kant is arguing against. One might think that Kant is continuing his attack on Leibniz by arguing that relations are not necessarily reducible to monadic, or intrinsic, properties, as Leibniz had (perhaps) thought. Kant's argument would attack the principle of the reducibility of relations by showing that at least relations of one kind, namely causal relations, are not reducible insofar as causal relations do not follow from the “mere existence” of a substance, that is, from the intrinsic properties of a substance. In other words, since the mere existence of substances does not ground their relational properties, the relational properties must be added as a distinct kind of property not reducible to their intrinsic features.⁴¹

However, the assumption that Kant's argument is directed against Leibniz (and his principle of the reducibility of relations) is problematic for a number of reasons. First, although Kant's position (in both the principle of succession and the principle of coexistence) commits him to denying the reducibility of relations, throughout his discussion of the principle of coexistence Kant never even explicitly mentions, much less focuses on, the issue of whether relations are *reducible* or not. Rather, Kant's concern seems to be with whether the mere existence of a substance is capable of *grounding* a relational property and his argument

⁴¹ See Langton, *Kantian Humility*, pp. 107–123, for such an interpretation.

seems to be based on the idea of the mere existence of a substance, that is, on what one understands merely by positing the existence of an independent substance.⁴² Though these two issues may be related, it would take argument to establish any strong entailment relation between them, argument Kant does not provide as such.⁴³

Second, it is important to keep in mind what Kant thinks he has already accomplished at this point in the *Nova dilucidatio* and what still lies ahead of him. Kant thinks that the principle of succession has already refuted pre-established harmony. If substances undergo change (as Kant thinks they do), then they cannot be causally isolated. What the principle of coexistence needs to show is what it is that makes causal interaction possible, an issue that is in fact irrelevant to Leibniz given his denial of causal interaction. Furthermore, Kant's claim that the mere existence of substances is not sufficient to make causal interaction possible is something that Leibniz himself explicitly affirms. For Leibniz infers from pre-established harmony that each substance (by means of its mere existence as a self-sufficient entity) is like a "world apart," that is, stands in no causal connections. As he puts it in "A New System of Nature": "This is what . . . makes the perceptions or expressions of external things occur in the soul at a given time, in virtue of its own laws, as if in a world apart, and as if there existed only God and itself."⁴⁴ Even Kant's reasoning sounds similar to Leibniz's. The self-sufficiency that attaches to a substance by means of its very existence precludes relations with other substances.⁴⁵ Finally, Leibniz would also agree with Kant's positive claim that God is necessary to establish harmonious relations between substances, whether

⁴² Although Leibniz may have been interested in the reducibility of relations, the issue received little attention in early to mid-eighteenth-century Germany.

⁴³ To see that reducibility and grounding are different notions, consider that reducibility concerns properties, whereas grounding concerns the relationship between grounds and properties. (Langton consistently speaks of grounds as if they were properties, which is dubious, both textually and philosophically. Langton herself seems to recognize it as philosophically problematic when she remarks, in a comment on the principle of succession, that "the assumption that for every intrinsic property there is another intrinsic property that is the 'reason' for the first seems to imply an infinite regress," *Kantian Humility*, p. 105.) Consider also that if grounds *a* and *b* posit relation *c*, it does not necessarily follow that *c* is reducible to *a* and *b* or, for that matter, to any intrinsic properties that *a* and *b* might (or might not) posit. Further, what is at stake in Leibniz's claim about the reducibility of relations is not the logical question of whether relations can reduce to intrinsic properties, but rather the metaphysical question of whether relations should be considered real. It is clear that Kant does not raise this latter question at all in his discussion here.

⁴⁴ See Leibniz, *Philosophical Essays*, p. 143.

⁴⁵ Granted, Leibniz might accept the argument only for real, not for ideal relations.

those relations are causal or not. Thus, as far as the content of the principle of coexistence goes, the Leibnizian would seem to be Kant's ally, not his opponent. If Kant has thus already refuted pre-established harmony in the principle of succession and if Leibniz would immediately agree with Kant's claim in the principle of coexistence that substances stand in interaction not by means of their very existence, but rather only with God's help, it becomes considerably less plausible to assert that Kant is targeting Leibniz (or any of his orthodox followers) with the principle of coexistence.⁴⁶

The final piece of evidence against thinking that Kant is attacking Leibniz here lies in the fact that this would entail attributing to Kant an implausible argument for the assertion of the irreducibility of relations. While it is of course possible that Kant offers bad arguments, it is hardly desirable to interpret him in such a way as long as the text does not demand it, and there is a plausible alternative reading. Consider first the following reconstruction:⁴⁷

A relational property cannot be reduced to an intrinsic property of a substance because if a relational property could be reduced to an intrinsic property of a substance, then the intrinsic property of that substance would imply the existence of another substance (namely, that substance to which it is related). However, one substance cannot entail the existence of another substance (unless it creates it, which Kant has ruled out), since that would violate the idea that substances are capable of independent existence. Therefore, a relational property cannot be reduced to an intrinsic property of a substance.

While this argument may show that a property that relates two substances cannot be reduced to an intrinsic property of one of the substances, it does not establish that it could not be reduced to the intrinsic properties of both of the substances. For example, the relational property "is taller than" that holds between two people cannot be reduced to the height of only one of the individuals, but would seem to be reducible to the heights of both individuals. Thus, this reconstruction does not represent a plausible argument.

⁴⁶ One might still think that Kant is responding to Leibniz by developing replies to the objections that Leibniz had raised against physical influx. For in order for Kant's replies to have any force against Leibniz, he could not appeal to doctrines that Leibniz would not accept. However, the resources Kant is drawing on seem irrelevant to the objections that Leibniz had raised against physical influx (e.g., against the idea of a literal transfer of accidents from one substance to another).

⁴⁷ The following reconstruction parallels that given by Langton, *Kantian Humility*, pp. 112–115.

Consider next a reconstruction that focuses rather on what might follow from the “mere existence” of two substances (a reconstruction that would thus not be subject to the objection raised against the previous argument):

If a substance must be “completely understood independently of any other substance,” then the only properties a substance could have essentially, that is, the only properties that a substance could have by means of its mere existence, would be intrinsic properties. However, if relations were reducible to intrinsic properties and two substances existed with their intrinsic properties, then it would be possible for the two substances to have a relational property essentially, namely in those cases where their essential intrinsic properties “reduced” a relational property. To block the unacceptable conclusion that a substance could have a relational property essentially, one would have to reject one of the premises. If one assumes that a substance can have only intrinsic properties essentially, the reducibility of relations must be rejected.

However, this argument is clearly fallacious as well. Even if one grants that both substances have their intrinsic properties essentially, it does not follow that either one of them would have the relational property that would reduce to their intrinsic properties *essentially*. For it would be possible for either one to exist without that relational property, namely in those cases where only one of the two substances exists at all. Rather than attribute either one of these fallacious arguments to Kant, it seems preferable to reject the assumption that Kant is attacking Leibniz.⁴⁸

If Kant is not attacking Leibniz’s position, then whose position is he attempting to refute with the principle of coexistence? In light of our discussion of the debate between proponents of pre-established harmony and physical influx discussed in Chapter 1, it is much more plausible

⁴⁸ Langton develops yet a different argument for the irreducibility of causal relations. Her idea is that the causal relations between two substances could be different, even if their intrinsic properties remained the same, citing the possibility that the laws of nature might make a difference to the kinds of effects intrinsic properties might have. (“In a world where the laws of nature were different, things might not have an attractive power, despite having the very same intrinsic properties that attractive things actually have,” *Kantian Humility*, p. 118.) Her evidence that Kant is making this argument stems from his considerations about “what God could or could not do” (*ibid.*) with respect to the laws of nature. However, in addition to the fact that the textual evidence Langton presents for this argument can also be read as a statement of what follows from rather than supports Kant’s argument, the argument appears to be question-begging in this context by assuming that the laws of nature are distinct from what follows from the intrinsic properties of substances, since Leibniz would not have thought that the laws of nature are distinct from what follows from the intrinsic properties of substances.

to think that Kant is targeting Crusius's position.⁴⁹ For Crusius introduced the notion of an existential ground and asserted that an existential ground could bring about effects in a substance distinct from itself. In other words, Crusius thought that the mere existence of a substance could connect it with other substances and thereby change their states. However, that is precisely what Kant is objecting to in the principle of coexistence.

As we saw, Crusius defines the world as a real whole, that is, as substances that stand in real connection with one another, where a real connection is established when a real ground in one substance posits a determination in another (and vice versa, which thereby keeps God from being a part of the world). That Kant agrees with this conception of the world can be seen by consulting the earliest preserved transcripts that we have of his metaphysics lectures, transcripts that Herder took down as Kant's student sometime between 1762 and 1764. Since Kant used Baumgarten's *Metaphysica* as a textbook in his lectures, their structure reflects the standard division of metaphysics, including sections on ontology, cosmology, and psychology. In the cosmology section of these transcripts, but, with some minor variations, in later transcripts as well, Kant explains:⁵⁰

§354 . . . The world is a real whole: all things in it stand in real connection.

The world is a whole which is not part of something else: otherwise this would be only a piece of the world. . . .

The world is therefore a (real) whole of actual things, which is not part of another.

357. All things are in real connection: they are conjoined in certain determinations, whichever those may be.

358. (In this world) the world is present, of which I am a part. There is a reciprocal connection, either mediately or immediately.

361. . . . All parts of the whole are as parts in real connection with one another as component parts: because they are grounds of the whole, and the whole cannot subsist without them. A part thus depends in some determinations on the others: consequently no part in the whole is independent – the whole [is] not independent – [but] contingent. (28:39–40)

While Kant accepts the general idea behind Baumgarten's definition of the world insofar as both stipulate that the world is a whole that is not

⁴⁹ The strategy of charting a middle course between Leibniz and Crusius is one that Kant pursues in several other pre-Critical works, including the *False Subtlety* (2:61) and the *Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality* (2:277 and 2:293).

⁵⁰ See, for example, 28:196, 29:850, 28:581, 28:657.

a part of another, he clearly sides with Crusius against Baumgarten in thinking that the world is a real rather than ideal whole.

However, despite his agreement with Crusius about how a world should be understood in terms of the notion of a real whole, Kant objects to the way that Crusius develops the idea of a real connection. For Crusius's division of real grounds into existential and active grounds implies that even existential grounds would connect substances in such a way that they would belong to one and the same world. In the principle of coexistence Kant objects, on essentially Leibnizian grounds, that a plurality of substances cannot form a single world due to their mere existence.⁵¹ Considered in itself, that is, considering its intrinsic properties alone, a substance does not stand in a real connection to other substances, since any given substance could belong to any one of a number of different worlds.⁵² In the passage just quoted from Herder's transcripts of Kant's metaphysics lectures, Kant makes the same point from a different angle. A substance can be a part of a real whole only if it depends in some respect on the other parts. That is, a substance that does not depend on others, an isolated substance, a substance considered merely insofar as it exists independently of any other, is not part of a common world.

To see the force of Kant's argument more clearly, consider the possibility that God creates two substances with their intrinsic properties. Do these two substances belong to one and the same world or to two distinct worlds? Kant's position is that they would belong to two separate worlds because none of their intrinsic properties entails a real connection between them. Moreover, none of them *could* entail such a connection, because that would violate Kant's understanding of substances as independently existing entities. In other words, if the intrinsic properties of one substance were to entail the existence of another substance, then the one substance would entail the existence of another substance and thus

⁵¹ Kant anticipates this point as early as the *True Estimation* when he argues: "A substance is either connected with and related to other substances external to it, or it is not. Because every self-sufficient entity contains within itself the complete source of all its determinations, it is not necessary for its existence that it should stand in any connection with other things. That is why substances can exist and none the less have no external relation to other substances, or have any real connection with them" (1:21–22).

⁵² By understanding Kant's argument in this way, one can see that he is not guilty of the fallacy Guyer attributes to him: "This argument surely turns on nothing other than simply equating considering a thing absolutely, in the sense of considering it in its ultimate reality, with considering it absolutely in the sense of *considering it in isolation*" (Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987), p. 352).

would be incapable of existing independently of the second substance. As Kant puts it, "there is nothing inhering in it which proves the existence of other substances distinct from itself" (1:413). For two substances that exist with only their intrinsic properties to belong to the same world, it is necessary that something connect them.⁵³

Another way to put the argument is to say that if God created the two substances as distinct worlds, then they would have no relational properties, whereas if God created them as belonging to a single world, then they would have relational properties. Kant is quite clear about such a distinction in the Herder lectures: "But is there only one **metaphysical world**? . . . Cannot the existence of a thing be thought without nexus? Can there not be single things that are not at all connected with these wholes? Can there not be wholes of series that stand in no connection with this world?" (28:40). What Kant's rhetorical questions are supposed to illustrate is that Crusius is committed to what Kant believes is an untenable position, namely that God could not create causally isolated substances, or, to be more precise, that God could not choose between creating a substance as either causally isolated or as standing in connection with others.

If two substances are thus not connected in a single world by means of their mere existence, as Crusius thought, then what kind of connection must be added so that they form a single world? Given the principle of succession, it is clear that Kant thinks that mutual interaction suffices, but it will be instructive for our discussion of Kant's reason for thinking that God is necessary for such a connection to see that the point rests not on a special fact about causality, but rather on a very general fact about the structure of substances. For in order for substances to belong to the same world, they must have some sort of "position" or "location" in the world in addition to the intrinsic properties God creates them with. Although "position" and "location" have strong spatial and temporal connotations for us, the point does not depend on the substances being either spatial or temporal. Rather, what is of concern is that substances can be created with all of their intrinsic properties in such a way that they have no relations to any other substance or in such a way that they are related to other substances.

⁵³ For an interesting discussion of how Leibniz attempts to deal with these intuitions in a sophisticated way (in part by distinguishing between expression and what is contained in the complete concept of a substance), see Donald Rutherford, *Leibniz and the Rational Order of Nature* (New York: Cambridge University Press, 1995), esp. pp. 185–197.

Not only can substances be created such that they relate to other substances, but they can relate to substances in different ways, where the differences do not depend exclusively on the intrinsic properties of the substances. For example, it would seem that God could create the two substances with all their intrinsic properties as standing in a variety of different relations, without thereby changing their intrinsic properties. If God creates two balls (of equal size, color, mass, etc.) either two feet apart or three feet apart, it would seem that God could do so without any difference in their intrinsic properties. The relations or, to put the point more precisely, the fact that substances are capable of relating and then do relate to each other in some particular way is thus something beyond the intrinsic properties of the substances and consequently must be added for the substances to be connected.⁵⁴ Kant's suggestion is that only God can be responsible for this feature of substances.

If Kant's argument against Crusius is persuasive, does it entail that he has, or at least would be required to have, an argument against Leibniz's doctrine of the reducibility of relations as well? After all, if substances are not connected by means of their mere existence, that is, in terms of their intrinsic properties, must it not be because their relational properties do not reduce to their intrinsic properties? Kant not only does not, but also need not present such an argument. The reason lies in the fact that Crusius and Kant both define a world differently from how Leibniz and his Wolffian followers understand that concept. Crusius and Kant define the world in terms of real connections, whereas Leibniz and Baumgarten think that not real, but rather "ideal influences"⁵⁵ connect substances that belong to the same world. Thus, all Kant is committed to establishing is that *if* the world is a real whole, then the mere existence of a substance does not establish a real connection to other substances. While even Leibniz could agree with this claim – simply by believing that the antecedent is false – Crusius cannot, and since Kant's target is Crusius, this is all he must prove. Thus, Kant does not need to establish what would

⁵⁴ Kant makes this point quite clearly: "But, since the reciprocal connection of substances requires that there should be, in the effective representation of the divine intellect, a scheme conceived in terms of relations, and since this representation is entirely a matter of choice for God, and can therefore be admitted or omitted according to His pleasure, it follows that substances can exist in accordance with the law which specifies that *they are in no place* and that they stand in no relation at all in respect of things of the universe" (1:415).

⁵⁵ See Leibniz's use of this term in, for example, §51 of "The Monadology" (Leibniz, *Philosophical Essays*, p. 219). Several Wolffians in eighteenth-century Germany, such as Baumgarten and Meier, pick up "ideal grounds" as well.

contradict Leibniz's position, namely that if the world is an ideal whole, then the intrinsic properties of a substance do not establish its ideal connection to other substances, that is, that the relational properties of substances reduce to their intrinsic properties.

God's Role in the Interaction of Substances

If Kant succeeds in refuting Crusius's position by arguing that substances do not stand in real or causal relations by means of their very existence, we can then turn to Kant's positive explanation of how substances stand in interaction. Kant claims that God is required in order to enable interaction, arguing:

But it does not follow from the fact that God simply established the existence of things that there is also a reciprocal relation between those things, unless the same schema of the divine understanding, which gives existence, also established the relations of things to each other, by conceiving their existences as correlated with each other. It is most clearly apparent from this that the universal interaction of all things is to be ascribed to the concept alone of this divine idea. (1:413)

Kant then explains this "divine idea" as follows:

The schema of the divine understanding, the origin of existences, is an enduring act (one calls it preservation); and in that act, if any substances are conceived by God as existing in isolation and without any relational determinations, no connection between them and no reciprocal relation would come into being. If, however, they are conceived as related in God's intelligence, their determinations would subsequently, in conformity with this idea, always relate to each other for as long as they continued to exist. That is to say, they would act and react, and the individual substances would have a certain external state. But if you abandoned this principle, no such state could exist in virtue of their existence alone. (1:414)

One might interpret Kant as claiming that the properties constituting the real connection between substances so that they form a single world depend exclusively on God, which would entail that for Kant "facts about intrinsic properties place no constraints at all on facts about causal powers."⁵⁶ One might be tempted into such an interpretation by thinking either that since relations do not reduce to intrinsic properties, intrinsic properties must be irrelevant to these relations or that since the choice of what causal relations to bring about is, as Kant puts it, "entirely a matter

⁵⁶ See Langton, *Kantian Humility*, p. 118. She repeats the same point on the next page: "Intrinsic facts do not constrain relational facts in any way."

of choice for God" (1:414), the intrinsic properties of things do not constrain God's decision as to what relational properties to bring about.

However, one might take certain considerations (both textual and philosophical) to suggest a different understanding of Kant's position, one according to which God "makes things interactive in the very act that makes them what they are."⁵⁷ As Kant states, "there is . . . a real reciprocal action between substances; in other words, there is interaction between substances by means of truly efficient causes. For the same principle that establishes the existence of things, also brings it about that they are subject to this law. And, hence, reciprocal interaction is established by means of those determinations which attach to the origin of their existences" (1:415). In other words, one might suggest that substances could not even exist or be what they are if God were not to place them in mutual interaction as well. For God not only causes the existence of things and their interaction, but he does so by means of the *same* principle, and if one and the same principle establishes the existence and the relational properties of a substance, then one might conclude that its relational properties cannot be divorced from its intrinsic properties.

Yet one can take up a middle position between these two readings. On the one hand, even if Kant does clearly think that the relational properties of substances do not reduce to their intrinsic properties, it does not follow that their intrinsic properties must be completely irrelevant to their relational properties. For their relational properties could depend in part, even if not entirely, on their intrinsic properties. For example, even if the attractions and repulsions of two substances are not reducible to their intrinsic properties – perhaps since the attractions and repulsions depend on an essentially relational property, namely the distance between them – it does not follow that an intrinsic property of the substances, such as their masses, could not also be relevant to what attractions and repulsions occur.⁵⁸

On the other hand, even if one claims that the *same* principle or divine concept is responsible for both the intrinsic and the relational properties

⁵⁷ See Karl Ameriks, "The Critique of Metaphysics: Kant and Traditional Ontology," in *The Cambridge Companion to Kant*, ed. P. Guyer (New York: Cambridge University Press, 1992), p. 262. Ameriks also stresses that "the 'external' changes of a thing, its interactions with other things, are *just as* immediately attributable to it as any internal changes" (*ibid.*).

⁵⁸ The fact that what relational properties a substance has is "a matter of choice for God" does not immediately imply that its relational properties do not depend on its intrinsic properties, since which intrinsic properties a substance has is a matter of God's choice as well.

of a substance, it does not immediately follow that intrinsic and relational properties have exactly the same status in every regard. One could still say, for example, that the intrinsic properties (or perhaps only a subset of them) are essential, while the relational properties are accidental. Moreover, Kant's claim that the same principle causes both sets of properties is clearly meant to distinguish his own position from occasionalism. For after claiming that his position is distinct from Malebranche's occasionalism, he explains that "the same indivisible act, which brings substances into existence and sustains them in existence, procures their reciprocal and universal dependence, so that the divine act does not need to be determined, now one way, now another, according to circumstances" (1:415). That is, whether God creates an intrinsic or a relational property in a substance, God is not determined to do so because of some change in the world. Rather, both the intrinsic and the relational properties in the world are the result of God's constant and indivisible act of creation rather than his miraculous intervention. It is primarily in this sense that intrinsic and relational properties follow from the same principle in God and therefore have the same status. Yet this sense is consistent with relational properties depending at least in part on intrinsic properties.

But what is Kant's argument for claiming that only God can enable relational properties such as mutual interaction between substances? If the intrinsic properties of a substance, that is, the properties that the substance has in virtue of its "mere existence," regardless of what other substances there might or might not be and thus regardless of what world it might belong to, do not entail which relational properties the substance has, then something in addition to the substance must be partly responsible for whatever relational properties the substance is to have. It cannot be in the substance with which the first substance is supposed to stand in mutual interaction, because it is the possibility and the efficacy of that ground that need to be explained in the first place. In other words, simply adding a second substance with its intrinsic properties would be insufficient to ground the relational properties of the substances, because intrinsic properties taken in isolation cannot ground relational properties. To revert to the case of bodies endowed with attractive and repulsive forces – which Kant himself develops at 1:415 of the *Nova dilucidatio* as well as in Propositions VI–X (1:480–485) of the *Physical Monadology* and in section two, chapter 1 (1:264–268) of the *Universal Natural History* – even if one were to posit the "mere existence" of two such bodies, one would still not be able to explain how they are related and can stand in mutual interaction. What is needed is the property of distance, but of

course distance, according to Kant (at 1:414 and, indirectly, at 1:415), is not an intrinsic property of a substance. If the substances that are to stand in mutual interaction are not sufficient to ground their relations or their mutual interaction, one must appeal to something else.

Kant's claim is that only what is the "common cause" of the two substances would suffice to help ground the relational properties of substances. In his "demonstration," however, Kant seems almost to presuppose this point rather than argue for it:

Since, therefore, in so far as each individual substance has an existence which is independent of other substances, no reciprocal connection occurs between them; and since it certainly does not fall to finite beings to be the causes of other substances, and since, nonetheless, all the things in the universe are found to be reciprocally connected with each other – since all this is the case, it has to be admitted that this relation depends on a communality of cause, namely on God, the universal principle of all beings. (1:413)

Even if it is clear that the finite substances that stand in mutual interaction are not in a position to make that interaction possible, why go so far as to think that what makes the interaction possible must be the cause of the very existence of the substance as well?

Kant's answer depends on the fact that the relational properties of substances concern the very structure of a substance and therefore that whatever causes them must also be the cause of the existence of the substances insofar as what causes the existence of a substance would also cause its basic structure. In what sense do relational properties concern the very structure of a substance? As we saw, not only do relational properties go beyond intrinsic properties, but it is also the case that whether a substance will have relational properties at all (and if so, which ones it will have) is a fundamental question, affecting what a substance can cause in other substances, just as its intrinsic properties do. Thus, in this sense, too, the relational properties of substances are just as fundamental as their intrinsic properties.⁵⁹ Yet if the intrinsic properties of substances must be determined by what causes the existence of the substances (insofar as they would follow immediately from the immutable grounds that are posited when it is posited), then it would stand to reason that what is equally fundamental, namely relational properties, would be determined by what causes the existence of the substances as well.

⁵⁹ This may be the sense in which Ameriks asserts that relational properties are just as primitive as intrinsic properties.

Kant returns to this complex of issues in the Herder lectures, though it is, again, in the context of criticizing Crusius's understanding of the world as a real whole:

To the connection of the things in the whole belong **not merely the existences of the things**. With this the question still remains whether they constitute a whole, since each could exist alone. If they are many and coexistent, then they do not immediately have community on that account. Thus for the conjunction *something special, reciprocal action*, is yet required. For two substances without connection do not possibly have any effect on each other: without connection nothing that takes place in A can have a consequence in B. Thus if a substance in its existence does not depend upon another, then substances could exist without connection: when two *substances effect each other, then A and B must necessarily depend upon C*, otherwise nothing in existence could follow in B from A: but from the fact that their existence depends upon a third: it does not yet follow that they must be in connection: *their connection still requires a special ground: a special action still of the creator, since he connected them. Thus the state of different substances in which one has an effect on the other and suffers* (interaction) thus has a *special ground in God, who willed that they should depend upon one another.* (28:51)

While this statement is still sketchier than one might like, it does clearly suggest that Kant continues to accept the kind of argument that he seemed to be developing in the *Nova dilucidatio*. Moreover, as the argument continues, Kant adds important details as follows:

If a substance suffers, then it must contain in *itself* by its own power the ground of the inherence of the accident, because the accident otherwise would not inhere in it. But the ground of this must also be *in the efficient power of the substance*, because it otherwise would not have an effect, consequently *the powers of the substances are harmonious*. One, in respect of the powers of the other, contains the ground of the inherence of the accident. This doctrinal edifice is called established harmony, and since God willed it previously, pre-established

Synthetic preparation. Each subject in which an accident inheres must itself contain a ground of the inherence. For if, e.g., God could produce a thought in a soul merely by himself: then God, but not a soul, would have the thought: because there would be no connection between them. . . .

If two substances have a reciprocal effect on one another: then the suffering, the *inherence of the accident*, happens *not merely by its own but rather also by alien power*: for otherwise it would *not be suffering*. . . .

An accident thus inheres by its own power, which contains the *sufficient inner ground* of it yet also by *alien* power, as by *an outer ground* of the inherence without which it would not have inhaled. Now *no substance can really contain the ground of the accident in the other, if it does not at the same time contain the ground of the substantial power and of the existence of the other*. . . .

If a substance is active by its own power under an external condition, then it suffers.

§450. If we want to conceive that one power simply suffers from the other, without its own power and thus without harmony, then that is called physical or real influx. . . . But what explains this connection? Since suffering always requires one's own power: then this influx is impossible, even for God. (28:51–53)

Although it is particularly difficult in this passage to discern what Kant is repeating from Baumgarten and what is his own critical commentary on it, one can see Kant making two points more clearly here than in the *Nova dilucidatio*. First, Kant makes it quite explicit that to explain the inherence of an accident in a substance, one must posit not only a second substance to cause that accident, but also a power in the substance that “suffers.” That is, in addition to the causal efficacy of a second substance, a substance requires a sufficient inner ground for an accident to inhere in it, because there must be some connection between the substance and the accident in order to make the accident be an accident *of that substance*.⁶⁰ Even God cannot produce an accident in a substance without a corresponding power in the substance, since God cannot cause, for example, a thought in a nonsentient creature.⁶¹ God can, of course, cause a thought, but then he, not the creature, would be thinking the thought.⁶² Thus, at least two causally efficacious substances are required for an accident to inhere in a substance.

Second, since there are two causally efficacious substances, something must guarantee that (1) they can act on each other at all and that (2) their powers act in harmony. Since no causal interaction follows from the “mere” existence of substances as such (namely as isolated entities), a ground is required for causal interaction to occur at all (establishing (1)). However, this initial argument does not immediately establish that this ground could not be a being less perfect than God. In addressing the issue of harmony, Kant seems concerned to eliminate the possibility that one substance exercise its (active) power to cause effect A while another substance exercises its (passive) power to exemplify effect B, where A and

⁶⁰ Although Kant calls this inner ground sufficient, he must mean by that not that it is sufficient as a ground (which would imply that no other grounds are necessary) but rather that it is sufficient as an *inner* ground (which implies merely that no other inner grounds are necessary).

⁶¹ Kant also expresses his commitment to this point in the *Negative Magnitudes*: “And an inner accident, a thought of the soul, cannot cease to be without a truly active power of exactly the *same* thinking subject” (2:191).

⁶² Karl Ameriks develops this point in his article “The Critique of Metaphysics: Kant and Traditional Ontology,” where it goes under the guise of the “Restraint Argument” (p. 263).

B are not identical, thus creating what one might call “causal discord.” In other words, because any change in the intrinsic properties of a substance is possible only due to change in its relational properties and because any change in the relational properties of substances can be brought about only through the joint involvement of both substances, they must act harmoniously in bringing about an effect that causes changes in their intrinsic properties, that is, if they are to bring about what Kant calls in the principle of succession “reciprocal changes” (as in the case of motion).⁶³ Only something with control over these substances’ powers would be in a position to guarantee their joint production of a reciprocal change, but since no finite substance can contain the ground of the substantial power of any other finite substance, only an infinite substance, that is, God,⁶⁴ is in a position to harmonize the causal connections between substances (establishing (2)). Not only is something needed to connect substances, given that they are not connected by means of their “mere existence,” but if substances have different powers, then something is also needed to make sure that the different powers act harmoniously, that is, in such a way that they can *jointly* bring about a reciprocal change. While something less powerful than God might be able to establish some sort of connection between substances, only God can establish that a harmonious connection arises, since only God can control what powers substances have. Thus, these passages from Herder’s transcripts of Kant’s metaphysics lectures represent an important clarification of Kant’s argument in the *Nova dilucidatio* that causal connections between substances are possible only with God’s assistance.

Physical Influx, Pre-established Harmony, and Occasionalism

If Kant’s principles of succession and coexistence thus represent the fundamental features of his pre-Critical theory of causality, it might seem completely unambiguous that Kant is developing a version of physical

⁶³ In a sense, Kant is simply taking what underlies Leibniz’s position one step further. Leibniz thinks that the states of all substances must harmonize with each other, despite the fact that substances are completely independent of each other and do not act on each other causally. As a result, God is responsible for ensuring that substances are set up such that their states harmonize. Kant agrees that the states of independent entities must harmonize with each other, but because he holds that substances can act on each other, he realizes that the actions that bring about the harmonious states of substances must be in harmony as well.

⁶⁴ Kant may have taken for granted at this point – on the basis of his argument in *The Only Possible Argument* at 2:83–85 – that there can be only one infinite substance: God.

influx. At the same time, Kant is not simply repeating what earlier advocates have said, since his view differs from theirs in a variety of ways. He rejects Crusius's real existential grounds, argues for the necessity of the causal efficacy of two substances in mutual interaction, and requires the assistance of God in explaining how such interaction is to be possible.

However, at the end of his discussion of the principle of coexistence in the *Nova dilucidatio*, Kant might seem to be making assertions that call into question the idea that he accepts physical influx at all (much less a unique version thereof). For he explains how his view relates to the three traditional causal theories as follows:

Physical influence, in the true sense of the term, however, is excluded. There exists a universal harmony of things. Nonetheless, this does not give rise to the well-known *Leibnizian pre-established harmony*, which is properly speaking *agreement* between substances, not their reciprocal *dependency* on each other. For God does not make use of the craftsman's cunning devices, carefully fitted into a sequence of suitably arranged means designed to bring about a concord between substances. Nor, moreover, is there an ever special influence of God, that is to say, an influence through which the interaction of substances is here established by means of *Malebranche's occasional causes*. For the same indivisible act, which brings substances into existence and sustains them in existence, procures their reciprocal and universal dependence, so that the divine act does not need to be determined, now one way, now another, according to circumstances. (1:415)

Needless to say, this passage is not what one would have expected, given what Kant has argued for at length in the *Nova dilucidatio*. For at first glance, he seems to be rejecting *all three* traditional causal theories. To understand what Kant is trying to do, it is helpful to consider his objections to each of the traditional causal theories in more detail.

Kant's objection to occasionalism is that it would require two separate acts, one by means of which God creates the existence of substances and another by means of which God is determined, presumably by (the intrinsic properties of) finite substances, to bring about ever-changing relations between them. Kant does not specify fully what is objectionable about such a view. On the one hand, Kant may be objecting to the idea that God would have to perform two acts rather than one in determining the existence and properties of substances. If, however, one distinguishes two acts, then the second act might appear to be a kind of continuous miraculous intervention and thus Kant's objection would collapse into one of Leibniz's standard objections to occasionalism. On the other hand, Kant may be objecting to the idea that finite substances might determine or, to put it more starkly, cause God to do different things at different

times, an idea that might be thought incompatible with the immutability of God or with the idea that God is in no way acted on by his creatures. Whichever of these objections Kant has in mind, he is clearly distancing himself from occasionalism.

Unsurprisingly, Kant does not describe his theory as a version of pre-established harmony either, though in this case his reasoning is of necessity subtler and more refined, since there might seem to be a way of understanding Kant's position that would locate it much closer to pre-established harmony than is typically the case. For not only is Kant willing to accept the idea that substances harmonize with each other, but the idea is also quite crucial to one of the most distinctive features of his theory, namely the necessity of God in enabling harmonious relations between substances. Kant also does not object to the idea that the harmony that exists between substances would be pre-established. For in the second section of the *Nova dilucidatio*, Kant explicitly accepts divine foreknowledge (1:405), even of the future free acts of human beings, which seems to imply that Kant thought he could deal satisfactorily with the temporality of the harmony between substances.

So how is Kant's position different from Leibniz's? Though one might expect Kant to reiterate the content of the principle of succession, namely that only causally connected substances can undergo change, he says rather that "this does not give rise to the well-known *Leibnizian pre-established harmony*, which is properly speaking *agreement* between substances, not their reciprocal *dependency* on each other" (1:415). In short, Kant repeats Crusius's objection, discussed in Chapter 1. But in explaining this point, Kant also takes recourse to a point that we had seen Knutzen making against pre-established harmony, namely that God acts in the shortest, most natural way. Kant puts the point as follows: "For God does not make use of the craftsman's cunning devices, carefully fitted into a sequence of suitably arranged means designed to bring about a concord between substances" (1:415). Unfortunately, Kant does not expand on the differences between the notion of a real causal connection that Crusius and he accept and the merely ideal one that Leibniz and Baumgarten adopt, but rather seems content, at this point, simply to follow Crusius's and Knutzen's criticisms.

What should we make of Kant's criticism of the ideality of the relations between substances that are to form a world, that is, of their "mere agreement" rather than "real dependency"? Is there genuine support for Kant's point in our common-sense understanding of what a world is and is that support significant enough to amount to a reason for understanding

the world as a real whole rather than an ideal one? If Crusius's criticism that ideal relations are insufficient to distinguish possible worlds from each other can be met, then it would seem that the main motivation for thinking of the world in terms of real rather than ideal connections stems from differences in the phenomena that Leibniz and Kant are attempting to explain. As an idealist, it is natural for Leibniz to think that ideas, as what are ultimately real, should suffice to explain what unites minds into a world. For Kant, however, who is primarily a natural philosopher interested in explaining motion and change and who is following philosophers (such as Wolff, Knutzen, and Baumgarten) not committed to idealism in Leibniz's sense (e.g., insofar as they either could or do accept physical points as the ultimate constituents of reality), it is much more natural to think of the world as a system of bodies that really act on each other in their collisions by means of the force of impenetrability and in their attractions through gravitational pulls.

If Kant thus distances his own position from both occasionalism and pre-established harmony, what is one to make of his rejection of what would appear to be the only remaining option, namely physical influx? Appreciating the historical context is crucial to understanding what Kant is attempting to say here. Since Kant's view is significantly different from that of previous advocates of physical influx, it makes sense that he would see the need to distance himself from their views. Accordingly, Kant does not reject physical influx *per se*, but rather physical influx "in the true sense of the term." Although Kant does not clarify this restriction, he never asserts that accidents might literally be transferred from one substance to another, as the term "influx" or "*influeret*" might suggest. A few lines later, Kant again remarks that his view, "the system of universal interaction of substances, constituted in this way, is certainly somewhat superior to the popular system of *physical influence*, for the former, to be sure, reveals the very origin of the reciprocal connection of things; and this origin is to be sought outside the principle of substances, considered as existing in isolation. And, in this respect, that threadbare system of efficient causes could not be further from the truth" (1:415–416). Here Kant comes as close to identifying the version of physical influx that he is distancing his own from as he can without naming names. For the "system of efficient causes" (*systema causarum efficientium*) Kant refers to here corresponds exactly to the title of Knutzen's treatise that makes an extended case for a particular version of physical influx. And his complaint against Knutzen's account is that it does not "reveal the very origin of the reciprocal connection of things," that is,

does not explain how it is that God is necessary for substances to interact in the first place by creating them such that they can belong to a single world.⁶⁵

Confirmation that Kant is rejecting only a particular version of physical influx can be found in the Herder lectures as well: "If we want to conceive that one power suffers simply from the other, without its own power and thus without harmony, then that is called physical or real influence. . . . Since one's own power to suffer is always required, this influence is impossible, even [for] God" (28:53). Later, in his Inaugural Dissertation (1770), he characterizes what should count as the unacceptable version of physical influx in a slightly different way as follows:

If a plurality of substances is given, *the principle* of a possible *interaction* between them *does not consist in their existence alone*, but something else is required in addition, by means of which their reciprocal relations may be understood. For they do not necessarily relate to anything else simply in virtue of their subsistence, unless, perhaps, they relate to their cause. But the relation of caused to cause is not interaction, but dependence. Therefore, if any interaction should occur between them and outer things, a special ground, which determines this interaction precisely, will be needed.

And it is in this, indeed that the *πρωτον ψευδος* of the theory of physical influx, in the vulgar sense of that term, consists. It rashly assumes, namely, that there is an interaction of substances and transeunt forces, which can be cognised by means of their existence alone. . . . If we free this concept from that blemish, we have a kind of interaction, which is the only one which deserves to be called real, and, in virtue of which, the whole, constituted by the world, deserves to be called real, rather than ideal or imaginary. (2:407)

Here Kant seems to be describing Crusius's version of physical influx in terms that are strongly reminiscent of Kant's argument for the principle of succession (which supports the idea that it is Crusius rather than Leibniz whom Kant is targeting there). Regardless of which theory Kant identifies as the vulgar version of physical influx, given that his view is both different from it and, in his eyes, superior, one can see why he might think that it would be misleading to say that he accepts physical influx without any explicit qualification. Further, given that Kant stresses the harmony between substances in an attempt to win over those who might have been attracted in certain ways to pre-established harmony, it would be natural

⁶⁵ While the fact that Kant is criticizing Knutzen's position here is consistent with Kuehn's view that the early Kant was critical of Knutzen, it also suggests that Kant does not reject every aspect of Knutzen's position and thus that the differences between their positions may not be as great as Kuehn seems to suggest.

for Kant to want to highlight his accomplishments by representing his view as an exciting new alternative to all three traditional theories of causality. While one can agree that Kant's view departs in significant respects from other proponents of physical influx and admire the fact that Kant has attempted to incorporate the Leibnizian desire for harmonious relations between substances into his account, it should nonetheless be clear that, seen from within the philosophical framework of the three traditional theories of causality, Kant's view is ultimately a version of physical influx given his assertion that finite substances act on each other causally.

KANT'S PRE-CRITICAL REACTION TO HUME

The Implication of Hume's Position

If Kant's pre-Critical theory of causality takes the form of the novel version of physical influx just described, how should we understand Kant's reaction to Hume's position in the *Inquiry*, which he became acquainted with some time after it was translated into German in 1755?⁶⁶ Recall that the main difference between Kant's theory of causality and Leibniz's pre-established harmony is that Kant asserts that one substance causes a change in another substance, whereas Leibniz denies that a finite substance could act on another finite substance, asserting that a substance can act only on itself. In short, Kant asserts, while Leibniz denies, causal relations between substances. Further, since Kant takes grounds to be akin to logical principles, he believes that a cause is necessarily connected with its effect. Accordingly, Kant's version of physical influx commits him to the claim that there is a necessary connection between two substances when they act on each other causally.

Kant would thus find Hume's views relevant to his own position insofar as Hume claims in an especially clear and forceful way that distinct entities cannot be related by a necessary connection, which is precisely what Kant is affirming in asserting that substances are necessarily connected by means of their causal relations to each other. Putting the point in terms closer to Hume's own in the *Inquiry*, matters of fact can never be necessary because their denial would not violate the principle of contradiction. As a result, one substance cannot be necessarily connected to another because denying the existence of the second after the first has

⁶⁶ The translation is *Philosophische Versuche über die menschliche Erkenntniß von David Hume* (Hamburg and Leipzig, 1755).

been posited would not violate the principle of contradiction. Hume, of course, prefers to speak of events, matters of fact, or objects rather than substances, but that would not have obscured for Kant the relevance of Hume's claim to his own theory. In truth, the fact that Kant accepts substances and understands them as self-sufficient entities capable of independent existence only makes Hume's point more pressing. How can substances be necessarily connected to each other if they are supposed to be completely self-sufficient?

It is also important to note that Hume's point is a problem for Kant only because Kant rejects pre-established harmony in favor of physical influx. That is, Hume's critique of necessary connection need not present a difficulty for Leibnizian pre-established harmony. For by asserting that a finite substance can act only on itself, a Leibnizian is in no way committed to claiming that two distinct entities are necessarily connected. If a Leibnizian accepts the idea that a cause is necessarily connected with its effect, then pre-established harmony would entail only a necessary connection between various states *within* a substance. Yet necessary connections within a substance are not obviously problematic in the way that necessary connections between substances are, since what Leibnizians think is self-sufficient and thus independent is not the states of a substance, but the substance itself (along with all of its states). Thus, Hume's views would not present the same challenge to Leibnizians as they would to Kant and other proponents of physical influx.

In fact, Hume and Leibnizian proponents of pre-established harmony could even form a united front against proponents of physical influx such as Kant. For Leibnizians could plausibly claim that the point Hume is putting in his own distinctive way is one that they have been using as support for their position all along.⁶⁷ While Hume rejects thinking of the world in terms of Leibnizian substances, what underlies his critical point about causality, from a Leibnizian perspective, is simply the idea that distinct existences (i.e., for the Leibnizian, substances) cannot be related by means of a necessary connection. If two things are substances, that is,

⁶⁷ In commenting on the development of Kant's thoughts on causality in his pre-Critical period, Lewis White Beck points out that "most of the Humean ideas in [*Dreams of a Spirit-Seer*] were very much in the air in Germany, and it would be difficult or impossible to trace them to one source" ("A Prussian Hume and a Scottish Kant," in *Essays on Kant and Hume* (New Haven: Yale University Press, 1978), p. 115). Whether such an undertaking is in fact impossible or not, it suggests that at least some of Hume's points could have been welcomed as independent support for positions developed prior to the arrival of Hume's work in translation.

independently existing, distinct entities, then Leibnizians, too, will infer they cannot be necessarily connected. As we have seen in Chapter 1, Leibniz himself goes further than this by claiming that physical influx is metaphysically incoherent insofar as it is committed to the idea that accidents would have to be transferred from one substance to another but that criticism may derive a considerable portion of its force from the point that substances and substances alone are capable of independent existence.

Real versus Logical Grounds

To see that Kant is aware of the relevance of Hume's critique of causality for his own account, it is important to consider how and when he draws a distinction between real and logical grounds. As we saw above, Kant had already explicitly accepted grounds as an integral part of his account as early as 1755 in the *Nova dilucidatio*. However, in that work Kant does not describe grounds as real. Starting around 1762 – presumably after having read Hume's *Inquiry* in translation – and continuing up throughout the rest of his pre-Critical period, Kant draws a distinction between logical and real grounds and makes real grounds into a fundamental feature of his metaphysics as he comes to see how important they are in providing an adequate account of a series of metaphysical issues.

Let us begin by considering the way in which the distinction appears in the Herder transcripts from Kant's lectures on metaphysics, *Reflexionen* that date from that time period, and *The Only Possible Argument* (1763). We can then turn to Kant's most direct characterization of real grounds in his *Attempt to Introduce the Concept of Negative Magnitudes into Philosophy* (1763), which will allow us to see how real grounds are pivotal to Kant's immediate response to Hume.

In the Herder transcripts, Kant introduces the distinction between logical and real grounds as follows:

A *ground* is thus something by which, having been posited, something else is posited. *Crusius* describes the ground through which something is brought about. The word *bring about* is much too composite: for not all effects are consequences and not all powers [a] ground.

Every *ground* is either logical, by means of which the consequence that is identical to it is posited as a predicate according to the rule of identity, or *real*, by means of which the consequence that is not identical to it is not posited according to the rule of identity. (28:11)

Baumgarten's definition of ground (*ratio*) in §14 of his *Metaphysica* makes no such distinction, so it is clear that Kant is the source of this distinction. Doctrinally, it is also clear that Baumgarten, like Leibniz, thinks that,

logically speaking, all predicates must be contained in the subject in order to be true, even if the connection is perceived only confusedly. By implicitly suggesting that the predicate (consequence) is not identical to (or contained in) the subject (ground), Kant is departing from Baumgarten in introducing this distinction.

The transcript continues by noting that the distinction between real and logical grounds is not the same as the distinction between ideal and real grounds, which he attributes to Crusius:

Crusius divides grounds into ideal and real grounds. This division is entirely different. E.g., the world is the ideal ground of God. For the ideal ground is merely the ground of knowing. They are thus subordinated to each other so that a real ground can at the same time be an ideal ground: but no real ground can be a logical ground and vice versa. For they are precisely opposed. (28:12)

While a real ground can be an ideal ground insofar as one can, at least in principle, know something as a real ground (though perhaps not in particular cases, as Kant's example illustrates), a real ground cannot be a logical ground, because a real ground is defined in opposition to the principle of contradiction. It is thus clear that the notion of a real ground does not stem from either Baumgarten or Crusius, making it quite plausible to think that it originates with Kant.

Kant continues in his lectures with a series of remarks that further spell out his concept of a real ground. For example, he distinguishes between logical and real repugnance, or opposition, noting that logical repugnance is impossible, since that would amount to a contradiction, whereas real repugnance is possible, resulting merely in a privation. Kant also notes: "The connection between the logical ground and consequence is comprehensible, but not that between the real ground, that when something is posited, something else would be posited at the same time: example: God wills! – There became the world!" (28:12). In other words, we cannot rationally grasp the connection between a real ground and its effect. That is, we cannot infer solely on the basis of reason or the principle of contradiction that an effect necessarily follows from its real ground. Later, Kant seems to go even further by claiming that "every determination of things, however, that requires a real ground, is posited by something else, and the connection of a real ground with the real consequence is thus not comprehended [*eingesehen*] from the rule of identity, also cannot be expressed by a judgment, but is rather a simple concept. . . . Only through experience can we have insight [*einsehen*] into the connection of the real ground, not logically" (28:24). Here we see Kant thinking of the concept of a real ground as being something

primitive and simple that cannot be understood on the basis of reason alone, but rather requires experience.

In a series of other passages in the Herder transcripts, Kant incorporates real grounds both into his criticisms of Baumgarten and into his own positive account. First, in commenting on Baumgarten's principle that no action is without counteraction (or reaction), Kant remarks:

But the word reaction, which the author uses here merely for reciprocal action, implies the concept of the *against*, of real opposition, a consequence of which is privation. And this really opposed action is not general, for actions which are really opposed must be homogeneous. Thoughts cannot be cancelled by motions, for motion is posited and cancelled merely by motive power and thus if the body reacted in this manner on the soul, then willing and moving would have to be the same. (28:45)

Here Kant seems to maintain implicitly that real grounds divide up into kinds and each kind of real ground can act only on substances having the same kind of ground. Kant uses this point to claim that thoughts and motions do not have the kind of homogeneity required for one to act on the other causally. While this explanation suggests that Kant may have rethought his position on the mind-body problem since the *True Estimation*, it supports the idea that the target of Kant's principle of coexistence is Crusius, since not only do substances not stand in interaction by means of their mere existence, but "causal discord," that is, a lack of fit between the types of real grounds in each substance, can preclude causal interaction as well.

Moreover, Kant uses real grounds in his positive explanation of the connection between a substance and its accidents:

§192. The respect of a substance to its inherent accidents is the real ground, or power [Kraft]: consequently, the grounds of inherence are: real grounds. . . .

196. With the concept of substance a subject is taken together with all of its inhering accidents. Yet these accidents must have their real ground in the substance. Distinguishing this from one another constitutes the substantial and the essential. The *substantial* contains the first real ground of all inhering accidents. It is not [a power], but rather has a power: we can never have insight [*einsehen*] into this first real ground. The *essential* contains the first logical ground. (28:25)

Kant now uses the concept of real grounds to explain the relationship between a substance and its own accidents, not merely for explaining the relationship between a substance and accidents in another substance. Yet he also uses it to distinguish between the substantial and the essential.

Kant also finds the distinction between logical and real grounds to be of help in attempting to attain an adequate grasp of God's existence. In

Reflexion 3725 (dated ε to ζ , thus prior to 1764) Kant shows how real grounds apply to God's existence as follows:

Absolute necessity is either logical: on account of the principle of contradiction, or real: not on account of the principle of contradiction.

The former is the necessity of judgments. Or the necessity of the relation of the predicate and the subject.

The latter is the necessity of the beings. 1. God is omnipotent. 2. God exists. The latter cannot be known (in itself) through the contradiction of opposites. The opposite of existing is not being. But not being, alone, does not contradict itself. Existence is not a predicate, therefore its opposite is not a predicate opposed to anything. (17:270)

Since Kant holds that no logical contradiction arises in claiming that God does not exist (given that existence is not a predicate and thus cannot contradict the concept of God), the necessity of God's existence cannot be established (as the ontological argument attempts to) by means of considerations pertaining to logical grounds.

In *The Only Possible Argument* Kant extends his use of considerations pertaining to real grounds to prove God's existence. Kant's proof rests on a distinction in the concept of possibility between formal and material elements. To be possible, not only must something not violate what he calls the formal element of possibility, namely the principle of contradiction, but a material element, namely the concepts that are involved, must also be given so that the principle of contradiction can be applied so as to establish the formal consistency of what is to be possible. As Kant puts it, "in every possibility, we must first distinguish the something which is thought, and then we must distinguish the agreement of what is thought in it with the law of contradiction" (2:77). On the basis of this distinction Kant then argues that "the actuality, by means of which, as by means of a ground, the internal possibility of other realities is given, I shall call the first real ground of this absolute possibility, the law of contradiction being in like manner its first logical ground, for the formal element of possibility consists in agreement with it" (2:79–80). In the ensuing argument, Kant maintains that only God could be the first real ground of possibility and must therefore exist if anything is to be possible at all. Thus, real grounds are crucial for Kant's innovative argument for the existence of God in his pre-Critical period.⁶⁸

⁶⁸ Although Kant develops in very abbreviated form a similar argument for the existence of God in his *Nova dilucidatio*, he does not use the concept of a real ground, which is thus further indirect evidence for the idea that it was Kant reading Hume's *Inquiry*

Real Grounds as a Response to Hume

While Kant thus applies the concept of real grounds to a wide variety of different issues in the early 1760s, it is the primary focus of his *Attempt to Introduce the Concept of Negative Magnitudes into Philosophy* (1763). After introducing the concept of negative magnitudes in terms of real opposition and illustrating it through examples taken from mathematics (in the first section) and philosophy (in the second section), Kant turns to developing philosophical principles for real opposition. Kant concludes his account with the following general remark, which, due to its significance for his understanding of the relevance of Hume to his own pre-Critical theory of causality, will be quoted in full:

I fully understand how a consequence is posited by a ground in accordance with the rule of identity: analysis of the concepts shows that the consequence is contained in the ground. It is in this way that necessity is a ground of immutability; that composition is a ground of divisibility; that infinity is a ground of omniscience, *etc., etc.* And I can clearly understand the connection of the ground with the consequence, for the consequence is really identical with part of the concept of the ground. And, in virtue of the fact that the consequence is already contained in the ground, it is posited by the ground, in accordance with the rule of agreement. But what I should dearly like to have distinctly explained to me, however, is how one thing issues from another thing, though not by means of the law of identity. The first kind of ground I call the logical ground, for the relation of the ground to its consequence can be understood logically. In other words, it can be clearly understood by appeal to the law of identity. The second kind of ground, however, I call the real ground, for this relation belongs, presumably, to my true concepts, but the manner of the relating can in no wise be judged.

As for this real ground and its relation to its consequence my question presents itself in the following simple form: How am I to understand **the fact that, because something is, something else is?** A logical consequence is only really posited because it is identical with the ground. Human beings are capable of error: the ground of this fallibility is to be found in the finitude of man's nature, for if I analyse the concept of a finite mind, I see that fallibility is to be found in it. In other words, I recognise that fallibility is identical with what is contained in the concept of a mind. But the will of God contains the real ground of the existence of the world. The will of God is something. The world which exists is *something completely different*. Nonetheless, the one is posited by the other. The state of mind in which I hear the name *Stagirite* is something, and it is in virtue of that something that something else, namely my thought of a philosopher, is

sometime after 1755 that prompted him to think about how grounds might help him in articulating a coherent metaphysical position.

posited. A body A is in motion; another body B, lying in the direct path of A, is at rest. The motion of A is something; the motion of B is something else; and yet the one is posited by the other. Now, you may subject the concept of divine willing to as much analysis as you please: you will never encounter in that concept an existent world as something which is contained with the concept of God's willing, or as something posited by that concept through identity. Likewise in the other cases. Nor am I willing to be fobbed off by the words "cause" and "effect," "force" and "action." For if I already regard something as a cause of something else, or if I attach the concept of force to it, then I am already thinking of the cause as containing the relation of the real ground to its consequence. (2:202–203)

After explaining how his distinction between real and logical grounds is distinct from Crusius's distinction between real and ideal grounds – a distinction also discussed in the context of the Herder transcripts – Kant concludes the *Negative Magnitudes* as follows:

Now, let the attempt be made to see whether real opposition in general can be explained. Let us see whether we can offer a distinct explanation of how it is that, *because something is, something else is cancelled*, and whether we can say anything more than I have already said on the matter, namely that it simply does not take place in virtue of the law of contradiction. I have reflected upon the nature of our cognition with respect to our judgment concerning grounds and consequences, and one day I shall present a detailed account of the fruits of my reflections. One of my conclusions is this: the relation of a real ground to something, which is either posited or cancelled by it, cannot be expressed by a judgment; it can only be expressed by a concept. That concept can probably be reduced by means of analysis to simple concepts of real grounds, albeit in such a fashion that in the end all our cognitions of this relation reduce to simple, unanalysable concepts of real grounds, the relation of which to their consequences cannot be rendered distinct at all. (2:203–204)

These passages are extremely important for understanding Kant's complex relationship toward Hume.⁶⁹ For in them Kant indicates that, in 1763, he not only recognizes the problem he sees Hume as presenting, but believes that he is able to present a partial solution to it.⁷⁰ What the first passage makes clear, as I read it, is that Kant sees an immediate connection

⁶⁹ We return to the complexities of Kant's relationship to Hume in Chapter 6.

⁷⁰ While Crusius is undoubtedly of great importance to Kant during this period – the interpretation presented above of the *Nova dilucidatio*'s principle of coexistence in fact emphasizes Crusius's position – I am not convinced by Beck's suggestion ("A Prussian Hume and a Scottish Kant," p. 114) that it was Crusius rather than Hume who precipitated this line of thought. Rather, I would suggest that Kant initially *sees* the problem due to his reading of Hume and that Crusius is important in this particular context insofar as

between the challenge that Hume's view presents for his own theory of causality and real grounds. Hume asks, and Kant reiterates the question: How can one infer from the existence of one thing to the existence of another? Like Hume, Kant also recognizes quite clearly that the inference cannot be justified by means of the principle of contradiction and thus sees the need to introduce a new principle to warrant the inference. Unlike Hume, who thinks that the inference is not justified except perhaps in virtue of sentiment-based custom, Kant holds that the inference is justified, but only on the basis of a *metaphysical* concept of real as opposed to logical grounds.⁷¹

In the second passage, Kant concedes that his solution to Hume's problem is only partial, given that a real ground has been characterized only negatively via its opposition to logical grounds (and the principle of contradiction). The positive line of thought that seems to be most promising to him is to think of real grounds as represented not by judgments (which would rely on the principle of contradiction), but rather by "simple, unanalysable concepts." While thinking further about how he might characterize such simple concepts, it is not implausible to conjecture that Kant would have discovered other such concepts. As we saw in the Herder transcripts, Kant also invokes real grounds to explain the inherence relationship between a substance and its own accidents. If real grounds as applied to connections *between* substances are to be represented by simple concepts, it stands to reason that real grounds as applied to connections *within* a substance would be represented by such concepts as well. After discovering several such concepts, it would be natural to wonder whether there might not be a way to present them systematically along with some guarantee of their completeness, and, more importantly, to show which objects they can and cannot refer to and how one could show this to be the case, as is evidenced by Kant's Inaugural Dissertation (2:395) and his letter to Herz in 1772 (10:130).

he develops one possible line of *response*, a line Kant clearly rejects as early as 1770–1771, as is shown by *Reflexion* 4275 (17:491–492).

⁷¹ Giorgio Tonelli, "Die Anfänge von Kants Kritik der Kausalbeziehungen und ihre Voraussetzungen im 18. Jahrhundert," *Kant-Studien* 57 (1966): 417–460) argues: "daß Kant, in dieser Phase seines Denkens, das Eigentümliche in Humes Kausalitätslehre entweder übersehen oder abgelehnt hat. Daraus folgt, daß ein Einfluß Humes auf Kants Kausalitätslehre in dieser Zeit wenigstens nicht feststellbar ist, da alles bei Kant durch den Einfluß anderer Denker leicht erklärbar ist, dagegen das Spezifische von Humes Lehre bei ihm völlig fehlt" (p. 453). It should be clear from above that Kant does not overlook Hume's position, but rather rejects it, and that one can determine with reasonable probability that Hume did exert a significant influence on Kant.

While this sketch does not even begin to reveal the wide range of complexities that lead Kant to his Critical view (insofar as it has not so much as touched on the distinction between the understanding and sensibility, synthetic and analytic judgments, Transcendental Idealism, and the various subjects treated in the Transcendental Dialectic), Kant's autobiographical remarks in the preface to the *Prolegomena* support its broad outlines.

So I first tried to see whether Hume's objection could not be put into a general form and soon found that the concept of the connection of cause and effect was by no means the only concept by which the understanding thinks the connection of things *a priori*, but that metaphysics consists altogether of such concepts. I tried to determine their number and when I had attained adequate success in this by starting from a single principle, I proceeded to the deduction of these concepts that I was now certain one could not deduce from experience, as Hume had done, but arose from the pure understanding. This deduction (which seemed impossible to my acute predecessor and which had never even occurred to anyone else, though everyone had confidently used the concepts without investigating the basis of their objective validity) was the most difficult task which ever could have been undertaken on behalf of metaphysics; and the worst thing about it was that metaphysics, such as it then existed, could not assist me in the least, because this deduction was supposed to make metaphysics possible in the first place. But as soon as I had succeeded in solving Hume's problem, not merely in a particular case, but with respect to the whole faculty of pure reason, I could proceed safely, though slowly, to determine the whole sphere of pure reason completely and from universal principles, in its limits as well as in its contents. This was required for metaphysics in order to construct its system according to a secure plan. (4:260–261)

Without considering the details of how real grounds lead Kant to reconceptualize what is necessary for what will eventually be a fully Critical metaphysics, one might still ask how real grounds can solve in a philosophically satisfying way the challenge Hume presents for Kant's pre-Critical theory of causality.

Kant "solves" Hume's problem at this point by asserting that the necessary connection that is involved in the causal relations between substances is based on his concept of a real ground. Because real grounds do not rest on the principle of contradiction, Kant can agree with Hume that denying one matter of fact (e.g., the effect) would never violate the principle of contradiction, as would be the case if an effect in one substance were to follow from its cause in another by the principle of contradiction. Moreover, since real grounds are not based on the principle of contradiction and thus are not "comprehensible" by reason, he can also still agree with

Hume that one cannot infer the existence of one thing from the existence of another on the basis of reason. Also, by introducing real grounds as distinct from logical grounds and asserting that causal relations are based on the former, Kant is committed to the necessary connections between causally interacting substances as being a form not of logical necessity, but rather of causal, or perhaps metaphysical, necessity (insofar as real grounds form part of the natures of substances). That Kant rejects thinking of the necessary connections involved in causality in terms of logical connections thus represents a significant departure from Wolff's and his own earlier conception of grounds and an important advance in his understanding of causality as well.

One might think that Kant's "solution" of Hume's problem is seriously incomplete as long as he has not specified positively what principle real grounds are based on, if not the principle of contradiction. Yet what Kant has said – even as early as 1763 – is not necessarily as incomplete as it might appear. If logical grounds depend on the principle of contradiction, it is not implausible to think of real grounds as depending on the *Nova dilucidatio's* principle of determining ground. Granted, Kant had not explicitly used the notion of real grounds in the *Nova dilucidatio*, yet there is no obvious reason to think that one could not revise that principle in light of his distinction between logical and real grounds. Thus, as a friendly amendment to that principle one might now call it the principle of determining real ground. If this suggestion is plausible, then one can say that just as the principle of contradiction (or identity) depends on the logical use of reason, the principle of determining real ground relies on its real use. This is not to say that Kant has a clear conception of what the real use of reason would amount to at this point or whether it might be better to distinguish between faculties (e.g., the understanding and reason) to be able to provide a better account of the unique functions that are carried out in each case. Views such as these are not present in Kant's writings in the early 1760s, but he does end up reflecting on them throughout the next several decades as he considers the implications of his disagreements with his rationalist background.

NEW DEVELOPMENTS IN THE INAUGURAL DISSERTATION

Seven years later (1770), in his Inaugural Dissertation, entitled *De mundi sensibilis atque intelligibilis forma et principiis* (*On the Form and Principles of the Sensible and Intelligible World*), Kant's interests ostensibly turn to the more comprehensive and systematic task of explaining the form and principles

of any world whatsoever, where he presents several new discoveries concerning the principles of the sensible world, such as that space and time, as a priori intuitions, are “not something objective and real” (2:400) but rather “subjective conditions.”⁷² However, his explanation of these principles not only will involve the theory of causality he had developed earlier in his pre-Critical period, but will also develop it in greater detail, by criticizing pre-established harmony in a novel way and presenting two new arguments for physical influx. Accordingly, Kant, unlike most other German philosophers of the period, continues to view causality as a crucial topic throughout the 1760s and thus for his entire pre-Critical period.

In discussing the generic concept of the world, which covers the more specific concepts of the intelligible and sensible worlds, Kant distinguishes between the matter and the form of the world, noting that substances constitute the matter of any world. What is of special note is Kant's specification of the form of any world:

II. Form, which consists of the *co-ordination*, not the subordination, of substances. For *co-ordinates* are related to one another as complements to a whole, while *subordinates* are related to one another as caused and cause, or, generally, as principle and that which is governed by principle. The former relationship is reciprocal and *homonymous*, so that any correlate is related to the other as both determining it and being determined by it. The latter relationship is *heteronymous*, for on the one side it is a relation of dependence only, and on the other it is a relation of causality. (2:390)

This explanation of the mutual interaction of substances fits in nicely with what Kant has already stated. For it suggests that God stands in a one-way causal bond with finite substances, which are thus subordinate to him, while finite substances relate to each other as causal co-ordinates, which is another phrase for mutual interaction (given that such co-ordinates are both “determining and being determined”).

Kant continues with a brief critique of pre-established harmony and occasionalism, which we have already seen in earlier works in a slightly less developed form.

This co-ordination is conceived of as *real* and objective, not as ideal and depending on the subject's power of choice, by means of which any multiplicity whatsoever may be fashioned into a whole by a process of adhering together at

⁷² As Giorgio Tonelli notes in “Die Anfänge von Kants Kritik der Kausalbeziehungen und ihre Voraussetzungen im 18. Jahrhundert,” pp. 451–452, Kant's views on causality in his *Träume eines Geistesehers* (1766) are unchanged from his earlier pre-Critical views.

will. For by taking several things together, you achieve without difficulty a *whole of representation*, but you do not, in virtue of that, arrive at the *representation of a whole*. Accordingly, if there happened to be certain wholes consisting of substances, and if these wholes were not bound to one another by any connection, the bringing of these wholes together, a process by means of which the mind forces the multiplicity into an ideal unity, would signify nothing more than a plurality of worlds held together in a single thought. (2:390)

Just as in the Herder transcripts, Kant claims that the form of the world is real, not ideal. However, here he describes an ideal connection as one which a subject (even, possibly, God) creates at will. As a result, he can explicitly criticize the possibility that an ideal connection could accurately describe the form of the world, since an ideal whole does not necessarily represent a whole *in itself* at all.⁷³ From Kant's perspective, there is no greater ontological connection between substances that are supposed to belong to a single world on Leibniz's account than there would be between substances that belong to different possible worlds. For one can think of causally isolated substances together regardless of whether they are claimed to belong to the same possible world or not. Accordingly, simply thinking causally isolated substances together does not make them into anything other than what they are, namely a plurality of isolated substances; such a mental act does not make them into a single world. As Kant succinctly puts it, a whole of representations (i.e., bringing several representations together in thought) does not make a representation of a whole (i.e., a representation of a single world). While this criticism represents the flip side of Crusius's criticism, Kant's novel rhetoric reveals that he has thought through the issue in an independent way.

After developing this familiar point, Kant presents a novel argument for the real connection of substances in a world:

But the connection, which constitutes the *essential* form of a world, is seen as the principle of the *possible influences* of the substances which constitute the world. For actual influences do not belong to the essence but rather the state of the world, and the transeunt forces themselves, which are the causes of the influences, suppose some principle by which it may be possible that the states of several things, the subsistence of each of which is nonetheless independent of that of the others, should be mutually related to one another as states determined by a ground. If you abandon this principle, you are debarred from positing as possible a transeunt force in the world. And, indeed, this *form*, which is *essential* to the world, is for that reason *immutable* and not subject to any change. And

⁷³ Kant first suggests this criticism in the *True Estimation*: "for otherwise there would be no discernible difference between a real and an imagined union" (1:22).

this is the case, first of all, on account of a *logical ground*. For any change presupposes the identity of the subject, whereas determinations succeed one another. Hence, the world, remaining the same throughout all its successive states, preserves the same fundamental form. For the identity of the *parts* is not sufficient for the identity of the whole; the identity of the whole requires an identity of characteristic *composition*. But, above all, the same result follows because of a *real ground*. For the nature of the world, being the first internal principle of each and every one of the variable determinations which belong to its state, cannot be opposed to itself; consequently, it is naturally, that is to say, in virtue of itself, immutable. Accordingly, in any world there is a certain constant and invariable form, which, as the perennial principle of each contingent and transitory form belonging to the state of the world, must be regarded as belonging to its nature. (2:390–391)

Kant's arguments in this passage rely on a distinction between actual and possible influences that he does not elaborate on here, but which he had explained more clearly in the *Negative Magnitudes*. There Kant says:

So far I have merely considered the grounds of real opposition, in so far as they *actually* posit in one and the same thing determinations, of which one is the opposite of the other. . . . For this reason, I shall, for the time being, call this opposition *actual opposition* (*oppositio actualis*). On the other hand, to take predicates of the following kind: although they belong to different things and although the one predicate does not immediately cancel the consequence of the other predicate, nonetheless, they may be each legitimately so called in virtue of the fact that each is so constituted that it is either capable of canceling the consequence of the other, or it is capable of canceling something which is determined like that consequence and which is equal to it. This opposition may be called *possible opposition* (*oppositio potentialis*). Both oppositions are real; that is to say, they are both different from logical opposition. (2:193)

To illustrate this distinction Kant gives two examples, one with two bodies in motion toward each other that results in a collision (a case of actual opposition), and one with two bodies in motion away from each other, so that no collision occurs, but where if a collision with a third body were to occur, it would cancel as much as the collision in the first case (a case of possible opposition).

In light of this distinction, Kant's intent in the passage cited from the *Inaugural Dissertation* can be described as follows. While in the *Nova dilucidatio* Kant had maintained that God was needed for substances to stand in interaction (given that they do not stand in interaction by means of their very existence and, taken on their own, they could not exclude the possibility of causal discord), Kant now refines his view by clarifying the nature of the connection required and by presenting new arguments for his position. More specifically, Kant suggests that in addition to the

actual interactions that occur between substances, there must also be possible interactions, which are what are responsible for the real connection between substances that unites them into a single world. The distinction is important because actual influences belong to the state of the world and thus cannot make the world possible. In particular, actual influences cannot express what is required for substances to form a single world, because, at the level of generality at which Kant is arguing, there is no guarantee that all substances would actually be interacting with each other. To take the case of two noncolliding bodies moving in opposite directions that Kant describes in the *Negative Magnitudes*, if actual interaction were used to define the world, they would not necessarily belong to the same world despite the fact that they are located in different parts of one and the same space.⁷⁴

One might argue that Kant could rule out the possibility that substances belonging to the same world would not actually be interacting, because, as Kant understands the idea at this point, matter acts immediately at a distance on all other matter by means of its attractive forces, and thus all substances in a single world would actually be interacting with each other. However, in the Inaugural Dissertation, Kant is arguing about substances *per se* and is thus not restricting his discussion to worlds that consist exclusively of material substances. Moreover, in line with his argument in the *Nova dilucidatio*, Kant suggests that actual interaction presupposes a more fundamental principle that allows substances, understood as things that subsist independently of each other, to stand in necessary connections. Here, however, Kant clearly indicates that that principle is one that pertains essentially to the world. In other words, Kant is providing a more refined description or specification of the schema of the divine understanding he had posited in the principle of coexistence. Thus, even if actual interaction could be guaranteed by attractive forces, the attractive forces themselves would presuppose a principle of possible interaction that united substances into a single world.

Kant then argues that the essential form of the world is necessarily immutable. His first line of reasoning is based on a “logical ground.” The idea is that the essential form of the world must be immutable if the world is to maintain its identity through change. For the identity of the parts of the world do not suffice in explaining the identity of the

⁷⁴ One might argue that substances must interact causally even to belong to a common space, but Kant’s views on space have changed in the Inaugural Dissertation such that this argument might not apply.

whole. Yet Kant thinks that an argument based on a real ground is more important. For he suggests that the real ground that underlies the world as a whole must be immutable insofar as it cannot be opposed to itself. In other words, since the world encompasses all substances that could be interacting with each other, there is nothing outside the world that could oppose and thus interact with it. If there were something else opposing it, that something else would be interacting with the world and would thus be part of the world rather than something external to it. Therefore, the essential nature of the world cannot be changing and therefore must be immutable.

Kant develops another provocative argument for his theory of causality later on in the Dissertation, in §22, where he explicitly addresses the traditional trichotomy of pre-established harmony, occasionalism, and physical influx. He observes that if one can validly infer from a given cause of all things to their interconnection, that is, to the form of the world, then:

the fundamental connection of substances would not be contingent but necessary, for all the substances are *sustained by a common principle*. The harmony arising from their very subsistence, a subsistence founded on their common cause, would accordingly arise in accordance with common rules. Now, I call a harmony of this kind a *generally established harmony*, whereas the harmony that only occurs in virtue of the fact that each individual state of a substance is adapted to the state of another substance would be an *individually established harmony*. And the interaction arising from the former harmony would be real and *physical*, whereas that arising from the latter would be ideal and *sympathetic*. Thus all interaction of the substances in the universe is *externally established* (by means of the common cause of them all). And it is either established generally by means of physical influx (in its more correct form)⁷⁵ or it is obtained individually for the states of each substance. But, in this latter case, interaction between substances is either founded *originarily* through the primary constitution of each substance, or it is imposed *on the occasion* of some change. Of these in turn, the former is called *pre-established harmony* and the latter *occasionalism*. . . . For myself, indeed, although the former of these alternatives has not been demonstrated, it has nonetheless been rendered fully acceptable for other reasons. (2:409)

Unfortunately, Kant does not specify in the ensuing text why his reasons merely render his view acceptable rather than amount to a demonstration of his view. Nonetheless, this passage raises an important point.

⁷⁵ See the discussion above of the different theories Kant identifies as “vulgar” versions of physical influx.

While Kant divides up the three standard causal theories in a novel way, this kind of novelty is not, as we saw above, unusual for Kant insofar as he wants the theories to be demarcated from one another so that his theory is not mistakenly identified with other versions of physical influx. However, in the present case Kant's way of characterizing the difference between the competing causal theories indicates a significant refinement in his understanding of how substances interact. Kant uses the distinction between general and singular grounding in order to distinguish physical influx from pre-established harmony and occasionalism.⁷⁶ This characterization of physical influx is surprising, since the distinction between general and singular grounding might not appear to bear on the acceptance or denial of intersubstantial causation. However, Kant now sees an important connection between these issues, one that he had only hinted at obscurely at the end of the *Nova dilucidatio* by suggesting, in a way reminiscent of one of Knutzen's objections to pre-established harmony, that God does not use the cunning devices of a craftsman as an elaborate means to bring about concord between substances. For when Kant asserts that the state of a substance is established generally according to physical influx, he intends to exclude the idea that God has considered the nature of the best possible world and created each particular state of each individual substance accordingly.

Rather, God has endowed each substance with a general nature, that is, one with causal powers that always act on other substances in certain ways under certain conditions. The link between general grounding and intersubstantial causation is established through the nature God provides each substance. This nature is general, since the nature consists in the substance's essential, that is, noncontingent and immutable, properties. Further, this nature implies intersubstantial causation, since such a nature would be superfluous unless it acted or were capable of acting (directly) on other substances.⁷⁷ This point is brought out by considering that

⁷⁶ One might suggest that Kant may have this kind of distinction in mind as early as the *Nova dilucidatio* when he states that the *general* connection of things is due to the divine schema. However, I find no evidence supporting such an attribution in addition to this piece of evidence (which can also be read simply as establishing that there is some kind of single abstract relation that combines all things that constitute a world, much like the Inaugural Dissertation's essential form of the world).

⁷⁷ Actually, this inference requires more support. For example, a Leibnizian might maintain that having a nature is not superfluous since it can act not on other substances but rather on itself. Thus, in order for this objection to apply to a Leibnizian, further argument is needed. But the situation is not hopeless, since it is less clear that pre-established harmony has any genuine use for a *general* nature.

neither occasionalism nor pre-established harmony has any genuine use for a general nature in Kant's sense, since God (rather than any general nature) determines individually the state of each substance at each time either immediately or before the creation of each substance. Accordingly, Kant presents an interesting new argument for a connection between general grounding and physical influx via general natures.⁷⁸

CONCLUSION

Against the background presented above in Chapter 1, we can now begin to appreciate Kant's contributions on the issue of causality during the first part of his career (1746–1770) and see what Kant's starting point is as he begins the complicated set of considerations that constitute the "Critical turn." Unsurprisingly, the broadest question that Kant poses, the general approach that he adopts in pursuing his own position, the basic principles that he assumes as unproblematic, and even the main outlines of his answer are all part of the German tradition in which he was educated. Thus, if the central question of the mid-1740s and for some time thereafter in Germany was whether to accept pre-established harmony or physical influx, Kant comes down decisively against pre-established harmony in favor of physical influx. Moreover, Kant arrives at his position by assuming (with Wolff and his followers) that monads are not necessarily mental, but could be physical instead. That is, Kant holds that what is substantial and real are principles of unity that are located in space and are the seats of physical forces. Kant, like many others in this period, then attempts to provide a metaphysically satisfying explanation of the physical properties of bodies, ending up with the commonsensical view that substances (i.e., physical monads) interact with each other causally so as to bring about the physical properties of bodies.

While viewing Kant's pre-Critical position at such a general level allows one to see how it arises out of and is responsive to his immediate philosophical context, it should not obscure the fact that Kant ends up holding a very distinctive and original position in this context, in fact, a position that can be fully appreciated only by way of comparison and contrast with the views of his predecessors. For example, what is distinctive about Kant's *True Estimation* and what represents its real contribution is not what he

⁷⁸ Leibniz does assert, as we saw in Chapter 1, that God does not act according to particular volitions, but Leibniz does not connect that point with the character of the natures of finite things.

had hoped for, namely a resolution of the *vis viva* controversy, but rather his idea that a force, which brings about specific properties of bodies such as motion and impenetrability, should be understood not in terms of its specific effects, such as a *vis motrix* or moving force (which Wolffians had done), but rather as essentially active. Similarly, by suggesting (against Wolff and Baumgarten) that forces can be relational and thus need not be exclusively intrinsic to a substance, Kant can allow forces a sphere of activity and thus reconcile the infinite divisibility of space and the unity of monads in the *Physical Monadology*.

As we saw, however, it is Kant's *Nova dilucidatio* that represents his most detailed and significant contribution to the issue of causality during the first part of his pre-Critical period. In the principle of succession, Kant argues that pre-established harmony is inconsistent with change and that the possibility of change therefore requires mutual interaction between substances. While the case of bodies in motion represents Kant's model of change (in a way that is consistent with his focus on physical properties in his other pre-Critical publications), his argument is based centrally on his understanding of grounds. That understanding, adopted from Wolff and Meier, is that grounds (1) are essential to the substances they constitute and (2) are immutable principles that must posit their determinations as soon as they exist (i.e., grounds and their determinations must be simultaneous). This conception of grounds is distinct from Leibniz's, since he wants to distinguish between primitive and derivative forces in such a way that change is explained by changing derivative forces within a substance rather than unchanging primitive forces. Given the Wolffian conception of grounds, Kant wants to point out that change within a causally isolated substance is impossible, that is, a substance cannot determine itself so as to change itself. Rather, change is possible only if substances both stand in mutual interaction and are changing their relations to each other. Kant thus pursues a very different strategy in explaining change than does Leibniz or Wolff. While the latter two attempt, in different ways, to explain change in terms of the resources within self-sufficient substances, Kant thinks that the contingent relations between substances must, at least in part, be the source of the mutual changes that occur within substances.

In the *Nova dilucidatio*'s principle of coexistence, Kant turns to explaining how the kind of mutual interaction between substances that is necessary for change is at all possible. Though Kant's argument is complex and subtle, his main idea is that God alone can coordinate the causal interaction between substances such that harmonious states result. This

principle represents a significant departure from Crusius's position, because Crusius thought that substances could stand in real relations by means of their mere existence alone. Kant argues, however, that substances could exist and either not be related at all or be related ("placed" or "situated") in a variety of different ways. Kant infers from this that there must be a ground that determines which relational determinations hold between substances, just as there must be grounds for intrinsic determinations within substances. But given the conception of grounds that Kant has adopted in the principle of succession, it is clear that the substances cannot themselves contain the ground of their relational determinations and that therefore the "common cause" of the existence of the substances, that is, God, must contain the ground of their relational determinations, revealing that Knutzen's "threadbare system of efficient causes," which does not "reveal the origin itself of the reciprocal connection of things," is inadequate. Thus, in the course of responding to the views of his immediate predecessors (especially Wolff, Knutzen, and Crusius), Kant develops a sophisticated (metaphysical) account of causality in the *Nova dilucidatio*, especially as it pertains to explaining motion as an instance of reciprocal changes of the relational determinations of substances.

While Kant's earliest views on causality depend in essential ways on his specifically German philosophical context, Hume became important to Kant as well after the first *Inquiry* was translated into German in 1755. But Hume was not important to Kant in the same way that he might be relevant to our philosophical interests today, namely as expressing either a radically empiricist or a skeptical position. Rather, Hume is of immediate importance to Kant *at first* due to one of his critical insights about causality, namely that one distinct, self-sufficient entity cannot be related to another such entity by means of a logically necessary connection, which is precisely what Kant's version of physical influx seemed to be committed to in the *Nova dilucidatio*. To avoid positing a logically necessary connection between independently existing entities, Kant distinguishes between logical and real grounds so that causality can be understood in terms of real rather than logical grounds. Although Hume's direct and immediate influence on Kant might thus seem to be limited to the introduction of the notion of a real ground, if one takes into account what role real grounds end up playing in his mature pre-Critical philosophy and what issues they lead to, it is clear that, historically, a much broader construal of Hume's importance emerges – one that Kant himself suggests later in his career. For Kant immediately incorporates real grounds into his philosophical theology, his account of action and reaction, and his theory of

judgment. More importantly, because real grounds are not based on the principle of contradiction, Kant sees the need to identify which principle they are based on, which eventually leads to a number of other crucial discoveries that, taken together amount to the “Critical turn.” So Kant is fully justified in emphasizing Hume’s importance to him, though it is also true that Kant would have been equally justified in stressing his immediate German predecessors, since it is against the context they provide that the initial and then the continued importance of Hume is intelligible.

In the Inaugural Dissertation, Kant continues to reflect on the issue of causality by refining various aspects of the view he had been developing over the previous twenty-five years, discovering new arguments for his views on causality, and incorporating them into a more comprehensive philosophical account. Thus, Kant wants to argue that only his theory of causality has a need for the generality and necessity of natures, since pre-established harmony and occasionalism allow that the states of the world be established “individually” rather than “generally.” He also argues that the notion of mutual interaction that he had invoked in the *Nova dilucidatio*’s principle of succession is a fundamental principle for any world, and he presents detailed considerations that pertain to the essence and logical and real grounds of the world. It is thus clear that in comparison, by 1770 Kant has developed an account of causality that is at least as detailed, sophisticated, and comprehensive as any of his predecessors’.

Armed with both a picture of the varied contributions Kant’s immediate predecessors made on the issue of causality (Chapter 1) and a detailed understanding of the complexities of Kant’s reaction to it throughout his pre-Critical period (Chapter 2), we are now in a position to turn to Kant’s Critical views on causality. For the sophisticated position Kant arrives at in his Inaugural Dissertation in 1770 represents not an endpoint for him, but rather merely a turning point along the way to his Critical position, since Kant will see the need to revise his account in fundamental ways between 1770 and 1781.

PART TWO

CAUSALITY IN THE CRITICAL PERIOD

There is no doubt that the “Critical turn” initiated by the *Critique of Pure Reason* (and completed by the *Critique of Practical Reason* and the *Critique of Judgment*) represents a truly revolutionary achievement in philosophy, one that must have far-reaching consequences for Kant’s views on causality just as they do for any of his other major philosophical doctrines. It is, to be sure, true that certain aspects of the *Critique* are anticipated in his pre-Critical period; as we saw in Chapter 2, the Inaugural Dissertation develops the Transcendental Aesthetic’s insight that space and time are merely subjective principles of the sensible world and draws a distinction between the sensible and intelligible worlds (even if Kant’s later distinction between reason and the understanding is not yet present). And Kant’s view that we can have nothing more than merely “symbolic cognition” of the intelligible world is perhaps not far from the *Critique*’s outright denial that we can have any substantive knowledge of things in themselves. However, most of the major sections of the *Critique* are not present at all in his pre-Critical period: the Preface (along with the idea of restricting knowledge in order to make room for faith), the Introduction’s powerful reflections on synthetic a priori knowledge, the Metaphysical and Transcendental Deductions, the Schematism, much of the Analytic of Principles (including the Amphiboly chapter), virtually all of the Transcendental Dialectic’s specific arguments in the Paralogisms and Antinomies, and the rich considerations developed in the Doctrine of Method.¹ And no one should be under the illusion that a mere listing

¹ This is true even if the Dissertation’s “principles of convenience” are rough anticipations of his account of the regulative principles of reason. Also, that Kant anticipates several

of sections that first appear in the *Critique* would give an adequate sense of its profound novelty.

At the same time, without denying either the importance or the novelty of the “Critical turn,” it is crucial that we arrive at a balanced understanding of its primary features. While it is easy to stress the originality of Kant’s Critical position – especially when it is contrasted with the views of his predecessors – it would be wrong to deny the important elements of continuity that hold from his pre-Critical to his Critical period. To cite just one example that is of particular importance to our present interests and whose origins we are familiar with from Chapter 2, Kant retains many of the central features of his cosmology in the Critical period. More specifically, by means of his arguments in the Analogies of Experience and elsewhere he continues to maintain that the world as we know it consists of substances that interact with each other by means of their (attractive and repulsive) forces. Kant also continues to believe that these substances are created and maintained in their reciprocal existence by God, despite the fact that he comes to deny that we could have theoretical knowledge of such a position in the Critical period, relegating it to the status of belief, based on a postulate of (practical) reason.

If the Critical Kant thus continues to accept some (even if not all) of the principles that he had initially developed during his pre-Critical years, how is this compatible with the radical break that is commonly associated with the “Critical turn”? Although a full account of the nature of the “Critical turn” involves complexities that deserve more sustained treatment than is possible in this context, the central idea here is that continuity on several points in cosmology and metaphysics is fully compatible with discontinuity with respect to epistemological and metaphilosophical issues. In short, Kant’s world stays much the same, even if his account of the way that we can come to know it does not.

In light of this point, it is striking that the sections identified above whose contents are new to the *Critique* are concerned primarily with epistemological and metaphilosophical questions. The Metaphysical Deduction attempts to establish a table of categories, while the Transcendental Deduction is mainly devoted to demonstrating their normative status with

of his “Critical” objections to the “traditional” theistic proofs in the *Only Possible Argument* need not detract from this claim significantly. See Mark Fisher’s and my “Kant on the Material Ground of Possibility: From *The Only Possible Argument* to *The Critique of Pure Reason*,” *Review of Metaphysics* 52 (1998): 369–395.

respect to knowledge. The Schematism addresses the epistemological question of how general, nonempirical concepts can be applied to particular empirical objects. The Transcendental Dialectic focuses on establishing epistemic humility by showing that we cannot have knowledge of things in themselves. The Preface, Introduction, and Doctrine of Method are somewhat different insofar as they center on the unique status of the propositions of metaphysics. The Preface's reflections on "the first thoughts of Copernicus" (Bxvi) and on how objects must conform to our intuition in knowledge (rather than vice versa) are both metaphilosophical and epistemological. The possibility of synthetic a priori knowledge, which Kant describes in the Introduction, is a metaphilosophical device he uses primarily to illustrate the radical contrast between his own philosophical method and that of his predecessors. Finally, the Doctrine of Method is particularly metaphilosophical in its investigation of the various uses of reason. We thus find in the *Critique* a decisive emphasis on epistemological and metaphilosophical issues.

However, it is crucial to see that the epistemological and metaphilosophical aspects that Kant is emphasizing throughout the *Critique* are not at all incompatible with many of his more traditional metaphysical and cosmological views.² Rather, one central issue in the following is to see precisely how Kant thinks that his "traditional" metaphysical views can be developed within and supported by a new epistemology *after* he has jettisoned the dogmatic method of rationalism. Appreciating this point is essential to attaining a balanced view of the nature of the "Critical turn," that is, an interpretation that can account for both the continuities and the breaks that occur in Kant's position from his pre-Critical to his Critical period.

These interpretive remarks about the "Critical turn" are, at this point, speculative in certain respects, given that Kant's Critical doctrines have not been presented, much less discussed in detail. Moreover, a systematic and thorough investigation of the nature of the "Critical turn" must remain an unfulfilled desideratum even after the completion of this study. For what is required for current purposes is only that Kant's Critical views on causality and the relation they bear to their pre-Critical counterparts be seen in a clearer light. Thus, our current investigations do not replace

² For an extensive discussion of how Kant's Paralogisms retain an important metaphysical core, despite their criticisms of positions embraced on the basis of "rational exuberance," see Karl Ameriks's *Kant's Theory of Mind* (New York: Oxford University Press, 1982).

so much as anticipate what seems long overdue, namely a careful and extensive examination of the nature of the “Critical turn.”³

However, even if we possess only a rough and somewhat speculative understanding of the “Critical turn,” we are able to turn to our primary task, namely an investigation of Kant’s views on causality in his Critical period. Chapter 3 (“Kant’s Second and Third Analogies of Experience”) begins by reconstructing Kant’s main arguments for his most fundamental claims involving causality. Chapter 4 (“Kant’s Model of Causality”) then describes in detail the intricate model of causality explicitly or implicitly embedded in these claims. In this way we can illustrate the central features of Kant’s Critical account of causality.

³ Wolfgang Carl has made an impressive start on such an issue by investigating Kant’s complicated reflections throughout the 1770s in *Der schweigende Kant: Die Entwürfe zu einer Deduktion der Kategorien vor 1781* (Göttingen: Vandenhoeck & Ruprecht, 1989).

Kant's Second and Third Analogies of Experience

INTRODUCTION

While Kant's main argument concerning causality is typically located in the Second Analogy of Experience, it is clear that any interpretation that could be adequate to Kant's intentions must be able to render intelligible the text and argument of the Third Analogy of Experience as well, since it attempts to establish the necessity of mutual interaction, a notion that, as we have seen in Chapter 2, receives sustained attention throughout his pre-Critical period. Accordingly, the central task of this chapter is to reconstruct Kant's arguments in both the Second and Third Analogies in detail. After considering various aspects of the general context of the Second and Third Analogies in a first section, we develop detailed reconstructions of the arguments of the Second and the Third Analogies in the second and third sections of this chapter.

THE CONTEXT OF THE SECOND AND THIRD ANALOGIES

To reconstruct the arguments of the Second and Third Analogies properly, it is crucial to understand exactly what they are and are not supposed to accomplish. Thus, we must begin by describing both the distinctive context of the Analogies of Experience within the *Critique of Pure Reason* and the argumentative framework of the Analogies in general. More specifically, it is important that we understand (1) how the Analogies relate to the Transcendental Deduction, (2) that they are concerned with addressing what can be called the problem of time-determination, (3) that they are concerned with the possibility of experience in the form

of the unity of time, and (4) that they should be understood as making neither purely metaphysical nor exclusively epistemological claims, but rather claims that combine metaphysical and epistemological elements.

The Analogies within the “Analytic of Principles”

The Analogies of Experience are sections of the *Critique* that form part of the Transcendental Logic’s Analytic of Principles, whose main task is

to exhibit in systematic combination the judgments that the understanding actually brings about *a priori*. . . for which our table of the categories must doubtless give us natural and secure guidance. For it is precisely these whose relation to possible experience must constitute all pure cognition of the understanding *a priori*, and whose relation to sensibility in general will, on that very account, display all transcendental principles of the use of the understanding completely and in a system. (A148/B188–189)

In other words, the Analytic of Principles is supposed to establish in a systematic and exhaustive way all of the primitive *a priori* principles of which our understanding is capable when the categories are applied to sensible intuition. The systematicity and exhaustiveness of these Principles, of which the Analogies of Experience are only three, is to be guaranteed through the systematicity and exhaustiveness of the table of categories established earlier in the Analytic of Concepts, in the so-called Metaphysical Deduction. Kant explicitly contrasts the task of the Analytic of Principles with that of the Analytic of Concepts with the remark: “In the previous chapter we have considered the transcendental power of judgment only in accordance with the general conditions under which alone it is authorized to use the pure concepts of the understanding for synthetic judgments” (*ibid.*).

While these statements might naturally give rise to many questions about the status of Kant’s project in this part of the *Critique*, one immediate query concerns the relationship between the various Principles established in the Analytic of Principles and the Analytic of Concepts’ Transcendental Deduction, which similarly provides an “explanation of the way in which concepts [i.e., the categories] can relate to objects *a priori*” (A85/B117). In particular, one might wonder whether or not the arguments for the various Principles presuppose the conclusions of the Transcendental Deduction and if so, how.¹ This question becomes all

¹ Jonathan Bennett, *Kant’s Analytic* (New York: Cambridge University Press, 1966), offers the unlikely suggestion that Kant supplies the arguments of the Analogies because his

the more pressing as soon as one notices that Kant nowhere indicates explicitly that he is drawing on the Transcendental Deduction per se in support of the assumptions he makes in his arguments for the Principles.² While Kant sometimes contrasts the generality of the Transcendental Deduction with the specificity of the Principles (a contrast intimated in the quotation above), it is clear from the arguments Kant actually provides that he does not consider his task, as one might surmise, to be one merely of substituting each of the various categories into the general argument of the Transcendental Deduction in order to get a more specific result. Rather, the danger – if it is one – is the reverse, namely that it appears as if the most important arguments for the various Principles proceed completely independently of the Transcendental Deduction.

Without going into the details of the Transcendental Deduction, which are a matter of considerable dispute in their own right, it is clear that whatever its assumptions and prospects for success, its self-described goal is to establish that the categories are *normatively justified* in synthetic judgments if and only if they are applied to objects given in sensible intuition. In describing the task of a transcendental (rather than empirical) deduction in the introduction to the Transcendental Deduction, Kant asserts that what is at issue is establishing the *quid juris* of the categories, not their *quid facti* (A84–85/B116–117), which the Metaphysical Deduction had already accomplished. If the Principles attempt to establish that particular categories are *necessary* for certain sorts of experience, then they must still presuppose the results of the Transcendental Deduction to be justified in asserting that the categories are *normatively justified* in those contexts. For even if the categories are necessary in certain contexts, that does not immediately establish that they can accurately represent (or structure) the world. That is, it is in principle possible that the categories, as discursive

previous arguments (in the Transcendental and Metaphysical Deductions) “have not the faintest appearance of following” (p. 93).

² Paul Guyer raises this concern in *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987): “Kant now writes as if the argument for the principles must go back to the very foundations of the deduction itself – as if it must restart the argument of the deduction – rather than just apply the conclusions already reached in the deduction and the schematism” (p. 178). But Guyer also notes: “When he comes to the central paragraph of this introduction to the principles, Kant does not write as though there had been no transcendental deduction at all. But he certainly does write as though the deduction had established at most the very general point that experience or empirical knowledge requires *the concept of an object*, and as if the schematism had added only the equally general point that this must be the concept of an object capable of being given by means of the kind of sensibility we possess” (p. 209).

representations of objects, might be unavoidable distortions bound up with our perception of the world rather than accurate reflections of it, and it is a matter of argument to establish that the latter case holds rather than the former. Thus, if the Principles contain arguments that specific categories must be applied in certain contexts, then they implicitly assume from the Transcendental Deduction that such applications are also normatively justified (to the extent that they satisfy the conditions established in the Transcendental Deduction for such applications, namely that they be applied to objects given in intuition).

The Problem of Time-Determination

It is also important to understand what is entailed by the fact that the Second and Third Analogies are Analogies of Experience and thus centrally concerned with the issue of time-determination.³ The structure of the Analogies of Experience is set up such that they are intended to establish that certain substantive principles involving the relational categories (of subsistence-inherence, causality-dependence, and community) are necessary for determining the temporal relations of objects. Thus, the First Analogy argues that a substance-accident ontology is necessary for the determination of persistence or duration, the Second Analogy tries to show that a causal rule is necessary for determining the succession of an object's states, while the Third Analogy attempts to establish that mutual interaction is required to determine the coexistence of substances.

To understand why substantive principles involving causally interacting substances are required for the determination of duration and the temporal relations of such substances, it is crucial to see why Kant thinks that time-determination is problematic in the first place. After all, one might naturally think that time-determination presents no difficulty at all: Look at a table, then look at a chair standing next to it and you simply know without further ado that they coexist.⁴ However, this intuitive picture of time-determination is naive, according to Kant, in several respects.⁵ First, in the proof of the general Principle of the Analogies Kant

³ Kant discusses the term "analogy" and the most significant differences between mathematical and philosophical analogies at A178–180/B220–222.

⁴ In fact, the issues are no different if we consider two objects that are in view at the same time.

⁵ Guyer adds the further aspect to the problem of time-determination that one cannot take for granted "that the order of *perceptions* is known independently of anything else" (Guyer, *Kant and the Claims of Knowledge*, p. 256). In other words, Guyer thinks that even "the subjective order of time" is not immediately given.

emphasizes that we cannot perceive either “time itself” (B219) or, as he calls it elsewhere, “absolute time.”⁶ In other words, we cannot determine the times at which an object exists just by looking either at time itself or the object; this table in front of me does not lie within a space-time entity with clearly marked coordinates (as a very crude Newtonian might picture it), nor does it have listed on it anywhere its temporal properties, the times at which it exists in this world (a picture perhaps suggested by a vulgar version of Leibniz’s theory of complete concepts). Thus, one cannot determine the place of objects in time (i.e., the times “at which” objects exist) through either of these methods. But since the places of objects in time cannot be determined through either of these methods, one cannot derive their temporal relations (i.e., whether as successive or coexistent) through these methods, either. Rather, to determine the temporal order of (the states of) objects, some kind of indirect, relational procedure must be performed. The Analogies of Experience state the rules for this procedure.⁷

Second, there is an important distinction between subjective and objective time. One of Kant’s most famous examples, that of the ship and the house in the Second Analogy (A192/B237), illustrates this distinction nicely. Although our apprehension is always successive, the various states of the object apprehended may not be, since, for example, the parts of a house, although apprehended successively, are not successive but rather coexist. As a result, one must distinguish between the temporal relations in which the states of objects stand to each other and the temporal relations of the representations by means of which the states of objects and their temporal relations are apprehended.

But given these two points, determining the temporal relations of objects now appears to pose a serious problem, not at all trivial as it might seem at first glance. How can one determine the temporal relations of objects in the world, if they cannot be perceived directly but rather can be presented only through the subjective temporal order of apprehension?

⁶ Crusius may be hinting at this problem when he remarks that “it is obvious that according to our linguistic practices we do not call the sequence of things one after another time itself. For by time one means something that relates uniformly to every thing that actually does or might occur in it. And time, which we represent to ourselves according to a useable sensible measure, since we cannot distinguish its moments intrinsically by anything, remains the same, whatever may happen. We thus understand by time not the actual succession of things, but rather merely something in which they can and do follow each other” (*Entwurf der nothwendigen Vernunft-Wahrheiten* (Leipzig, 1745), §54, pp. 90–91).

⁷ It is not necessary, for current purposes, to defend the claim that Kant’s other Principles state rules for this procedure as well. See Guyer, *Kant and the Claims of Knowledge*, pp. 184f., 194, and 202, for the difficulties such a defense would have to overcome.

The structure of the Analogies of Experience implies that the relational categories – which involve substantiality and causality – must be used to determine the objective temporal relations of the world.⁸ While claiming that relational categories can solve the problem of time-determination may sound extravagant and appear to be driven solely by architectonic considerations, reflection on commonsense examples (such as what is actually involved when we not only look at a clock, but also interpret it as indicating objective time) makes it increasingly plausible to think that our everyday practices depend on an object's properties and causal mechanisms, which is an informal take on what Kant is arguing.⁹

However, the problem of time-determination raises more issues than this portrayal might suggest. For the considerations just presented might seem to pertain to purely epistemological issues that would leave the ontological question of the temporal determinations of objects untouched. In other words, they might leave one with the impression that it is a question of how we could come to *perceive* the temporal relations that already obtain between objects rather than how such relations could exist in the first place. Yet Kant is clearly concerned with both issues. Kant distinguishes between the Axioms of Intuition and the Anticipations of Perception as “mathematical” Principles and the Analogies and Postulates of Empirical Thought as “dynamical” Principles, noting that the former are constitutive of the objects that are given to us in intuition while the latter are merely “regulative.” In explaining the regulative status of the Analogies, Kant makes it clear that the regulative principles concern the “relation of existence” between objects (A179/B222) rather than the existence of objects *per se*, since, as he puts it, “if a perception is given to us in a temporal relation to others . . . it cannot be said *a priori* **which** and **how great** this other perception is, but only how it is necessarily combined with the first, *as regards its existence*, in this *modus* of time” (A179/B222, emphasis added). That is, the dynamical Principles state not which objects must exist, but rather which necessary connections they must enter into in order to constitute their temporal relations in the first place.

The idea that temporal relations must not only be discovered, but rather constituted, is also a consequence of Kant's arguments for Transcendental Idealism in the Transcendental Aesthetic. There Kant argues

⁸ Properly speaking, of course, it is not the relational categories *per se*, but rather their schemata that must be employed in the Analogies, since the categories proper (i.e., unschematized) have no explicitly temporal meaning.

⁹ The point is even more obvious if one takes “experience” in the restrictive sense of scientific experience.

that space and time are merely subjective forms of intuition constituted by us, not properties of or relations between things in themselves that we might simply discover. But if the Aesthetic establishes that time and thus all temporal determinations are not independently given, then the Analogies cannot assume that objects exist with their temporal determinations, and assert that the relational categories merely aid in the process of discovery, especially given the fact that even our passive faculty of intuition provides only subjective and not objective temporal relations. Accordingly, it is clear that the Analogies argue for substantive principles that concern the very constitution of the temporal relations of objects.¹⁰

The Unity of Time

Let us consider another issue that pertains to the place of the Analogies of Experience within the architectonic structure of the *Critique*. As we saw above, the systematicity and completeness of the Principles of Pure Understanding is supposed to be guaranteed by that of the table of categories and it thus stands to reason that the three Analogies of Experience parallel the three relational categories. Despite the fact that the categories of

¹⁰ This point does not hinge on whether one interprets Transcendental Idealism as an ontological or an epistemological doctrine. To cite just one example each of the ontological and epistemological interpretations of Transcendental Idealism, see Karl Ameriks, "Kantian Idealism Today," *History of Philosophy Quarterly* 9 (1992): 329–342, and Henry Allison, *Kant's Transcendental Idealism: An Interpretation and Defense* (New Haven: Yale University Press, 1982). In the case of the ontological interpretation, I take it to be obvious that the Analogies concern the ontological dimension of time-determination (since what is at issue, on this interpretation, is how phenomenal substances can actually have, rather than simply be considered to have, determinate temporal states). In the case of the epistemological interpretation, it may not be as immediately clear. If the epistemological interpretation of Transcendental Idealism dictates that we take up different standpoints on objects by abstracting or not abstracting from space and time, one might think that there is some sense in which the objects exist with all their determinations and the only question that arises is what we can do with the resources inherent in each standpoint, and from the standpoint of knowledge it would merely be a matter of coming to know the temporal properties that objects already have. However, this interpretation disregards an essential feature of the epistemological interpretation of Transcendental Idealism by presupposing that objects would have all their determinations independently of any standpoint. A more appropriate epistemological interpretation would assert that objects and their determinations must be defined with respect to, and are constituted by, standpoints. To speak of objects and their determinations independently of standpoints (or epistemic conditions, to use Allison's preferred phrase) would in reality be to presuppose a divine standpoint, which is both unavailable to us and atemporal, which precludes the possibility that the Analogies could be concerned with it insofar as they are essentially about time-determination from an essentially human perspective.

quantity and quality do not map neatly onto individual Principles in the same way, this correspondence clearly holds for the Analogies of Experience.¹¹ However, in his introduction to the Analogies of Experience Kant also refers to three *modi* of time: persistence, succession, and simultaneity (A177/B219). While Kant may (or may not) have thought that the three *modi* of time are systematically derivative of the relational categories, the crucial point to see is that the three *modi* of time make possible what Kant sometimes calls “the unity of time.” By the “unity of time” Kant means to refer to the fact that there can be only one time, that all moments of time must be successive parts of it, and that it is in this single time that all temporal objects that we could know along with their states or determinations must be located. Time is thus an all-encompassing unity in this sense.

Although Kant does not explicitly develop this line of thought, one can nonetheless understand how the three modes of time make the unity of time possible as follows.¹² Persistence is required for the unity of time insofar as it expresses the idea that time stays the same while its moments and the states of affairs that obtain at each moment do not. That is, time does not go into and pop out of existence as its moments (or contents) do. As a result, for there to be a single time, something must persist so as to relate the various moments as moments of one and the same time.¹³ In virtue of its persistence, time can have an enduring magnitude. Kant refers to the First Analogy as expressing the relation of the existence of appearances “to time itself, as a magnitude” (A215/B262). Succession and simultaneity, by contrast, are the two basic temporal relations that the states of objects must bear to each other to be temporally related, and they jointly make such a persisting time possible. That is, for objects (or their states) to be located in one and the same time, the state of one

¹¹ Guyer comments on this point as follows: “There is no suggestion that we know that we can make just these three forms of time-determination *because* we *independently* know that there are three categories of relation, or that we must be able to make these three kinds of time-determination *because* only thus will we be able to use or lend objective reality to such categories. . . . Kant goes directly from the modes of time which have to be determined because of the structure of time itself and the possible relations among objects in it to the existence of principles of the understanding, without either an advance road map or detour” (*Kant and the Claims of Knowledge*, p. 180).

¹² A similar line of argument is developed in a more formal way by R. I. G. Hughes, “Kant’s Analogies and the Structure of Objective Time,” *Pacific Philosophical Quarterly* 71 (1990): 141–163.

¹³ It is this idea that might suggest that time should be understood substantively, as Newton did.

object must be either simultaneous with or before or after each state of every other object.

Notice that either simultaneity or succession taken in isolation will not guarantee the unity of time. If all states were related only by means of simultaneity, then there could be as many worlds as there are moments, since there would be no temporal link between any one set of simultaneous objects with any other. Similarly, if states were related only by means of succession, there could be several independent chains of successive states that still lack temporal relations to each other. Thus, both simultaneity and succession are required to guarantee that a plurality of states can be related in one and the same persisting time, and in this way the three modes of time make the unity of time possible.

But why should Kant be concerned with making the unity of time possible in the Analogies of Experience if the point of the Analytic of Principles is to lay out in a systematic and exhaustive way the fundamental synthetic a priori principles of our understanding? In addition to his architectonic interests (which stem from showing how each of the categories can be applied to sensible intuition in a specific context), one of Kant's most fundamental endeavors in the *Critique* is to explain how experience is possible, but when Kant says that he wants to explain the possibility of experience, it is really the possibility of temporal experience that he has in mind.¹⁴ Insofar as what he means by temporal experience is the experience of objects in one and the same time, Kant's general task transforms into the more specific task of explaining how the unity of time is possible. Kant clearly draws this connection in the following remark: "There is only **one** experience, . . . just as there is only one space and time. . . . If one speaks of different experiences, they are only so many perceptions insofar as they belong to one and the same universal experience" (A110).¹⁵

Yet the unity of time is also connected in an important way to the unity of the world. If Kant is going to appropriate any of the traditional

¹⁴ It is tempting to view the *Metaphysical Foundations of Natural Science* as similarly attempting to explain the possibility of spatial experience or, to be more precise, the experience of bodies in motion. For a sense of how such an explanation might go, see my "The Argumentative Structure of Kant's *Metaphysical Foundations of Natural Science*," *Journal of the History of Philosophy* 36 (1998): 567–593.

¹⁵ This reading thus contrasts with Pierre Keller's interesting interpretation in *Kant and the Demands of Self-Consciousness* (New York: Cambridge University Press, 1998), which invokes a universal point of view or perspective of which all individual perspectives are limitations.

cosmological doctrines of the pre-Critical period, he is going to have to provide some motivation for thinking that its central concept, that of the world, is not “dogmatic” and thus can have a place in his Critical project. For if cosmology were to presuppose a concept of the world and then simply derive analytic truths from it, the question of why one should accept that particular concept of the world (or, for that matter, any concept of the world at all) becomes quite pressing. It is clear that the unity of experience and thus the unity of time provide Kant with precisely the motivation he needs for his concept of the world.

The idea is as follows. If the temporal relations of succession and simultaneity are necessary for the unity of time and thus the unity of experience and if the Second and Third Analogies argue that causality and mutual interaction are necessary for succession and simultaneity, then it follows that causality and mutual interaction are necessary for the unities of time and experience. But if mutual interaction between substances constitutes the form of the world (as Kant maintained in the Inaugural Dissertation), then the world is as necessary for the unities of time and experience as mutual interaction and simultaneity are. The unity of experience, the unity of time, and the unity of the world thus all go hand in hand.¹⁶ Given this connection, it should come as no surprise that (1) after presenting all the Principles of Pure Understanding (A227–230/B280–282), Kant accepts several traditional cosmological principles and (2) in the introduction to the Antinomy of Pure Reason (A418–420/B446–448) he employs the concepts “world” and “nature” in describing the sum of causally interacting appearances.¹⁷

Kant expands on this point in his general note to the Analogies of Experience as follows:

By nature . . . we understand the combination of appearances as regards their existence, in accordance with necessary rules, i.e., in accordance with laws. There are therefore certain laws, and indeed *a priori*, which first make nature possible. . . . Our analogies therefore really exhibit the unity of nature in the combination of all appearances under certain exponents, which express nothing other than the relation of time . . . to the unity of apperception. . . . Thus together they say: All appearances lie in one nature, and must lie therein, since without this *a priori* unity no unity of experience, thus also no determination of the objects in it, would be possible. (A216/B263)

¹⁶ More precisely, the unity of the world is a necessary condition of the unity of time.

¹⁷ For a discussion of the cosmological principles of no chance, no fate, no gap, and no leap, see my “Kant on Rational Cosmology,” in *Kant and the Sciences*, ed. E. Watkins (New York: Oxford University Press, 2001), pp. 70–89.

Kant asserts a strong connection here between the unity of nature, the unity of experience, the unity of time, and the unity of apperception, which simply underscores the idea that Kant would continue to accept the importance of cosmology in the *Critique*, even if his primary focus there is on other, newly discovered kinds of unity. More specifically, insofar as Kant attempts to establish different aspects of cosmology as necessary conditions of these “new” unities, an important dimension to Kant’s “Critical turn” can be understood in terms of an argument from relatively “traditional” and uncontroversial cosmological concepts and principles to novel epistemological aims in order to carry out a project that combines metaphysical and epistemological aspects in a revolutionary way.¹⁸

But it is also important to see a further dimension of the relation between the Analogies of Experience and the basic framework of the *Critique*. As we have seen above, at an extremely high level of generality, the *Critique* attempts to provide an explanation of how experience is possible. But what kind of experience is at issue? Perceptual commonsensical experience or scientific experience? While both of these possibilities have found supporters who garner plausible textual and philosophical evidence in their favor, this reading of the Analogies of Experience suggests that neither of these options properly captures the exact notion of experience that Kant has in mind. For, as we have seen, what is essential to experience in the context of the *Critique* is that the objects of experience belong to one and the same time. But because such a requirement holds of both perceptual experience as expressed in common sense and scientific experience as represented in, for example, Newtonian science, Kant would seem to be interested in a much more general notion of experience than either of these possibilities taken separately.¹⁹

¹⁸ While one might be concerned that the concepts of “nature” and “world” are being conflated here, it is clear that Kant thinks that they can, in most contexts, be taken to be identical, given that the differences between the two are quite minor. As Kant explains in the beginning of the Antinomy of Pure Reason: “We have two expressions, **world** and **nature**, which are sometimes run together. The first signifies the mathematical whole of all appearances and the totality of their synthesis in the great as well as in the small. . . . But the very same world is called nature insofar as it is considered as a dynamical whole and one does not look at the aggregation in space or time so as to bring about a quantity, but looks instead at the unity in the **existence** of appearances” (A418–419/B446–447).

¹⁹ This position is thus different in an important respect from the view of Karl Ameriks, “Kant on Science and Common Knowledge,” in *Kant and the Sciences*, ed. E. Watkins (New York: Oxford University Press, 2001), pp. 31–52. For Ameriks thinks that Kant views the role of philosophy as mediating between the apparently conflicting claims of the “manifest image” and the “scientific image,” whereas I am suggesting that philosophy’s role is simply to articulate certain common necessary conditions of both of these “images,”

The Status of the Claims: Analytic or Epistemological?

These issues are closely related to yet one more question about the exact meaning of the claims of the Analogies of Experience and the precise status of their arguments. As we have just seen, they are involved in making the unity of time and thus experience possible, but that does not specify precisely enough what kind of claim they are making and by means of what kind of argument. For instance, are the Analogies asserting that the relational categories are necessary for us merely to make *judgments* about the temporal relations of objects? In other words, does the very idea of objective succession contain the idea of causality and is the argument of the Second Analogy therefore based exclusively on conceptual analysis of the *meaning* of the idea of objective succession? Strawson would seem to be a proponent of this view by assuming that Kant is a descriptive metaphysician and by interpreting the Second Analogy as asserting that causality is involved in what we understand an object (and perhaps the succession of its states) to be.²⁰ As Strawson explains: “The conception of an objective world is bound up with the conception of alternative possible experiential routes through it, with the distinction between subjective experience and the world of which it *is* experience, and with the very possibility of empirical self-consciousness.”²¹ In light of Strawson’s analytic-descriptive mode of argument and stress on the very meaning of our idea or conception of objective succession, I call this the *analytical model* of the Analogies.

with no conflict required. Thus, the point of philosophy (as represented in the *Critique*) is not primarily to *mediate* between the *conflicting* claims of common sense and science, but rather to specify conditions that either one would have to meet for its claims to be even possible in the first place.

²⁰ See Peter Strawson, *Individuals: An Essay in Descriptive Metaphysics* (New York: Methuen, 1959), p. 9, and *The Bounds of Sense* (New York: Methuen, 1966), pp. 140–146. However, the view would seem to go back at least to A. C. Ewing, *Kant’s Treatment of Causality* (Hamden, Conn.: Archon, 1924), p. 83, though Ewing emphasizes that the analysis is not of mere concepts, but of *experienced* concepts, without fully clarifying what difference that makes. Henry Allison, too, seems to accept the analytical model when, for example, he distances his own position from Guyer’s and Friedman’s interpretations, which he characterizes as epistemological (“Causality and Causal Laws in Kant: A Critique of Michael Friedman,” in *Kant and Contemporary Epistemology*, ed. P. Parrini (Dordrecht: Kluwer Academic, 1994), pp. 291–307, esp. p. 301). It is, I think, doubtful that Friedman actually ascribes to an epistemological interpretation insofar as he stresses the idea that Principles, like the Analogies, are necessary for the very meaning of terms such as “true or absolute motion.” See Michael Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992), e.g., pp. 46–47.

²¹ Strawson, *The Bounds of Sense*, p. 121.

In contrast to the analytical model, one might think that Kant is arguing that the relational categories are necessary for us to have *knowledge* of objective temporal relations. There are several versions of such an *epistemological model* of the Analogies that stem from different answers to the following question: If the relational categories are necessary for knowledge, what aspect of our knowledge generates the need for the categories? One response, call it the *psychological* or *phenomenological* version of the epistemological model, would assert that our knowledge is always based on perception and that perception is possible only if the categories are employed in our psychological or phenomenological apprehension of the world in reflection.²² Yet another response, the *justification* version of the epistemological model, would claim that in order for experience to be knowledge it must be justified and that Kant's arguments attempt to show that any justification must involve the relational categories.²³ Paul Guyer, a leading proponent of this version, explains the status of the claims of the Analogies as follows: "Kant is dealing strictly with principles that would have to be appealed to in the justification of empirical claims to knowledge," and he repeatedly contrasts such an epistemological framework with questions about any purported psychological processes for the generation of representations or beliefs.²⁴

Unfortunately, each of these interpretations encounters significant difficulties. Against the analytical model, it is clear that Kant's arguments must extend beyond mere conceptual analysis for two reasons. First, if Kant's argument were based solely on conceptual analysis of, for example, "knowledge (or experience) of objective succession," then it is plausible that the first step in such an analysis would result in the concepts of knowledge (or experience), object, and succession. But it is clear that (1) the very concept of knowledge (or experience) is unlikely to contain the concept of causation, since a priori knowledge of mathematics would seem to contain no causal elements, (2) the concept of an object would not seem to contain the concept of causality insofar as we can conceive of causally inert objects, and (3) succession is a temporal rather than a causal concept, even if they are closely related. As a result, mere analysis of these three concepts is unlikely to deliver the concept of causation (whether that of "simple" causation or mutual interaction). Second, if

²² See Beatrice Longuenesse, *Kant and the Capacity to Judge* (Princeton: Princeton University Press, 1998), and Pierre Keller, *Kant and the Demands of Self-Consciousness*.

²³ See Paul Guyer, *Kant and the Claims of Knowledge* (e.g., p. 27).

²⁴ *Ibid.*, pp. 258–259.

Kant's argument were based merely on conceptual analysis, then, unless one were to accept a strong version of verificationism, one would, ultimately, still be uncertain as to whether this analysis had any application to the world (a problem that has been repeatedly raised against Strawson's own attempts at deploying transcendental arguments).²⁵ Accordingly, the analytical model, if taken in isolation, seems problematic.

Yet the purely epistemological models fare no better. The psychological or phenomenological version of this model, which does find clear textual support in, for example, the first edition *Transcendental Deduction*, focuses too heavily on the "subjective" side of Kant's argument, and thus leaves no distinct task for Kant's "objective" argument. That is, it is true that Kant sees the need to explain how experience is possible in terms of how our distinctive faculties of understanding, sensibility, and imagination are able to represent substances, causes, and so on, by means of concepts, intuitions, and syntheses. While this subjective dimension is crucial to Kant's overall project, it is hard to see that it could exhaust Kant's project.²⁶ Though one may not need to go so far as to claim that "the argument of the second analogy is entirely free from any reference to real or imagined psychological processes for the generation of particular representations or beliefs," the underlying point, namely that it does not consist *exclusively* in such processes, must be granted.²⁷ In particular, Kant's interest in Newtonian science in the *Metaphysical Foundations of Natural Science* strongly suggests that Kant is thinking that the notion of experience – the possibility of which he is trying to explain – must at least include, if not be exhausted by, scientific experience, which need not be purely psychological, phenomenological, or perceptual in any narrow sense.

However, the justification version of the epistemological model faces questions of its own. In claiming that knowledge of objective temporal relations can be justified only on the basis of causal laws, what is really being claimed is that such knowledge can be justified only if we *know* what causal laws hold, and the question then arises as to what justifies

²⁵ Strawson's *Individuals* sparked an intense debate about the nature and viability of transcendental arguments. For one prominent line of criticism of Strawson's version of transcendental argument, see Barry Stroud, "Transcendental Arguments," *Journal of Philosophy* 65 (1968): 241–256.

²⁶ For a detailed account that stresses the importance of Kant's transcendental psychology, see Patricia Kitcher, *Kant's Transcendental Psychology* (New York: Oxford University Press, 1990).

²⁷ Guyer, *Kant and the Claims to Knowledge*, p. 258.

our knowledge of causal laws.²⁸ The apparent answer must be that it is justified by our knowledge of particular objective temporal relations. At this point, worries about circularity immediately arise, since it appears that we are appealing to knowledge of objective temporal relations in order to justify knowledge of objective temporal relations. One can avoid the circularity by stipulating that one may not appeal to the very same sequence of representations as being both derived from a particular causal law and as evidence for that law, but that makes it extraordinarily difficult, though perhaps not impossible, to see what could be the first independent element in the chain of evidence and derivation.

If each of these models faces significant questions, what kind of claim might Kant be advancing instead in the Analogies of Experience? A decisive clue lies in a footnote buried at the end of his conclusion to all three Analogies. In a paragraph devoted to “the method of proof which we have employed in the case of these transcendental laws of nature” (A216/B263), Kant argues, in a familiar vein, that the arguments of the Analogies can be based not on “mere concepts, . . . no matter how much one analyzes them” (A217/B264), but rather on the conditions of the possibility of experience. After noting that previous philosophers had erred in trying to prove the principle of sufficient reason analytically, he claims that they were also mistaken in not even thinking “of the other two analogies, though one always tacitly employed them” (A217/B265). Kant explains this mistake further in a footnote:

The unity of the world-whole, in which all appearances are to be connected, is obviously a mere conclusion from the tacitly assumed principle of the community of all substances that are simultaneous: for, were they isolated, as parts they would not constitute a whole, and were their connection . . . not already necessary on account of simultaneity, then one could not infer from the latter, as a merely ideal relation, to the former, as a real one. Nevertheless we have shown, in its proper place, that community is really the ground of the possibility of an empirical cognition of coexistence, and that one therefore really only infers from the latter back to the former as its condition. (A218/B265)

Note first that Kant describes the inference from an objective temporal relation (simultaneity) to a causal relation (community) as being an inference from something *ideal* to something *real*. If an ideal relation is to

²⁸ For the sake of argument, one can grant that introducing causal laws is the most natural way to make sense of our knowledge of particular instances of temporal succession. What Guyer's interpretation needs is that such causal laws are not merely natural, but also necessary.

be understood as one that exists merely in thought and a real relation as one that exists between objects (as Baumgarten, Crusius, and the pre-Critical Kant all held), then Kant is characterizing his argument in the Analogies as inferring from what obtains in thought to what obtains in things as a “condition” and “ground” of this cognition. That is, the Analogies are arguing that something ontological or metaphysical is required as a condition to ground something epistemological, and that the one can thus be said to make the other possible.

Understanding the claims of the Analogies in this way fits in neatly with the problem of time-determination. As we saw above, the problem of time-determination is not merely an epistemological issue concerning the discovery or justification of the temporal relations between objects, but rather an ontological matter concerning the constitution of the relations between objects. While the term “determination” might have epistemological connotations for us, Kant clearly thinks that it can also have ontological import. As we saw in Chapter 2, the pre-Critical Kant, following Baumgarten, defines the notion of “determination” by means of the notion of a ground as follows: A ground “converts things which are indeterminate into things which are determinate” (1:392). That is, grounds posit determinations. Although the pre-Critical Kant distinguished between epistemological and metaphysical grounds (in the form of antecedently versus consequentially determining grounds) so that the notion of a ground could be taken either epistemologically or metaphysically, the *Nova dilucidatio*’s principles of succession and coexistence make it obvious that in the context of the temporal relations of succession and coexistence he is concerned with metaphysical grounds. The fact that Kant continues to refer to the notion of a causal relation (community) as a “real” relation that “grounds” “ideal” temporal determinations (our knowledge of coexistence) reveals that he has not abandoned the metaphysical dimension of that notion, even if it is transformed by being placed in a context that has an epistemological dimension.

Accordingly, rather than thinking of the claims of the Analogies as either purely analytical/conceptual or purely epistemological, we can argue that the Analogies combine epistemological and metaphysical aspects. More specifically, the idea is that Kant is claiming that *knowledge* of objective temporal relations requires substantive *ontological* principles. It is important to note what is meant by “ontological” or “metaphysical” in this context. The “ontological” principles asserted in the Analogies are not about things in themselves, but rather hold for phenomenal substances and their relations and are thus made from within an

epistemological framework that is defined in terms of possible experience. To illustrate this point, let us continue with the example of the Third Analogy. Kant wants to claim that knowledge of objective coexistence requires not only the coexistence of phenomenal substances, but also “community” between them. In an explanatory remark to the Third Analogy, Kant notes: “The word ‘community’ is ambiguous in our language, and can mean either *communio* or *commercium*. We use it here in the latter sense, as a dynamical community, without which local community (*communio spatii*) could never even be empirically cognized” (A213/B260). Accordingly, the coexistence of objects in space may be prior to the causal relation of community (or mutual interaction) in the order of knowing, but the latter is prior to the former in the way of being, without it being the case that the latter is thereby meant to apply to what might have ultimate metaphysical priority, namely things in themselves.

To see more clearly how epistemological and ontological considerations are combined in the Analogies, consider how this interpretation diverges from those described above. It shares in common with the analytical model the idea that there is a metaphysical connection between temporal and causal relations.²⁹ Yet it is distinct from that model in two ways. First, because this model, unlike the analytical model, is concerned with the conditions for our *knowledge* of objective temporal relations, it is not restricted to conceptual analysis of the meaning of the relevant concepts. Rather, one can attend to whatever is required for the *justification* of our experience that involves those concepts in a particular way. Accordingly, current scientific practice or reliable testimony could be as helpful as conceptual analysis. Second, because this model starts with *knowledge*, it is not vulnerable to the charge that our ideas of reality might not adequately reflect reality. It is important that knowledge of temporal relations be the starting point of the argument since knowledge has ontological import (insofar as it is normatively justified, a feature mere concepts do not have).³⁰

Yet this hybrid model is also different from the model that it is otherwise closest to, namely Guyer's purely epistemological model, because it involves metaphysical considerations. Guyer describes the argument of the Analogies in its most general form as follows: “[B]ecause we take one

²⁹ See below for a more detailed description of this connection.

³⁰ The fact that this interpretation starts with knowledge does not mean that it starts with absolutely certain knowledge, so there is still a limited sense in which it is possible that our “knowledge” might not adequately reflect reality, which coheres with Kant's usage of “*Erkenntnis*,” according to which “*falsche Erkenntnis*” is possible.

set of judgments to be true, another set of judgments, which offer indispensable evidence for the former, must also be taken to be true. . . . The argument is [thus] concerned. . . with the conditions under which we might be able to justify empirical claims to knowledge.”³¹ At this level of generality, the hybrid interpretation can be contrasted with it as follows: “because we take one set of judgments to be true, another set of judgments, which are required on metaphysical grounds by the truth of the first set of judgments, must also be true.” In other words, on this understanding the Analogies do not primarily concern evidence at the level of empirical justification, but rather are concerned with the kind of ontological structure that is required for our empirical knowledge to be true.³² And because the focus is not *exclusively* on epistemological justification, Kant need not be committed to having prior knowledge of causal laws in order to justify knowledge of particular temporal relations. Rather, Kant’s idea in the Analogies is that the world, understood as a whole of mutually interacting substances, must possess ontological unity of a very specific sort, since if it did not, the epistemological unity of time, the singularity of experience, and the unity of apperception would not be possible.³³

³¹ Guyer, *Kant and the Claims to Knowledge*, p. 426.

³² Guyer unambiguously rejects this interpretation as follows: “Here [at A193/B238] – and almost nowhere else – Kant precisely delineates just what it *means* to call a principle such as that of causation a principle of the possibility of experience. It is not to say that such a principle is one which constitutes an empirical object in any ontological sense, nor that it is one which is somehow a psychological precondition of the occurrence of a representation. . . . Rather, to call a principle a condition of the possibility of experience is to say no more and no less than that it is a necessary condition for the *justification, verification, or confirmation* of the judgments about empirical objects that we make on the basis of our representations of them” (*Kant and the Claims of Knowledge*, pp. 245–246). But Kant’s “precise” delineation at A193/B238 does not unambiguously support the epistemological interpretation, since it supports the hybrid interpretation equally well. For after concluding the necessity of one state following another according to a rule, Kant infers “only by that means [*nur dadurch*] can I be justified” in asserting objective succession. But saying that causality is what allows for claims to objective succession to be justified does not imply that causality is itself nothing other than an epistemological “inference ticket,” even if one grants that it must be that as well. According to the hybrid view, it is the ontological status of a cause as a ground that licenses such an “inference ticket,” and neither Guyer nor Kant has said anything at all to exclude that possibility. In fact, as we saw above, the textual evidence, subtle as it may be, actually speaks in favor of it, not to mention the evidence supplied by Kant’s immediate historical context.

³³ This interpretation could be developed in greater detail. For example, one could consider how it relates to the question concerning whether Kant’s ontology is naturalistic, raised, for example, by Georg Sans, *Ist Kants Ontologie Naturalistisch?: Die “Analogien der Erfahrung” in der “Kritik der reinen Vernunft”* (München: Verlag Kohlhammer, 2000), pp. 16–18.

THE ARGUMENT OF THE SECOND ANALOGY

Against the background of this understanding of the status of the Analogies of Experience and their place within the *Critique*, we can now turn to the Second Analogy. Before presenting what I take to be the most adequate reconstruction of Kant's argument in the Second Analogy, consider briefly several ways in which the claim of the Second Analogy has been understood and which argumentative strategies have been pursued in attempting to establish its claim.

Preliminaries

First, while the claim of the Second Analogy – “All alterations occur in accordance with the law of the connection of cause and effect” (B232) – might seem to be relatively straightforward, it has been interpreted in two very different ways. The weak reading, developed in different forms by Buchdahl, Beck, Strawson, and Allison, suggests that it asserts merely that every event must have a cause. Since the notion of causality that Kant attempts to establish contains the notion of necessity that Hume rejects in favor of “constant conjunction,” the conclusion of the “weak” reading is still stronger than at least some would be willing to accept. By contrast, the strong reading, advanced in distinct ways by Guyer and Friedman, holds that the Second Analogy is committed not just to causes, but to causal laws as well.³⁴ It is important to note, however, that Kant distinguishes between different types of laws, in particular, between universal transcendental principles and empirical laws of nature. The latter, Kant says, have an empirical element that cannot be derived from the former, but rather can be discovered only through experience. At the same time, these empirical laws would not be laws if they did not possess some kind of necessity, and Kant thinks that empirical experience is incapable of providing the requisite necessity. Accordingly, the transcendental laws “ground” the empirical laws by supplying them with their necessity. As a result, according to the strong interpretation, Kant also has a reply to Hume's doubts about justifying causal laws that would hold in the

³⁴ Gerd Buchdahl, *Metaphysics and the Philosophy of Science* (Oxford: Basil Blackwell, 1969); Lewis White Beck, *Essays on Kant and Hume* (New Haven: Yale University Press, 1978); Strawson, *The Bounds of Sense*; Allison, *Kant's Transcendental Idealism*; Guyer, *Kant and the Claims to Knowledge*; Michael Friedman, “Causal Laws and the Foundations of Natural Science,” in *The Cambridge Companion to Kant*, ed. P. Guyer (New York: Cambridge University Press, 1992), pp. 161–199.

future. In this chapter, we see that Kant's *explicit* argument in the Second Analogy supports only the weak interpretation. However, in Chapter 4, we discover that considerations deriving from Kant's Second and Third Analogies will *ultimately* support a version of the strong interpretation as well, though one that is different from both Guyer's and Friedman's.

Second, there has been disagreement about the role of the irreversibility of our perceptions in Kant's argument. One way of thinking starts with the distinction between subjective and objective time and notes that one cannot immediately infer the succession of states of an object from succession in the perceptions that constitute our apprehension. But then, it is asked, what does allow us to make the inference, if the *de facto* order of our perceptions is insufficient? The temptation is to think that it is a *modal* feature of our perceptions: In the case of objective succession, the order of our perceptions could not have been other than it is, whereas in the case of objective coexistence, the order of our perceptions could have been reversed. Causality then enters the picture to explain why the order of our perceptions is irreversible. Strawson summarizes the idea quite eloquently: "Briefly, any succession of perceptions is a perception of objective change only if the order of those perceptions is necessary; but the order of the perceptions can be necessary only if the change is necessary, i.e., causally determined."³⁵ As a result, objective succession can be known only if the order of perceptions is irreversible, which can itself be known only if there is some cause that makes that order irreversible. On this line of reasoning, irreversibility thus serves an intermediate link in a chain of necessary conditions for objective succession.

As has been pointed out repeatedly, however, this particular reconstruction of Kant's argument has the disadvantage of attributing to Kant a fallacious argument, in fact, "a non-sequitur of numbing grossness."³⁶ As Strawson explains, this kind of argument "not only shifts the *application* of the word 'necessary' [from expressing a relation between two successive representations to expressing a relation between a cause and a change from a first to a second state], but also changes its *sense*, substituting one type of necessity for another."³⁷ For the sense of necessity involved in representing objectively successive states as successive is conceptual, whereas the sense of necessity according to which the second state actually follows the first is supposed to be causal.

³⁵ Strawson, *The Bounds of Sense*, p. 138.

³⁶ *Ibid.*, p. 137.

³⁷ *Ibid.*, p. 138.

But where exactly does this reconstruction go astray? The problem stems from the way in which it attempts to connect the irreversibility of the order of our perceptions to the necessity of causality. For the illicit slide in both the application and the sense of necessity comes from moving from the conceptual necessity involved in representing successive states as successive to the causal necessity by which a cause brings about a change from an earlier to a later state. Accordingly, to develop a cogent reconstruction of Kant's argument, one must not assume that the irreversibility of the order of our perceptions entails, or is a criterion for, causality.

However, if one does not assume that the irreversibility of the order of our perceptions is supposed to serve as a criterion for necessity, then irreversibility no longer stands at the center of Kant's argument (between causality and knowledge of objective succession). As a result, two questions immediately arise. First, how is Kant's argument supposed to work if irreversibility does not play this mediating role? Second, why does Kant bring up irreversibility if it does not mediate between the two crucial elements in his argument, namely causality and knowledge of objective succession?

In response to the first question, Kant's argument can be understood as suggesting a direct connection between causality and knowledge of objective succession, one that has no mediating link in the way that irreversibility was invoked in Strawson's reconstruction. That is, causality is itself an immediate necessary condition of knowledge of objective succession. While it is true that we may be able to notice the irreversibility of the order of our perceptions more easily than the presence of a causal connection, that is irrelevant to Kant's purposes. For as Kant makes clear in the *Transcendental Analytic*, he needs to show (against the empiricists) that there are nonempirical concepts, or categories, that must be applied to sensible intuition for us to have knowledge (i.e., to make experience possible), and the primary task of the Analogies must likewise be to show that the relational categories must be invoked and knowledge of persistence and temporal relations is simply the means that Kant wants to use to establish their necessity. What is crucial to note is that this construal of the task of the *Transcendental Analytic* does not require that we have any systematic or universally applicable criterion for knowing when the categories must be applied (apart from the knowledge that their application makes possible).

Second, even if irreversibility is thus irrelevant to Kant's attempt to establish a direct link between causality and knowledge of objective

succession, one need not view Kant's discussion of it as confused or superfluous. For Kant recognizes that one must also give an account of the relation between the subjective order in which we apprehend objective sequences and the objective sequences that are thereby represented. And while it was common for early modern philosophers to think that the objective temporal order must be derivative on features of our subjective representations, Kant holds that the converse is the case, that is, that one must "derive the **subjective sequence** of apprehension from the **objective sequence** of appearances" (A193/B238).³⁸ As a result, rather than being a *criterion* from which objective succession can be inferred, irreversibility is a *consequence* of objective succession.

This debate about the role of the irreversibility of the order of our representations in the Second Analogy leads to a final general consideration about what its argument presupposes. Again, there are two radically different lines of thought. A first line, suggested by Strawson, but developed in greater detail by others (including Henrich, Allison, and Guyer), suggests that Kant is undertaking the quite ambitious attempt of refuting the skeptic and can thus presuppose no substantial knowledge that could not be had with apodictic certainty.³⁹ Although this strategy is often coupled with the Transcendental Deduction or the Refutation of Idealism, the Second Analogy in particular is supposed to assume nothing more than either our awareness of the order of our perceptions or our conception of an object (whether characterized as reidentifiable or otherwise) and attempt to deduce causality from it, thereby making experience (in the form of objective succession) possible. In this way, it is hoped, the argument of the Second Analogy can refute skepticism.

By contrast, a second line of interpretation does not presuppose that Kant is attempting to refute the global skeptic (i.e., one who is skeptical with regard to all claims to knowledge about the external world), but rather suggests that Kant is presupposing that we have experience (in the sense specified above, namely experience of a single world) and simply attempting to display various necessary conditions of that experience.⁴⁰

³⁸ Paul Guyer, *Kant and the Claims to Knowledge*, expresses this point well (pp. 246–248).

³⁹ Dieter Henrich, *Identität und Objektivität: Eine Untersuchung über Kants transzendente Deduktion* (Heidelberg: Carl Winter Verlag, 1976).

⁴⁰ Karl Ameriks, "Kant's Transcendental Deduction as a Regressive Argument," *Kant-Studien* 69 (1978): 273–285, and Manfred Kuehn, *Kant: A Biography* (New York: Cambridge University Press, 2001). According to this account, it is important to distinguish different basic concepts of experience – experience, experience of space and time, experience of a single spatio-temporal world – in order to establish what certain principles are supposed to be necessary conditions of.

For the context of the Analogies this interpretation would mean that every object that we could know must be locatable in one and the same time, which is possible only if they are related by temporal relations of succession and coexistence. Thus, by assuming that we have experience, Kant is presupposing that we have knowledge of objective succession and is simply asking how that knowledge is possible. The answer of the Second Analogy is that such fundamental knowledge is possible only if we apply the category of causality, that is, if the world (or the succession of the states of its objects) is governed by causality. While such an interpretation is obviously not as ambitious as the antiskeptical line, its chances of success may be considerably greater and it is by no means a trivial or unimportant undertaking. Establishing that the causal order is more fundamental than or prior to the experienced temporal order is neither easy nor uninteresting, and, if successful, shows that knowledge of the world must employ the categories, a conclusion that Hume certainly would not have wanted to accept. Thus, on this interpretation, even if the Second Analogy is not designed to defeat global skepticism, it can at least attempt to defeat someone who is skeptical about the category of causality.

The Introductory Argument

Although commentators have identified as many as six separate arguments in the text of the Second Analogy or Experience, they can all be seen as different versions of no more than two main argument types. One short, introductory argument, featured most prominently in text added in the second edition (B232–234), relies heavily on Kant's epistemological doctrines, whereas the other, main argument, which Kant formulates and reformulates several times, emphasizing different aspects of it in different contexts, is both more complicated and more controversial.

Kant states his first argument as follows:

I perceive that appearances succeed one another, i.e., that a state of things exists at one time the opposite of which existed in the previous state. Thus I really connect two perceptions in time. Now connection is not the work of mere sense and intuition, but is here rather the product of a synthetic faculty of the imagination. . . . This, however, can combine the two states in question in two different ways, so that either one or the other precedes in time; for time cannot be perceived in itself. . . . [T]hrough the mere perception the **objective relation** of the appearances that are succeeding one another remains undetermined. Now in order for this to be cognized as determined, the relation between the two states must be thought in such a way that it is thereby necessarily determined which of them must be placed before and which after rather than vice versa. The

concept, however, that carries a necessity of synthetic unity with it can only be a pure concept of understanding . . . and that is here the concept of the **relation of cause and effect**. (B233–234)

Kant's argument can be reconstructed as follows:

- P1 Objective succession is a connection between two appearances (i.e., between the states of an object that can appear to us).
- P2 Intuition does not provide knowledge of any connection.
- C1 Intuition does not provide knowledge of objective succession. (from P1 and P2)
- P3 The imagination's syntheses can represent a connection.
- P4 Objective succession is not just any connection, but a *necessary* connection according to which one state of an object must precede a second state of the object.⁴¹
- P5 The imagination's syntheses cannot represent a necessary connection between appearances (i.e., the states of the object), since it can represent only a *contingent* relation between perceptions [or states of an object].
- C2 The imagination's syntheses cannot represent objective succession. (from P4 and P5)
- P6 There is an exhaustive disjunction between sensibility's intuitions, the imagination's syntheses, and the understanding's categories.
- C3 Only the categories can represent a necessary connection between states of an object. (from C1, C2, and P6)
- P7 The only category that can represent a necessary connection between successive states of an object is that of causality.
- C4 Causality is a necessary condition for representing and thus knowing objective succession. (from C3 and P7)⁴²

This argument maps onto our preliminary considerations as follows. The first part of the argument, which establishes that intuition cannot give us knowledge of objective succession, turns on the fact that there is a distinction between subjective and objective time. Intuition gives us merely subjective temporal information and not objective knowledge. The second part of the argument, which establishes that the imagination cannot provide the requisite knowledge, turns on Kant's idea that objective succession involves the irreversibility of the order of our perceptions. For

⁴¹ Though Kant does not explicitly mention the irreversibility of the order of our apprehension in this passage, it could be thought to be implicit in his emphasis on what an objective relation of succession would naturally entail.

⁴² To be explicit about the conclusion of this argument, it is conditional in nature. If we represent or have knowledge of objective succession, then the category of causality must be employed. Also, while the argument focuses on the conditions for *representing* succession, insofar as representation is necessary for knowledge, the argument can still be viewed as attempting to establish the category of causality as a necessary condition for knowledge of objective cognition.

whatever the imagination might represent is contingent in the sense that the imagination could have represented it otherwise and thus is incapable of representing an objective sequence, that is, one that entails the irreversibility in question. The third step of the argument, establishing the necessity of the category of causality, proceeds on the assumption that it is the only viable representation that remains, as well as on the claim, established in the Metaphysical Deduction, that the categories can represent necessity.

This argument does not require extensive discussion. It does fit with our characterization of what the Analogies of Experience are supposed to accomplish and how they fit in with the rest of the first *Critique*. Also, it clearly presupposes (rather than establishes) knowledge of objective succession and attempts to argue that the category of causality (and not knowledge of causality) is necessary for such knowledge. At the same time, it is unclear how much weight this argument is supposed to bear. For one, it is an argument from elimination that presupposes Kant's distinctive division of our epistemic faculties and thus would seem not to carry any force independent of the argument for his account of our faculties, an account that is not justified by specific or detailed argument in the *Critique*. Also, although Kant does not explicitly indicate this point, what seems to carry the burden of the proof here is the Transcendental and Metaphysical Deductions. For those two arguments jointly show that there are a limited number of categories that can be legitimately employed to obtain knowledge. This argument in the Second Analogy would seem merely to apply the conclusion of that argument to the special case of objective succession. As a result, the argument here is not especially interesting in its own right, even if it is an argument that immediately suggests itself given what Kant thinks he has established earlier in the *Critique*.

The Main Argument

Rather than simply relying on his earlier arguments in the *Critique*, Kant devotes considerable attention to the details of causality in the course of developing a second, main argument throughout much of the rest of the text in the Second Analogy. I reconstruct the primary steps of this argument as follows.

P1 Apprehension of objects (the subjective order of perceptions) is always successive.

- P₂ There is a distinction between the subjective order of perceptions and the successive states of an object such that no immediate inference from the former to the latter is possible.
- C₁ One cannot immediately infer objective succession from the successive order of perceptions. (from P₁ and P₂)
- P₃ To have knowledge of objective succession, the object's states must be subject to a rule that determines them as successive.
- P₄ Any rule that determines objective succession must include a relation of condition to conditioned, i.e., that of the causal dependence of successive states on a cause.
- C₂ To have knowledge of the successive states of an object, the object's successive states must be dependent on a cause, that is, must stand under a causal rule. (from P₃, P₄, and C₁)

P₁ states an obvious fact about our experience, while P₂ follows straightforwardly from the problem of time-determination. (P₁ and P₂ together imply that we cannot perceive time itself.) P₃ introduces the idea of a rule that is supposed to make knowledge of objective succession possible. Kant seems to be expressing this point at A193/B238 when he argues that the kind of connection between states that is necessary for objective succession

must therefore consist in the order of the manifold of appearance in accordance with which the apprehension of one thing (that which happens) follows that of the other (which precedes it) **in accordance with a rule**. Only thereby can I be justified in saying of the appearance itself, and not merely of my apprehension, that a sequence is to be encountered in it.

P₄ then characterizes the rule introduced in P₃ as a causal rule. Kant states this point at A193–194/B238–239:

In accordance with such a rule there must therefore lie in that which in general precedes an occurrence the condition for a rule, in accordance with which this occurrence always and necessarily follows. . . . I must necessarily relate it [i.e., the succession] to something else in general that precedes, and on which it follows in accordance with a rule, i.e., necessarily, so that the occurrence, as the conditioned, yields a secure indication of some condition.

As it is clear from the context that the condition-conditioned relationship mentioned in P₄ must be understood as causal, we have adequate textual support for P₁–P₄.

Before we turn to evaluating the crucial steps in Kant's argument, P₃ and P₄, consider Guyer's summary of the crucial step of the argument: "[P₃'] Only from a rule which says that one of the represented states *must* succeed the other can it be inferred that it *does* succeed the other. . . . And

[P₄'] a rule which dictates that in a given situation one state of affairs must succeed another is just what Kant means by a causal law" (brackets and numbering added).⁴³ While P₃ and P₃' both introduce the idea of a rule and P₄ and P₄' both characterize that rule in causal terms, there are important differences between these two reconstructions of Kant's argument. The most fundamental difference is that Guyer's version explicitly indicates that the rule in question must be a *necessary* rule. Moreover, Guyer understands the necessity as obtaining between the first state and the state that follows it.

It is difficult to see – at this point in the argument and with the resources introduced so far – what justifies the necessity of the causal rule that determines objective succession. While Guyer is certainly right that one can infer that one state *does* succeed another from the fact that the one state *must* succeed the other, it is patently false that this is the *only* justified inference to objective succession. In particular, it would seem to be entirely possible that objective succession could follow from a *contingent* rule, for example, a rule that determines that state A does in fact happen to occur before state B, even if we can imagine worlds in which it does not occur in that order. In other words, the necessity of the rule that Guyer invokes in the argument on Kant's behalf appears to be "unjustified overkill" insofar as "necessary succession" seems to go beyond "actual succession" and no reason has been given why Kant could not have gotten by with less.⁴⁴ As a result, it is unclear that Guyer's reconstruction is capable of establishing causality as a necessary condition of knowledge of objective succession.

To gain a better understanding of the role of necessity in Kant's argument, consider the various ways in which he uses the term "necessary" in the text of the Second Analogy. In a majority of cases the term is used with respect to the order of our perceptions. The idea is simply that *given* a certain instance of objective succession, the order of our perceptions is necessary. That is, Kant repeatedly claims that irreversibility follows from objective succession. One might object that the modal element contained in the idea of irreversibility could not follow from a purely factual matter. However, such an objection would be based on a misunderstanding of the kind of modal claim Kant is asserting in these contexts. For the modal

⁴³ Guyer, *Kant and the Claims of Knowledge*, pp. 248–249.

⁴⁴ Specifically, a contingent generalization of the sort "All As cause Bs to follow" would seem to suffice to license an inference to the conclusion that, given the occurrence of an A, B will follow. "Necessarily, all As cause Bs to follow" would therefore seem to be more than is required.

operator involved in irreversibility is clearly conditional in nature. *Given* an instance of objective succession, the order of our perceptions of it could not have been otherwise. Of course, if we had encountered a different instance of objective succession, then the order of our perceptions could have been different.

Yet not all of Kant's uses of the term "necessary" in the Second Analogy pertain to the order of our perceptions. Some of them do apply to succession in the object. For example, Kant states: "the appearances themselves must determine their positions in time for each other, and make this determination in the temporal order necessary, i.e., that which follows or happens must succeed that which was contained in the previous state in accordance with a general rule" (A200/B245). What is to be made of uses such as this? If it could be established that necessity in the form of a causal law is the *only* means at our disposal in attaining knowledge of objective succession, then Guyer's reconstruction could perhaps be justified after all. This would fit well with the point of Kant's first argument, discussed above, which proceeded by a process of elimination and left the category of causality as the sole means for representing the necessity involved in objective succession. Since the categories, as pure concepts of the understanding, bring necessity with them, even though the element of necessity they have is not, strictly speaking, required for the purposes of his argument, it would suffice for the task at hand.

However, if the argument by elimination cannot carry any independent weight, then it becomes less clear that Guyer's reconstruction can be supplemented in this way. Moreover, even if Kant does speak of necessity with respect to objective succession, it is crucial to notice exactly how Kant applies the notion of necessity in that context. For what he says is that a cause necessarily brings about its effect. But if the cause need not be the first state of an instance of objective succession (as Guyer rightly notes in response to one of Schopenhauer's objections), then the relation of necessity need not obtain between the successive states, but rather between the cause – whatever it may be – and the effect, that is, the succession of states. But if the necessary relation obtains between the cause and the successive states of its effect rather than between the successive states themselves, then Kant is justified in saying that there are necessary connections in nature, but then the justification for asserting a necessary connection between the successive states has disappeared.

A similar problem arises with Guyer's P₄'. If P₃' goes beyond what is minimally required to account for knowledge of objective succession by positing necessity, then P₄' likewise goes too far in understanding this rule

as a causal law. To see this, recall that for Guyer the point of the Second and Third Analogies of Experience is to show that knowledge of objective temporal relations is justified only if knowledge of causal laws is assumed. Although knowledge of causal laws could license "inference tickets" to such knowledge, it is difficult to see that such knowledge is truly *necessary* to this end. After all, it is at least in principle possible that we could have other kinds of knowledge (e.g., expert or even divine testimony or purely descriptive accounts) that would still license such inference tickets.⁴⁵ In short, if the Analogies were focused exclusively on justification, then rules that licensed inference tickets to knowledge would not necessarily have to be causal laws.

In light of the specific difficulties that P₃' and P₄' encounter, we can now see more clearly what the relative advantages of P₃ and P₄ are. Since what was controversial about P₃' was the fact that it represented the successive states as necessary, and P₃ does not contain necessity in that place, there should be no problem with P₃. Moreover, because Kant's argument is to be understood as attempting to reveal ontological conditions for knowledge of objective succession, it is much easier to see why one might think that the laws required by P₄ would have to be causal. For what we mean by causality is not merely that it licenses inference tickets but that it necessarily brings about its effect.

To understand more fully the justification for P₄ – more specifically, why causality rather than any other kind of connection is required for knowledge of objective succession – it is helpful to compare the structure of the Second Analogy's argument with certain aspects of Kant's pre-Critical views. If the pre-Critical Kant consistently held (in fact, quite explicitly after 1763) that real grounds are responsible for positing determinations and the Critical Kant is investigating in the Second Analogy how temporal determinations are possible, then it appears that in the Second Analogy Kant is simply looking for real grounds of the temporal determinations of objects.⁴⁶ If, however, temporal determinations

⁴⁵ It is true that the only empirical judgments the justification of which is at issue in the Second Analogy are judgments about the temporal order of states of objects, and thus that the only inference tickets that could be relevant must concern the temporal sequences of states of objects, but that still falls short of requiring that such judgments must be about causal laws. For example, we could imagine a set of purely descriptive statements written in a book (titled *The Book of Nature*) from which we could be warranted in accepting inference tickets for judgments about the temporal order of states of objects.

⁴⁶ At A196/B241–242 Kant explicitly identifies the rule that allows knowledge of objective succession with a necessitating ground.

pertain to objects only via their states or features (just as was the case in his pre-Critical period), then the real grounds of the temporal determinations of objects Kant is seeking in the Second Analogy must be the real grounds of those determinations of objects that involve their states or features. But the real ground of the determination of the state or feature of an object is just what a cause is. Accordingly, it is simply an analytic truth that determining the temporal states of objects requires causality. In other words, because (1) determinations are posited only by means of grounds, (2) the temporal determinations Kant is concerned with in the Second Analogy are temporal determinations of the states of objects, and (3) grounds of the determinations of states of objects are simply the causes of those states, it follows that causality is required for the temporal determination of objective succession.

One might object, however, that there are important differences between Kant's pre-Critical account of real grounds and the kind of temporal determination involved in the Second Analogy and that these differences are improperly minimized if the latter is cast in terms of the former. More specifically, one might argue that in the pre-Critical period, Kant is arguing exclusively at the level of metaphysics, whereas in the Second Analogy he is undertaking an epistemological task insofar as he is explaining knowledge, and that the resources of the former cannot account for the latter. To relate this objection back to the reconstruction provided above, one might think that the term "determination" is used ambiguously in P₃ and P₄. If it is understood epistemologically, then P₃ is true, but then P₄, which clearly involves a metaphysical sense of "determination," does not follow. If it is understood metaphysically, then one could raise doubts about why P₃ must be true, since it might be unclear why metaphysical determination of the succession of states is required for knowledge. While Guyer does not explicitly raise this objection, one can easily surmise that the Analogies' emphasis on knowledge could have motivated him to think that only *knowledge* of causality could be adequate to account for our *knowledge* of objective succession.

The crucial premise for this objection is P₃, which specifies that a rule determining the succession of states is a necessary condition of knowledge of objective succession. One aspect of P₃ is uncontroversial. Knowledge of objective succession requires that objective succession actually occur, for it would otherwise not be an instance of knowledge.⁴⁷ Accordingly,

⁴⁷ As we saw above (in note 30), it is true that Kant's use of the term "*Erkenntnis*" allows that "knowledge" can be false (which perhaps makes the term "cognition" a better translation

any question about this premise must be about the rule that determines the object's successive states. Why should that rule be understood metaphysically and how can it explain something epistemological? As for the first question, Kant's idea is simply that any determination (and therefore successive determinations as well) requires a ground to posit it, since otherwise the object in question will be indeterminate in that respect. Moreover, in light of the problem of time-determination, it is clear that this indeterminacy is not purely epistemological (indicating simply a lack of knowledge), but metaphysical, given that time (and thus any state that it attaches to directly) is not an independently existing entity whose properties we merely have to discover, but rather something that must be constituted.⁴⁸ In response to the second question, one should keep in mind that Kant is merely attempting to show that causality is a necessary condition of knowledge of objective succession. Since there is no reason to think that something metaphysical could not be a necessary condition of something epistemological, no problem arises on this count either.

If Kant's argument in the Second Analogy can be reconstructed in this manner, it still remains to be considered whether it requires only that every event has a cause (or occurs according to a causal rule) or whether it also establishes causal *laws* (which would have strict universality and involve types of events). In the text of the Second Analogy, Kant typically restricts himself to use of the term "rule," leaving "law" for either the Principle of the Second Analogy itself or other substantive principles, such as that of (the) continuity (of change). At the same time, Kant does suggest that an effect "always and necessarily follows" from its cause (A193/B238) and that succession must occur "in accordance with a general rule" (A200/B245), both of which certainly suggest that he is thinking of causal laws (i.e., that a certain kind of thing must always

in some contexts). However, "knowledge" cannot be false in a systematic and global way in the context of Kant's overall project.

⁴⁸ In other words, I am suggesting that Kant's argument in the Second Analogy cannot be properly understood if it is interpreted as *neutral* between epistemological and metaphysical models, as Guyer's line could be interpreted. I have already raised objections to interpretations that come down entirely on the side of epistemological models (not because they do not assume enough, but rather because they assume too much). Now I would simply suggest that each move in Kant's argument must admit of either an epistemological or a metaphysical interpretation and that while some moves are clearly epistemological, others make good sense (e.g., can avoid the objections discussed above) only if understood metaphysically. Hence, we get an interpretation that is a "hybrid" of epistemological and metaphysical elements.

cause the same kind of effect).⁴⁹ Moreover, if one can somehow infer from a causal rule to a causal law, Kant would be in a position to reply to both of Hume's skeptical doubts about causality with one and the same argument.

However, in light of the reconstruction of Kant's argument provided above, it is not immediately obvious how the stronger claim could be justified at this point. For what the argument shows is that knowledge of an instance of objective succession presupposes some cause that contains the ground of the successive determinations that constitute the event. If causality is required for knowledge of the succession of states of an object, it is difficult to see how one would be justified in asserting anything about future instances of objective succession (beyond the claim that they too must have some cause or other). We see below, in Chapter 4, that Kant's model of causality does supply resources on the basis of which he could feel justified in asserting the necessity of causal laws.

Before turning to the Third Analogy of Experience, it may be helpful to conclude our discussion of the Second Analogy with a few brief remarks about the argument and the text of the Second Analogy. The first point to note is that the argument is not obviously invalid; it is not a non sequitur of any kind, much less of numbing grossness. To some, the argument might appear to be too weak insofar as it does not refute the global skeptic, but there is no compelling historical or textual reason to think that Kant is attempting that sort of refutation in the Second Analogy, and even the greatest of philosophers need not take on the most ambitious of all projects. To others, the argument might seem to be not too weak, but rather irrelevant insofar as it turns primarily on the problem of time-determination, but, again, as is clear from our discussion of Kant's predecessors in Chapter 1, the problem of time-determination was a real philosophical issue at the time and Kant can simply be seen as addressing it in a particularly powerful way.

Second, it is worth being explicit that the text of the Second Analogy extends well beyond the argument that has been reconstructed above. For example, Kant is also interested in distinguishing between subjective representations of apprehension and the objective representations of knowledge, given that both are simply representations and would thus seem to have the same status. He also seems to think that one moment in time necessarily determines the next moment in time, which runs parallel to the idea that the appearances at one moment in time determine the

⁴⁹ Cf. also A198/B243–244.

appearances at the next moment. Finally, the text of the Second Analogy contains discussions of more familiar topics such as the problem of simultaneous causation and the principle of continuity. We put off to Chapter 4 a discussion of some of the complex issues raised in these “nonargumentative” passages, since some of the resources necessary to understand them properly can be appreciated only on the basis of contributions made by the Third Analogy.

THE ARGUMENT OF THE THIRD ANALOGY

Despite minor differences in formulation between the first and second editions, the main claim of the Third Analogy is that knowledge of objective simultaneity (or coexistence) requires that substances stand in thoroughgoing community or mutual interaction.⁵⁰ In the second edition Kant adds the restriction that this claim holds only for spatial substances. Though we shall have reason to return to the issue of what exactly mutual interaction is, it should be obvious that Kant is hereby rejecting pre-established harmony and occasionalism, since mutual interaction is clearly a form of causal interaction between finite substances.

As was the case in the Second Analogy, Kant presents two separate arguments for his claim in the Third Analogy. While the shorter of the two arguments, added in the second edition, focuses exclusively on epistemological considerations, the longer, main argument is present in both editions and combines epistemological and metaphysical considerations. The parallels between the arguments of the Second and Third Analogies are quite striking.⁵¹ This makes it surprising that the Third Analogy has been curtly dismissed or even ignored altogether by the vast majority of commentators on the Second Analogy.⁵² As we see in Chapter 4, the

⁵⁰ For a different reading of Kant's argument, see Jeffrey Edwards, *Substance, Force, and the Possibility of Knowledge: On Kant's Philosophy of Material Nature* (Berkeley: University of California Press, 2000), who holds that Kant is attempting to exclude the possibility of a void by establishing a “material” ground of experience, an idea Kant develops explicitly much later in the *Opus postumum*.

⁵¹ It is true that the Second Analogy involves change, whereas the Third Analogy does not. However, even so, the text of the Third Analogy (esp. A211/B258) strongly suggests that the states whose simultaneity is to be known would endure for a period of time, during which one *could have* apprehended them in a different order.

⁵² Without providing a complete list of negative evaluations, virtually all of the main commentators on Kant's first *Critique* either neglect it altogether (e.g., Allison, Kitcher, and James van Cleve, *Problems from Kant* (New York: Oxford University Press, 1999)) or dismiss it with almost no sustained or independent analysis (e.g., Strawson, Bennett, and Arthur

argument of the Third Analogy turns out to be much more informative about Kant's views on causality than is the Second Analogy. In the rest of this chapter, Kant's arguments in the Third Analogy are reconstructed and discussed, before we turn to consider several qualifications Kant makes about the conclusion of the Third Analogy.

The Introductory Argument

Kant states a first, introductory argument as follows:

Things are **simultaneous** if in empirical intuition the perception of the one can follow the perception of the other **reciprocally**. . . . Now simultaneity is the existence of the manifold at the same time. But one cannot perceive time itself and thereby derive from the fact that things are positioned at the same time that their perceptions can follow each other reciprocally. The synthesis of the imagination in apprehension would therefore only present each of these perceptions as one that is present in the subject when the other is not, and conversely, but not that the objects are simultaneous. . . . Consequently, a concept of the understanding of the reciprocal sequence of the determinations of these things simultaneously existing externally to each other is required. . . . Now, however, the relation of substances in which the one contains determinations the ground of which is contained in the other is the relation of influence, and, if the latter reciprocally contains the ground of the determinations of the former, it is the relation of community of interaction. (B256–258)

Kant's argument can be reconstructed as follows:

- P₁ If the order of perceptions of two substances (or states of two substances) in intuition is (or can be known to be) reversible, then the substances (or their states) are (or can be known to be) simultaneous.
- P₂ One cannot perceive time itself in intuition.
- P₃ If one cannot perceive time itself in intuition, then one cannot know the reversibility of the order of one's perceptions in intuition from the perception of time itself.
- C₁ One cannot know the reversibility of the order of one's perceptions in intuition. (from P₂ and P₃)
- C₂ One cannot know the simultaneity of two substances (or their states) in intuition. (from P₁ and C₁)
- P₄ The synthesis of distinct perceptions in the imagination is successive.

Melnick, *Kant's Analogies of Experience* (Chicago: University of Chicago Press, 1973). Two exceptions to this are Paul Guyer (who is ultimately critical of the argument) and Beatrice Longuenesse (who is primarily interested in arguing for a particular interpretation of how to understand the relation between Kant's table of judgments and the possibility of perceptual experience).

- P₅ If the synthesis of distinct perceptions in the imagination is successive, the imagination cannot represent two substances (or their states) as simultaneous.
- C₃ The imagination cannot represent two substances (or their states) as simultaneous. (from P₄ and P₅)
- C₄ Thus, a pure concept of the understanding is required in order to have knowledge of objective coexistence and to warrant a claim of reversibility. (from C₂ and C₃)
- P₆ The relation of substances in which the one contains determinations whose ground is contained in the other and vice versa is the relation of community or mutual interaction and can be represented by a pure concept of the understanding.
- C₅ The simultaneity of substances can be known only if the substances stand in community or mutual interaction. (from C₄ and P₆)

With a few notable differences, this argument runs analogously to the introductory argument that was presented for the Second Analogy. Both arguments (1) assume (on the basis of the problem of time-determination) that intuition does not provide knowledge of time itself, (2) explain what knowledge of a particular kind of temporal relation between objects entails (irreversibility and reversibility, respectively), and (3) conclude that neither intuition nor the imagination is capable of giving us that knowledge. As a result, they both establish that a category of the understanding (causality and mutual interaction, respectively) is required. The most important differences between the arguments (to which we have reason to return) are (1) that this argument explicitly invokes substances and (2) that there are important differences between the categories of causality and community. However, neither of these differences changes the fact that the force of this argument is limited in precisely the ways in which the introductory argument of the Second Analogy is and that it thus does not bear significant independent argumentative weight.

The Main Argument

In light of the parallels between both the claims and the introductory arguments of Kant's Second and Third Analogies and given the main argument of the Second Analogy, it comes as no surprise when Kant presents an argument for the Third Analogy, the "main" argument, that can be reconstructed as follows:

- P₁ Apprehension of substances (the subjective order of perceptions) is always successive.
- P₂ There is a distinction between the subjective order of perceptions and the temporal relations (of the states) of substances.

- C₁ One cannot immediately infer objective coexistence from the successive order of perceptions. (from P₁ and P₂)
- P₃ To have knowledge of objective coexistence, the substances' states must be subject to a rule that determines their states as coexistent.
- P₄ Any rule that characterizes objective coexistence must include reciprocally conditioned conditions, that is, a relation of mutual interaction.
- C₂ In order to have knowledge of objective coexistence, substances must stand in mutual interaction.

The three central paragraphs of the Third Analogy argue that (1) the states of substances can be known to coexist if the order of our perceptions of their states in apprehension is indifferent (or reversible), (2) if substances are causally isolated, then their coexistence cannot be known, and (3) thoroughgoing community is required in addition to the mere existence of substances in order for their coexistence to be known. How do these three paragraphs relate to and support the above reconstruction?

The first, fairly brief paragraph (at A₂₁₁/B₂₅₈) simply explains which conditions must hold for us to have knowledge of coexistence by noting that if several substances coexist, then the order of our apprehension of them is "indifferent" or reversible. The second paragraph (A₂₁₂/B_{258–259}) then argues that this condition would not be satisfied if substances were causally isolated. For, as Kant puts it, "the perception that proceeds from one [substance] to the other in time would certainly determine the existence of the latter by means of a succeeding perception, but would not be able to distinguish whether that appearance objectively follows from the former or is rather simultaneous with it" (A₂₁₂/B₂₅₉). That is, if two substances were completely causally isolated from each other, we would have no reason to interpret the successive apprehension of first one and then the other as an instance of coexistence rather than succession. Accordingly, if we have knowledge of coexistence, the coexisting substances cannot be causally isolated. In this way, Kant has developed an argument that refutes both Crusius's and Leibniz's positions, despite the fact that it is importantly different from (though also similar to) his arguments in the principles of succession and coexistence in the *Nova dilucidatio*.

These two paragraphs thus accomplish the equivalent of P₁, P₂, and C₁, though the second paragraph in particular illustrates more clearly how the problem of time-determination applies to substances lacking causal relations. For the point is not merely that we cannot immediately know the coexistence of two substances on the basis of what is given to

us in subjective time, but also that we cannot attain such knowledge if, ontologically speaking, the substances were to lack causal relations to each other. For, as we saw in a slightly different form in the Second Analogy, Kant thinks that temporal relations are possible only on the basis of causal relations. The fact that these steps are not discussed in greater detail is not surprising, given Kant's earlier treatment in the Second Analogy. As a result, it is clear that P₃ and P₄ – the crucial premises here just as they were in the case of the Second Analogy – are supported in a single paragraph, namely, the third paragraph at A212–213/B259–260, which argues that since the “mere existence” of substances is insufficient for knowledge of coexistence, mutual interaction is necessary for such knowledge.

The basic idea behind this final step is quite simple. If the Second Analogy has established that a causal rule is necessary for time-determination and if simultaneity requires the time-determination of two substances, then it would be natural to suppose that each substance would bring about the time-determination of the other substance. If a substance could determine itself, then causally isolated substances *could* be known to coexist, which goes against P₁, P₂, and C₁, which had been established by the previous paragraphs. It is equally natural to think that such a two-way causal determination would be called mutual interaction. Accordingly, one can see fairly easily how these assumptions might have naturally led Kant to mutual interaction.

There are, however, important complications affecting the details of this argument. The first point to note is that Kant has switched from talking about objects and their successive states (or events) in the Second Analogy to explicitly invoking substances in the Third Analogy. Given that Kant has devoted considerable attention in the First Analogy to arguing for the necessity of substance, he has every right to speak of substances in the Third Analogy. However, the switch to substances leads to a potential objection to the whole strategy of Kant's argument in the Third Analogy. Because Kant defines substance in terms of permanence, one might wonder why one could not infer from this immediately (i.e., without invoking mutual interaction) that all actual substances coexist. After all, if all substances are permanent, that is, exist at all times, how could they *not* exist at the same time?

To see why this objection is mistaken, it is necessary to understand in more detail both Kant's conception of substance and the problem of time-determination. The crucial idea here lies in understanding properly the relation between a substance and its states. In particular, it is possible to know *that* two substances exist simultaneously without knowing *how* they

exist simultaneously, that is, without knowing what any of their states are at any given time. That Kant is aware of this kind of point is brought out by the passage in the second main paragraph of the Third Analogy that was just considered, where he remarks that in the case of two isolated substances, one could “certainly determine the existence of the second substance by means of a succeeding perception, but would not be able to distinguish whether that appearance objectively follows the former or is rather simultaneous with it” (A212/B259). That is, it is possible that we could know that a substance exists and even that it exists in a particular state, and still not know when it is in that state.

But why need one be concerned about *how* a substance exists, that is, what state it is in, at a given time? Two features of Kant’s conception of substance are relevant here. First, Kant thinks that a substance cannot exist without any determinations (or accidents), since a substance must exist in some way or other.⁵³ In other words, even if a substance were a bare particular, it could never exist without existing in some state or other, that is, without any positive determination. Even if Kant shies away from claiming that phenomenal substances could ever be completely determined (in the way that things in themselves must be), it is clear that they must be determined in some minimal way. Second, only the states, determinations, or accidents of a substance can be known. Therefore, it is the case not only that a substance can appear only along with at least one of its determinations, but also that it can appear only *through* such determinations. Accordingly, insofar as Kant is interested in phenomenal substances, that is, substances that we must be able to represent to ourselves empirically, it is impossible for us to represent the coexistence of two substances apart from their temporally determinate states. To put the point in terms of the problem of time-determination, determining the place of a substance in time can occur only if the states of that substance are determined in time, that is, if it is determined that a substance is in a certain state at a certain time (and not if it is determined that the substance simply exists at a certain time). Though this point was not emphasized in the Second Analogy, it was implicit there as well. For the crucial feature of the causal rule that allows for knowledge of objective succession is that it determines the successive *states* of an object (rather than determining the successive temporal indices of objects themselves, independently of their states).

⁵³ See A186/B229–230 for some related claims about accidents and positive determinations.

A second complication arises as follows. Even if one grants the first step of the argument – which establishes that substances cannot be causally isolated if we are to have knowledge of their coexistence – one might still question whether *mutual* interaction is required. In other words, why not think that one substance does “all the work,” determining the place in time of both the other substance and itself? The Second Analogy provides no help on this point, since it leaves completely undetermined what the cause of the succession of states might be and thus provides no resources with which to exclude the possibility that one substance might determine both itself and another substance at the same time.

The third main paragraph of the Third Analogy expresses the crucial step of the argument as follows:

In addition to the mere existence there must therefore be something through which A determines the position of B in time, and conversely also something by which B does the same for A, since only under this condition can those substances be empirically represented as **existing simultaneously**. Now only that determines the position of another in time which is the cause of it or its determinations. Thus each substance . . . must simultaneously contain the causality of certain determinations in the other and the effects of the causality of the other, i.e., they must stand in dynamical community. (A212/B259)

There are two separate steps in this argument. First, there is the idea that a substance cannot determine its own place in time. Second, there is the idea that time-determination can occur only through the (causal) determination of each substance's states. Putting these two ideas together entails that for a simple, “closed” system of two substances, for the place in time of both substances to be determined (which is obviously necessary to have knowledge of their coexistence) each one must act on the other, which is what mutual interaction is.⁵⁴

If we have already considered Kant's reasoning for the second premise, that is, for thinking that determination must be causal determination, only the first premise remains. Why does Kant think that a substance cannot determine its own place in time? Since the text of the Third Analogy does not provide any explicit resources to answer this question, it will be helpful to appeal to other means, in particular, to various lines of reasoning presented in the pre-Critical period. By taking the pre-Critical period into account, we can see two possible justifications for the claim that a substance cannot determine its own place in time.

⁵⁴ There are a host of complications involved in the details of this argument, but we defer discussion of them to Chapter 4 to focus on the crucial move in this argument.

The first possible line of justification is based on Kant's understanding of relations as developed in the *Nova dilucidatio's* principle of coexistence.⁵⁵ As we saw in Chapter 2, in the *Nova dilucidatio* Kant held that relations are distinct from the intrinsic properties of things in the sense that they are not grounded solely in the intrinsic properties of things. Something else must be added to two things so that they can stand in relation (e.g., causal or spatial) to each other. There he seemed to have something such as distance in mind as an example. For God could create a world with two substances three feet from each other, but he could also have created a world with those very two substances two feet from each other, without having to change any of the intrinsic properties of either substance. At the same time, relations are not completely independent of the intrinsic properties of substance. Given that two substances stand in a certain relation to each other at a certain moment in time, the powers of these substances will help to determine their relations to each other at the next moment in time. Thus, if two substances with attractive forces are three feet apart at one moment in time, both their distance and the strength of their attractive forces will determine how much closer to each other they will move.

One might think that Kant's argument in the Third Analogy results from applying a similar conception of relations to time-determination as follows. Although "determining a substance's place in time" could, under certain circumstances, be an intrinsic property (e.g., in those cases where time is defined relative to changes within a single substance), it is clear that in the case of the simultaneity of a plurality of substances, it must be a relational property. If relational properties depend, at least in part, on the intrinsic properties of substances, then simultaneity will depend on the intrinsic properties of substances as well. But just as it was the causal powers of substances that helped to determine the spatial relations of substances in the *Nova dilucidatio*, so, too, it will be causal powers here that help to determine the temporal relations of simultaneous substances. As a result, so the argument goes, simultaneity requires mutual interaction (in the form of the causal powers of the simultaneous substances).

However, as we saw above, such an argument fails to establish the necessity of mutual interaction. Even if relations do depend on intrinsic properties and even if simultaneity does depend on causal powers, the

⁵⁵ I am grateful to Desmond Hogan for suggesting (though not necessarily endorsing) a more elaborate version of this kind of argument in "Cosmological Unity and Universal Interaction in Kant's *True Estimation of Living Forces*" (manuscript).

argument does not show that the kind of causality that must be invoked is mutual interaction. More specifically, the argument, so understood, does not rule out the possibility that the causal power of just one of the two substances is required to establish the simultaneity of the two substances. Accordingly, using Kant's pre-Critical conception of relations in this way cannot establish the necessity of mutual interaction.

A second line of justification derives from the *Nova dilucidatio's* principle of succession. Recall that the principle of succession asserts that reciprocal change requires mutual interaction. The primary justification for this claim was the idea that a substance could not cause a change in itself, since (1) changing grounds are inconsistent with their role as fundamental constituents of a substance, and (2) unchanging grounds cannot posit first one set of determinations and then another if determinations are to be simultaneous with their grounds. In considering Kant's own positive explanation of change, we also faced the question of whether mutual interaction or rather a merely one-directional causal bond was required to account for change. Since Kant's primary concern was with *reciprocal* change, it was clear that mutual interaction would be required (insofar as reciprocal change entailed a change in intrinsic properties of both substances involved). In fact, the very possibility that one substance could cause changes in both another substance and itself appeared to be incoherent within Kant's framework. For if a substance contained a ground of change in the other and thereby changed itself, then the second substance would in fact contain a ground for the change in the first substance insofar as the change in the first substance would not have occurred, had the second substance not existed or been incapable of being determined by the first substance. To put this final idea in different words: Insofar as the second substance is a condition for change in the first substance, the first substance is not in fact capable of changing itself without a causal contribution from another substance.

If Kant continued to accept the essential features of his pre-Critical account of the role of grounds in explaining change, how might that be applied in the current context of attempting to establish the necessity of mutual interaction for simultaneity? Although there are obvious differences between mutual change and simultaneity, it is the similarities that can be put to use. What is crucial to establishing the necessity of mutual interaction in the *Nova dilucidatio* is the idea that a reciprocal change has implications for the intrinsic determinations of both substances (even if, in the Critical period, such "intrinsic" determinations are not absolutely intrinsic in the sense of being indivisible, but rather merely comparatively

so). More specifically, because those intrinsic determinations of both substances that are implied by reciprocal change must be accounted for, each substance must be understood as a condition, and thus as a cause, of determinations in the other (since it is granted that they cannot cause a change in their own intrinsic determinations). Simultaneity displays these same features. Because (1) simultaneity is a relational property that entails that the place in time of both substances be determined, (2) the place in time of each substance can be determined only if their states are causally determined, and (3) they cannot causally determine their own states, they must stand in mutual interaction.

To avoid misunderstanding, it should be noted that Kant does not rule out the possibility that a substance could act on itself per se. If a substance's essence consists of grounds that immediately posit its essential properties, then it is clear that Kant must hold that a substance can act on itself. The point here is simply that, apart from the causal efficacy of any other substance, a substance cannot act on itself either so as to *change* itself (as the principle of succession argued) or so as to determine a *relational* property (as the principle of coexistence implicitly argued). Since simultaneity is a relational property, Kant can be understood as claiming merely that a substance cannot act on itself so as to determine its simultaneity with another substance.

If this line of justification reflects Kant's reason for the crucial step in the argument of the Third Analogy, then one might naturally ask how it relates to the principle of inertia. For the ultimate premise of Kant's argument – that a substance cannot act on itself causally so as to determine its own place in time – bears some resemblance to Newton's law of inertia, according to which a body must remain in its state of (rectilinear) motion or rest, unless acted on by another. In particular, one might suspect that it is the law of inertia that leads Kant to accept the principle that a substance cannot cause its own place in time as simply a metaphysical generalization of a physical principle. Moreover, one might draw support for this idea by noting that Kant restricts the scope of the claim of the Third Analogy to substances that coexist *in space*.

While the fact that Kant restricts the scope of his conclusion to spatial substances is, *prima facie*, peculiar insofar as the arguments he presents do not explicitly invoke space in any of their premises, it is clear that the law of inertia cannot be used as a justification for his crucial premise in the "main" argument, because he provides an independent proof of the law of inertia in the Mechanics of the *Metaphysical Foundations of Natural Science* (4:543). If the principle of inertia can provide no independent

support for the crucial premise of his main argument, it would seem to be likely that his support for that premise must stem from the conception of grounds that he has developed in the pre-Critical period. This is not to say that he accepts everything about that conception, and, in fact, we have reason to return to his Critical evaluation of his pre-Critical account of grounds in Chapter 4, but it does provide indispensable help in ascertaining Kant's argument in the Third Analogy.

The Meaning of "Thoroughgoing" Mutual Interaction

The conclusion of Kant's argument is that knowledge of the coexistence of substances requires that they stand in thoroughgoing mutual interaction. But what exactly does it mean for substances to be in *thoroughgoing* mutual interaction? We have already given a brief characterization of what mutual interaction is and return to this in more detail in Chapter 4, but the term "thoroughgoing" has remained unspecified. Moreover, since we have restricted our discussion to only two substances, it leaves open the question of whether considering more complex scenarios changes how mutual interaction must be understood.

Consider two different interpretations of what "thoroughgoing" might mean if applied to a larger number of substances. A "strong" interpretation of this term would suggest that each substance acts on and is acted on by every other simultaneously existing substance immediately or directly. The model for such an interpretation might be Newton's universal attraction, since universal attraction implies that any given substance with mass (and located in a common space) acts on and is acted on by every other substance with mass from any point in space, regardless how large or small and how far apart or close together they may be.

A "weak" interpretation of "mutual interaction" would claim that each substance acts on (and is acted on by) every other simultaneous substance either immediately (as in the strong interpretation) *or* mediately. Kant may be using perception as the model for this latter interpretation. The example Kant uses in his explanation of the Third Analogy is that for me to perceive celestial bodies ("*Weltkörper*"), there must be intermediary light which establishes "a mediate community" (A213/B260) between myself and these bodies. Thus, although I do not interact immediately with these bodies, I do interact immediately with the light, and the light in turn interacts immediately with these bodies such that I can be said to be in interaction with these bodies mediately.

Does Kant adhere to the weak or the strong interpretation? The strong interpretation may appear to be supported by Kant's *Metaphysical Foundations of Natural Science*, since that work presents arguments for universal attraction.⁵⁶ But this fact alone does not immediately imply the strong interpretation of the Third Analogy, since, as we saw above, Kant clearly distinguishes between the transcendental principles of the *Critique* and the *Metaphysical Foundations's* principles. However, if Kant really does want to restrict the Third Analogy to spatial substances and if he establishes in the *Metaphysical Foundations* that every spatial substance is immediately related to every other substance via the mutual interaction of gravity, then Kant would in fact hold the strong interpretation, whether he intends to be arguing for it in the *Critique* or not. At the same time, the weak interpretation allows for both mediate and immediate mutual interaction, so that the weak interpretation is obviously not *inconsistent* with the *Metaphysical Foundations*. Thus, it is quite possible that Kant could be both holding the weak interpretation here, while elsewhere attempting to demonstrate universal attraction.

Two other considerations favor the weak interpretation. First, the text of the Third Analogy tends to favor the weak interpretation. At the beginning of A213/B259, Kant mentions, albeit in parentheses, that dynamical community can be "immediate or mediate," and below on that same page he claims, again in parentheses, that mediate community must obtain for objects far away. Second, the argument of the Third Analogy would seem *prima facie* to establish only the weak interpretation, since simultaneity is a transitive relationship. If one knows that substances A and B are simultaneous because they interact and similarly that B and C are simultaneous because they interact, then one would appear to be justified in inferring that A and C are simultaneous without it being the case that A and C interact directly (or in any way other than by means of their interaction with B). Such a claim regarding transitivity can be supported, however, only after we have determined the exact nature of mutual interaction. There are models of mutual interaction that would not allow for the transitivity of simultaneity. For example, if A interacts with B at t_1 and C interacts with B at t_2 , it does not immediately follow that A and C are simultaneous. Thus, if the weak interpretation is to be correct, Kant's model of mutual interaction must preclude the possibility of such gaps, that is, must be transitive. We return to this issue in Chapter 4.

⁵⁶ See 4:508–509.

CONCLUSION

In this chapter we have seen that Kant's Second and Third Analogies of Experience attempt to argue that causality and mutual interaction (metaphysical relations) are necessary conditions for knowledge of objective succession and coexistence (epistemological items), which are in turn required for several different unities: the unity of nature/world, the unity of time, the unity of experience, and the unity of apperception. We thus now have confirmation of the speculative thesis advanced at the beginning of this chapter that the nature of Kant's "Critical turn" ought to consist in a mixture of continuities and breaks compared with his pre-Critical views. The discontinuities are most apparent (1) in the addition of a series of epistemological concepts (e.g., the unities of apperception and experience) to the argumentative framework Kant employs (such that his primary concern is with, e.g., phenomenal rather than noumenal substances) and (2) in the fact that causality and mutual interaction are alleged to be necessary for *knowledge* of succession and coexistence. It is thus in line with his metaphilosophical rejection of mere analysis as a legitimate means for providing knowledge of synthetic a priori truths that Kant employs experience (of succession and coexistence) in establishing metaphysical principles (of causation and mutual interaction).

However, what has not yet been noted in any detailed way by commentators and what is thus of special significance for our understanding of Kant's views are the widespread continuities. The most obvious one pertains to the basic content of Kant's claims in the *Nova dilucidatio's* principles of succession and coexistence and in the *Critique's* Second and Third Analogies, namely that causal relations are required for the temporal relations of succession and coexistence. Another continuity, much less apparent to those who have focused almost exclusively on the Second Analogy, is Kant's assumption that a substance cannot act on itself so as to change itself or to determine its place in time, an assumption Kant makes most clearly in the *Nova dilucidatio's* principle of succession. The continuities do not, however, end with the general features of Kant's claims or with the most fundamental assumptions of his arguments. To see that they are more extensive still, we now turn to Kant's ontological commitments in the Analogies of Experience by investigating his model of causality.

Kant's Model of Causality

INTRODUCTION

Most commentators on Kant's views on causality have presupposed, whether explicitly or implicitly, that Kant adopts Hume's model of causality, according to which one determinate event (e.g., the motion of one billiard ball at one moment in time) causes another determinate event (e.g., the motion of a second billiard ball at a later moment in time). If pressed, they could cite several reasons in support of such a presupposition. First, if Kant were to employ a model that displayed significant differences from Hume's, how could he possibly hope to refute Hume's position without begging the question? Second, in light of the widespread success of Newtonian physics at the time, it might have seemed that science would require nothing more than causal laws that describe which events follow which other events. Third, this presupposition is apparently backed up by textual evidence from the *Prolegomena*, where Kant famously asserts (4:260) that it was Hume who first awoke him from his dogmatic slumber and, for that matter, precisely on the issue of causality.

If Kant were committed to causality being a relation between events, then the main relevant difference between Kant's and Hume's models of causality would be that Kant accepts, whereas Hume rejects, the idea that these relations are necessary and universal. Accordingly, commentators have focused their attention first and foremost on whether or not a justification can be found for accepting Kant's more ambitious claim. Moreover, their focus could easily appear to be perfectly consistent with the apparent intent of Kant's argument in the Second Analogy, since, if successful, it establishes that the temporal determination of events, which

Hume mistakenly took to be unproblematic, actually presupposes causal rules of the sort Hume wanted to deny.

We have reason to return to the more general question of how Kant should be seen as replying to Hume in Chapter 6. The present chapter prepares the way, however, by arguing, in a first section, that Kant's model of causality is *not* Hume's, that is, does not consist of one event causing another, whether events are understood in simple or more complex ways. The case against event-event causation is made by showing that an event-event model is inconsistent with Kant's claim in the Third Analogy that mutual interaction is necessary for knowledge of coexistence. For it is simply incoherent (in the context of Kant's philosophy) to assert that two events could stand in mutual interaction with each other.

If Kant rejects Hume's (and Humean) event-event model(s) of causality, what model does he accept? The primary aim of the second section of this chapter is to state the fundamental features of Kant's model of causality. The decisive clue to understanding Kant's model comes from appreciating how he draws on several basic aspects of his pre-Critical notion of a ground. For just as was the case for the pre-Critical Kant, the Critical Kant thinks that a substance can cause a change of determinate state in another only insofar as it contains a *ground* that *determines* the successive states of the other substance. In the case of mutual interaction, this means that the grounds or powers of two substances will jointly determine each other's states in such a way that these states can be understood as simultaneous (just as mutual interaction was held to be responsible for mutual changes in the *Nova dilucidatio's* principle of succession). Thus, rather than thinking of causality in terms of events, Kant can be seen as a proponent of a causal powers model of causality. In the Critical period, Kant further refines this notion of a causal power or ground such that it is consistent with the possibility of simultaneous causation and the principle of continuity. Yet perhaps the most distinctive feature of Kant's notion of ground is its inherent asymmetry. A ground that determines the state of another substance does not do so by virtue of a determinate state of its own or by means of its mere existence, but rather through an indeterminate activity that is incapable of ever becoming determinate itself.

Kant expresses this conception of activity in terms of the notion of "the causality of the cause," thereby distinguishing it not merely from Humean event-event models but also Leibniz's version of a causal powers model. The task of investigating and clarifying this notion is undertaken in the third section of this chapter. It is true that this notion fits

perfectly with Kant's view that several of the most fundamental properties of bodies (e.g., filling a determinate space and communicating motion) are to be explained in terms of the *exercise* of attractive and repulsive forces. However, since the exercise of such forces might seem to be unintelligible to some, the notion of activity invoked in Kant's *metaphysical* account of causality cannot be explained in terms of its *physical* instantiations. Nor can one truly be said to be *adding* any clarification to it by taking recourse to "architectonic" resources within the *Critique* (such as the schema or category of causality). Rather, this notion can ultimately be rendered intelligible by way of analogy with Kant's distinctive account of self-consciousness, where he develops a subtle response to Hume's insightful critique of the Cartesian position.

If Kant's model of causality is much more robust than event-event models in virtue of its inclusion of substances endowed with causal powers that actively bring about determinate states, we must consider, in the final section of this chapter, what implications such a model might have for aspects of Kant's position that have already been discussed in Chapter 3. In particular, it is natural to revisit the argument of the Analogies, since a better understanding of Kant's model of causality allows us to see two ways in which our understanding of these arguments can be improved. First, this model helps to explain why Kant restricts the conclusion of the Third Analogy to *spatial* substances. Second, it provides indispensable help on the question of causal laws, which was left without a definitive answer in Chapter 3. On the basis of the explicit argument of the Analogies of Experience, it was unclear as to why Kant might think that he would be justified in asserting the existence of causal laws rather than "rules" that apply to just a single case. However, Kant's notion of *unchanging* grounds can be seen as providing resources that explain why he would think that appealing to causal laws in this stronger sense is justified. Finally, once we are in possession of a detailed understanding of both Kant's arguments for causality and his model of causality, we can consider in less tentative ways how the "Critical turn" is to be understood, that is, how Kant can reconcile his account of causality with Transcendental Idealism.

EVENTS AND EVENT-BASED MODELS OF CAUSALITY

Humean and Kantian Events

Before comparing Kant's model of causality with Hume's, it is helpful to understand what events are for Hume and how they figure into his

views on causality.¹ Hume begins *A Treatise of Human Nature* by distinguishing between simple and complex ideas in order to state precisely his basic principle that every simple idea must be copied from a more vivacious simple impression. This basic principle is in turn necessary for Hume's express intent in the *Treatise*, which is "to introduce the experimental method into moral subjects." For, in the context of book I, Hume needs to determine what he can legitimately rely on in providing an account of how we could know the world; it is in this context that Hume singles out simple sense impressions.² Accordingly, simple impressions are a primitive epistemological element in Hume's project, and his ontological primitives, namely events, are defined in terms of them. Although Hume does sometimes talk about objects (e.g., in his definitions of causality in book I, part III, section XIV), it is apparent from his account of the continued and distinct existence of bodies (in book I, part IV, section II) that they are ultimately derivative from the content of simple impressions.³

What are simple impressions for Hume? In book I, part I, section I, of the *Treatise*, Hume defines simple impressions as those that "admit of no distinction or separation."⁴ One kind of distinction and separation he discusses at length in sections I and II of book I, part II, is spatial and temporal divisibility. There, he notes that although many different perceptions can be united in a variety of ways, only one kind of unity is truly real, namely that which "must be perfectly indivisible, and incapable of being resolved into any lesser unity."⁵ In applying this reasoning to time, he argues that "every moment must be distinct from, and posterior or antecedent to another. 'Tis certain then, that time, as it exists, must be compos'd of indivisible moments."⁶ In other words, for Hume, simple impressions as well as the primitive events that are to be defined by means of them cannot be further divided spatially or temporally. Accordingly, for Hume events are instantaneous states of affairs at particular moments in time.

¹ By the time of the *Critique*, it is likely that Kant had access to a much wider range of Hume's corpus, including his *Dialogues on Natural Religion*. We attend to the issue of Kant's reply to Hume in detail in Chapter 6.

² David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch (Oxford: Clarendon, 1978), p. xi.

³ *Ibid.*, pp. 170–172, 187–218.

⁴ *Ibid.*, p. 2.

⁵ *Ibid.*, p. 31.

⁶ *Ibid.*

But how does Hume's understanding of events as instantaneous states of affairs relate to his account of causality? After all, one might initially think that Hume's argument for the distinctness of cause and effect would be purely epistemological in character and rely simply on the fact that due to the coarseness of our senses we do not happen to have the ability to perceive necessary connections or "secret" powers in nature. We are built in such a way that we are able to perceive colors, motions, sizes, shapes, and so on but not necessary connections. Just as some species of animals are color-blind, we happen to be "power-blind." Accordingly, even if cause and effect were somehow necessarily connected in nature, we would still perceive them as distinct, since we do not possess the ability to have an impression of a necessary connection.

While such an epistemological argument would obviously carry significant weight for Hume, it is also clear that he would think that his account of events would provide a second, even more powerful argument for the radical distinctness of cause and effect. For Hume's position on the infinite divisibility of space and time in book I, part II, of the *Treatise* commits him to an understanding of events that entails the *necessary* distinctness of cause and effect.⁷ Since simple impressions and the events that are defined by means of them are necessarily spatially and temporally indivisible, they cannot endure, that is, span any measurable length of time. Thus, for Hume we experience the world by means of completely discrete, instantaneous mental snapshots. In support of this position, he argues that we get our idea of duration from our idea of succession, and our idea of succession is in turn derived not from an impression of a change of state of an object, but rather from noticing or feeling the succession of distinct impressions in the mind.⁸ In short, according to Hume we get our idea of succession not from an impression of succession (as one might expect in light of his principle that ideas are copies of impressions), but rather from a succession of impressions. Given that an event is a state of affairs at an instantaneous moment in time, along with Hume's requirement (stated in section XIV of book I, part III, of the *Treatise*) that the cause precede the effect, the cause and the effect must exist at different times. As events that we can perceive, cause and effect are, therefore, not contingently, but rather necessarily

⁷ The argument from the coarseness of our senses establishes only the contingent distinctness of cause and effect, whereas this argument would establish their necessary distinctness.

⁸ David Hume, *A Treatise of Human Nature*, pp. 35–37.

distinct.⁹ As a result, in attempting to construct the causally connected world out of such instantaneous events, Hume discovers that there could be no objectively necessary connections that might serve as “the cement of the universe” and thus that he must use constant conjunctions and subjective feelings or expectations instead.¹⁰

While Hume's account of events and the role that they play in his views on causality would obviously have been both important and interesting for Kant, our discussion in Chapters 1 and 2 should serve to remind us that the historical context for Kant's discussions of what model of causality could be appropriate for his own needs extends beyond Hume's empiricist account. In particular, as we saw in Chapter 1, even Leibniz's and Wolff's most radical critic, Crusius, agreed with them that causality would have to be explained in terms of substances and their causal powers. Moreover, insofar as Kant felt the need to develop *metaphysical* principles for Newtonian science by, among other things, providing an ontology that would be intrinsically intelligible and at the same time contribute in some way to the justification of such scientific laws, it is clear that simply invoking the events described by those laws would neither increase their intelligibility nor provide any argumentative support for them, and that whatever providing an intelligible account might amount to, it would require an appeal to resources more robust than the events invoked by such laws. Finally, as we saw in Chapter 2, the pre-Critical Kant invoked substances and mutual interaction as the most fundamental (material and formal) principles of any world, whether it be sensible or intelligible. It would thus be surprising if Kant were to attempt to develop a theory of causality in the Critical period that rejected his pre-Critical account entirely and invoked only events. In light of this fuller context, our expectation should rather be that it would be most natural for Kant to

⁹ §7 of the *Enquiry Concerning Human Understanding* directly supports such a claim. For helpful discussions of Hume's views on causality, see Barry Stroud, *Hume* (London: Routledge & Kegan Paul, 1977); Robert Fogelin, *Hume's Skepticism in the Treatise of Human Nature* (London: Routledge & Kegan Paul, 1985); and, more recently, Fred Wilson, *Hume's Defence of Causal Inference* (Toronto: Toronto University Press, 1997); Don Garrett, *Cognition and Commitment in Hume's Philosophy* (New York: Oxford University Press, 1997); and Graciela de Pierris, “Hume's Pyrrhonian Skepticism and the Belief in Causal Laws,” *Journal of the History of Philosophy* 39 (2001): 351–383.

¹⁰ Hume could view certain processes, such as the playing of Beethoven's *Eroica*, as an event, but they must be composed of instantaneous events. As a result, Humean events cannot overlap partially. One could develop a non-Humean theory of events, but then the contrast between Kant's and Hume's positions becomes more difficult to state in a clear way.

appeal to a model of causality that invoked substances, causal powers, and mutual interaction.¹¹ Obviously, such an ontology is much less minimalistic than Hume's events are, but it would be in keeping with the tradition in which Kant was educated and for whose figures he wrote to invoke such an ontology.

Faced with this fuller context, it would seem more promising to suggest (perhaps as a fallback position) that even if Kant does accept substances – those very substances that Hume forcefully criticized in book I, part I, section VI, of the *Treatise* – one could simply graft Humean events onto Kantian substances without all too significant modification. For example, one might suggest that a Humean event is, for Kant, simply the state of a substance at a particular time and that causal powers are reducible to determinate states of substances (and regularity relations between them). However, even before we consider the details that would bear on such an attempt, such a strategy cannot appear promising. In the Second Analogy, Kant states that an event (most often, a “*Begebenheit*” but, occasionally, an “*Ereignis*”) is not a state of a substance at a particular time, but rather a *change* from one temporally determinate state to another. Moreover, as we saw in Chapter 3, what Kant typically (though not exclusively) focuses on is not succession *between* the cause and the effect, as it was for Hume, but rather succession *within* the effect.

Finally, and perhaps most importantly, Kant's interest lies not in constructing a world out of instantaneous states, but rather in explaining how it is that we can know the various temporally determinate states of an object that would belong to a single spatio-temporal world. Thus, it is clear from the start that Kant's account of events does not map neatly onto Hume's insofar as for Kant an event is a change of state the temporal determination of which requires explanation, whereas for Hume change is impossible within an event, since change within an event would entail the divisibility of something that is, as we have seen, essentially indivisible. Because these and further comparisons and contrasts between Hume and Kant deserve much more attention, we return to a detailed discussion of them in Chapter 6. For now, it is enough if we simply remind ourselves of the fuller historical context to which Kant is reacting, since it is directly

¹¹ There are several Hume scholars who think that Hume, too, invokes causal powers. See John Wright, *The Sceptical Realism of David Hume* (Minneapolis: University of Minnesota Press, 1983), and Galen Strawson, *The Secret Connexion* (Oxford: Clarendon, 1989), for positive arguments for this view and Ken Winkler, “The New Hume,” *Philosophical Review* 100 (1991): 541–579, for a rebuttal.

relevant to understanding the terms in which Kant's model of causality is expressed.

Event-Event Models of Causality

In light of this understanding of the historical context of Kant's reflections on causality, we can now explicitly consider Kant's model of causality. In this context, it is striking how little the argument of the Second Analogy of Experience entails. As we saw in Chapter 3, the Second Analogy establishes that a cause is required to bring about the succession of states in a substance as its effect, but its argument does not explicitly require that a cause be an event, that it be prior to rather than simultaneous with its effect, or that it have any specific characteristics at all. Its central task is simply to determine the states of a substance as successive. Accordingly, taken in isolation, it rules out neither occasionalism nor pre-established harmony, since it is possible that the cause could be either God or internal to the substance whose state is changing. Thus, Kant's model of causality must be filled out on the basis of other considerations.

The argument of the Third Analogy, by contrast, is much more informative and is hence a crucial resource for developing a detailed description of Kant's model of causality.¹² In virtue of its focus on a plurality of substances it cannot leave unspecified the nature of the cause and therefore commits Kant to principles that rule out an event-event model of causality. As we saw in Chapter 3, if the argument of the Third Analogy is successful, it establishes that substances must stand in mutual interaction for knowledge of the simultaneity of their states to be possible. We also saw that mutual interaction must be understood as a two-way causal relation, where each causal relation holds between a substance and the states of a distinct substance, because (1) causality can determine a substance's place in time only by causing its *states* at those times rather than by causing either temporal determinacy per se (i.e., independently of its states) or the existence of the substances, and (2) a substance cannot act on itself so as to determine its own place in time.

Our current focus is on the precise nature of the relations that make up the two-way causal relation Kant calls mutual interaction. How exactly does one substance determine the state of another? In particular, does such determination require that an event occur in the substance that is

¹² The two other main resources are nonargumentative passages in the Second Analogy and Kant's pre-Critical views on causality.

the cause, as the event-event model would hold? As we saw in Chapter 1, Meier explicitly employs this idea as a premise in one of his arguments for pre-established harmony, so that it is safe to assume that Kant is familiar with such a model not just in its Humean form, but also in the context of Leibnizian metaphysics. If Kant does not accept this requirement, then by what means does Kant think that one substance can determine the state of another? What must a substance be like to be able to determine the state of another substance?

Simple Event-Event Models

To begin to determine the basic features of Kant's model of causality, take a simple event-event model of causality and apply the Third Analogy's notion of mutual interaction to it. As one can immediately see, attempting to understand mutual interaction in terms of a simple event-event model of causality generates an explicit contradiction. If causes are events and mutual interaction is a two-way causal relation, then mutual interaction would be a two-way causal relation between two events. Accordingly, one event would cause a second event, which would, in turn, cause the first event, which is obviously a contradiction. It is this problem that motivates Schopenhauer to "banish the concept of mutual interaction from metaphysics."¹³ To bring out some of the difficulties that arise in attempting to understand mutual interaction in terms of events, consider his formulation of the problem.

Only insofar as state A precedes state B in time, but their succession is necessary, not contingent . . . only to that extent is state A the cause and state B the effect. The concept of mutual interaction contends, however, that both are the cause and both the effect of the other: but this means the same as that each one is both the earlier and the later event: which is absurd [*ein Ungedanke*]. For that both *states* are simultaneous, and necessarily so, cannot be accepted: because as necessarily correlated and simultaneous, they constitute only *one* state.¹⁴

One of Schopenhauer's objections in this passage is that if the cause must be prior to its effect, and mutual interaction is reciprocal causation between events, then a single event would have to be both prior and subsequent to another event, which is clearly a contradiction. Since this particular contradiction arises due to the impossibility of one event being

¹³ Arthur Schopenhauer, *Die Welt als Wille und Vorstellung*, in *Sämtliche Werke*, vol. 2 (Wiesbaden: Eberhard Brockhaus Verlag, 1972), vol. 1, p. 544.

¹⁴ *Ibid.*, p. 545.

both *prior* and *subsequent* to another event, one could avoid it by simply rejecting the temporal asymmetry of causality.¹⁵

However, a second contradiction immediately takes its place, one that does not involve the temporal priority of the cause over the effect and is thus problematic for any simple event-event model of mutual interaction. If mutual interaction consists in two events causing each other, then the problem arises that if the first event causes the second event, then the second event cannot in turn cause the first event, as mutual interaction would claim, since the second event depends on the first one for its very existence and is thus unable to be that on which the first one depends for its existence. In other words, it is a contradiction to claim that one event is both the cause and the effect of another at one and the same time, because causality entails that the existence of the effect depends on the existence of the cause and it is impossible for one event to depend on a second event at the same time that the second event depends on it (in the same respect or with respect to existence). It is important to notice, however, that this contradiction (involving reciprocal existential dependencies) arises for *any* kind of simple event-event model of mutual interaction, that is, whether events are understood as instantaneous states of affairs at particular moments in time, as Hume would have it, or as changes of state, as Kant believes. As a result, Kant's main claim in the Third Analogy is incompatible not only with the model of causality Hume actually holds, but also with any simple event-event model of causality.

Complex Event-Event Models

What generates both of these contradictions for the simple event-event model of causality is the fact that in mutual interaction one event is supposed to be both the cause and the effect of another. To maintain mutual

¹⁵ Margaret Morrison, "Space, Time and Reciprocity," in *Proceedings of the Eighth International Kant Congress* (Milwaukee: Marquette University Press, 1995), vol. 2, pp. 187–195, pursues this strategy as follows: "If A causes B then B must be later than A and if B also causes A then A must be later than B. It is impossible for A to be both later and earlier than B, therefore they must be simultaneous." Now one might think that what is at issue here is neither momentary states nor changes of state – in which cases the contradiction does arise that the one momentary state or change therein would have to be both earlier and later than the other momentary state or change therein – but rather enduring states. There is textual support for this interpretation in the Third Analogy, where Kant seems to suggest that the simultaneous states endure from A to E, during which time they can be apprehended in any order, and philosophical support for it since it allows one to avoid this objection insofar as enduring states can be both before and after other enduring states. However, as we see below, this interpretation faces another, more serious difficulty.

interaction while also avoiding the contradictions just encountered by event-based models of causality, one would have to develop a more complex model of causality by splitting up each of the events that had served as the causal relata into distinct entities so that one and the same event does not perform both functions at the same time. As a result, for mutual interaction to be a coherent possibility, it is necessary that what one might call the “causal aspect” of a substance not be identical to the “effect aspect” of that same substance, where the causal aspect of the substance is that part of the substance by means of which one of the two relations constituting mutual interaction brings about its effect, and the effect aspect of that substance is the effect brought about by the other of the two causal relations constituting mutual interaction.

If we draw a distinction between the cause and effect aspects of a substance in this way, then one might attempt to construe Kant’s model in terms of events as follows. Event e_1 that occurs in one substance causes event e_2 in another substance, while this second substance, by means of event e_3 that occurs in it, causes event e_4 in the first substance. Because event e_1 brings about event e_2 and event e_3 brings about event e_4 , no event is both the cause and the effect of another and the contradictions faced by simple event-event models do not arise.

Does this more complex event-event model of causality represent a coherent interpretation of the Third Analogy’s notion of mutual interaction? There are two different basic versions of such a model. First, if one accepts the idea that a cause must precede its effect, no contradiction arises and mutual interaction at least represents a metaphysical possibility. For example, this model might be understood such that we would have event e_1 at t_1 causing event e_2 at t_2 , and then event e_3 at t_3 , in turn, causing event e_4 at t_4 . While such a version is apparently coherent on its own, it is unacceptable in the context of an interpretation of the Third Analogy because mutual interaction does not, in that case, enable knowledge of coexistence. We have knowledge of the first substance only at t_1 and t_4 and knowledge of the second substance only at t_2 and t_3 , and thus no knowledge of the states of both substances at the same time.¹⁶

¹⁶ C. D. Broad, *Kant: An Introduction* (New York: Cambridge University Press, 1978), would appear to hold a version of such an interpretation. Consider his following statement: “[T]o say that two substances A and B are in mutual interaction would seem to have the following meaning. Every alteration a_1 in A causally necessitates a *later* alteration in b_1 in B; this in turn causally necessitates a *later* alteration a_2 in A; this causally necessitates a *later* alteration in B; and so on” (p. 178, emphasis added). Broad finds textual support in Kant’s distinction between mediate and immediate community and the example used to

To ensure that knowledge of simultaneity is established, one might propose a second version of this model as follows. Event e_1 at t_1 causes event e_2 at t_2 , while event e_3 at t_1 causes event e_4 at t_2 , with events e_1 and e_4 occurring in a first substance and events e_2 and e_3 occurring in a second substance. This modified model is still distinct from the simple event-event models since it distinguishes causal aspects of substances (events e_1 and e_3) from their effect aspects (events e_2 and e_4), and no contradiction arises in virtue of any reciprocal existential dependency. Further, unlike the original version of this model, because events e_1 and e_3 are both at t_1 and events e_2 and e_4 are both at t_2 , the causal ties guarantee that we would be able to know the simultaneity of the two substances.

There is, however, a fatal difficulty with this version of this model. For using temporal indices, such as t_1 and t_2 , smuggles in coexistence illegitimately. To illustrate this difficulty more clearly, consider the same model using the terms “before” and “after” in place of t_1 and t_2 . Such a replacement is warranted on Kant's account of causality, since Kant's concern never extends beyond establishing the minimal notion of temporal *order* that is involved in succession (and not its measurable lapse, A203/B248). Now, on the version of mutual interaction just proposed, the first causal tie does *not* determine that event e_1 at t_1 causes event e_2 at t_2 , but rather that event e_1 is *before* event e_2 . Similarly, the second causal tie determines that event e_3 occurs before event e_4 . But stated in this manner, coexistence has not been established between any of the events, and the previous version's difficulty reappears. First, it has not been shown that events e_2 and e_4 coexist, but rather only that they occur after events e_1 and e_3 , respectively. Second, it has not been shown that events e_2 and e_4 occur an equal temporal distance after events e_1 and e_3 . It is entirely possible on this model that the one later event (e_2) occurs long after its causally related initial event (e_1), whereas the other later event (e_4) occurs just a split second after its causally related initial event (e_3) so that one cannot infer the simultaneity of the later events (e_2 and e_4) from their occurring at an equal temporal distance after the initial events (e_1 and e_3). Third, even if one could determine equal temporal distances between both sets of initial and later events, one could not infer the simultaneity of the later events from this fact, since this inference requires the simultaneity of the initial events, which has also not been shown. Therefore, one initial

illustrate it, namely that of the light mediating between celestial bodies and our eyes. In my view, this example shows simply that mediating entities may be required for distant objects to interact mutually.

event, which is determined to be prior to one later event, is not necessarily simultaneous with the other initial event, which is determined to be prior to the other later event.

Faced with another dead end, one might think that the source of the problem lies in introducing temporal asymmetry into the causal relations. Accordingly, if one rejects the idea that a cause must be prior to its effect (as Kant does at A203/B248), then one might think that the temporal disparity between the events could be avoided in such a way that mutual interaction could still be necessary for the simultaneity of the two substances without entailing any contradiction. The most promising complex event-event model now is as follows. Event e_1 , in the first substance, at t_1 causes event e_2 , in the second substance, at t_1 , which in turn causes event e_3 , in the first substance, also at t_1 . Since (1) events e_1 , e_2 , and e_3 all obtain at t_1 , (2) the simultaneity of the two substances can be known, and (3) there are two causal relations going in opposite directions, it would be natural to describe them as an instance of mutual interaction that also allows for knowledge of the coexistence of substances.

However, this model faces two new objections. First, if this model were possible, it would not so much support the Third Analogy, as rather be a devastating objection to it. For if it were possible to assert that event e_1 at t_1 causes event e_2 at t_1 , then there would be no need to assert that event e_2 causes event e_3 at t_1 to establish the simultaneity of the two substances. That is, there would be no need to assert *mutual* interaction between the two substances, since we would already know, on the basis of the first causal relation, that the two substances coexist at t_1 .¹⁷

Second, and even more seriously, this model presupposes a feature of the causal relations making up mutual interaction that is prohibited by the argument of the Third Analogy. In particular, it presupposes that a substance can determine its own place in time by assuming that event e_1 occurs in the first substance at t_1 . As we saw when considering the argument of the Third Analogy in Chapter 3, a substance cannot determine its own place in time and therefore requires the causal efficacy of a substance distinct from it, which ultimately generates the need for mutual interaction (since each substance requires its place in time to be determined by another). Because the first substance cannot determine

¹⁷ Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987), criticizes Kant's argument in the Third Analogy as being unable to establish *mutual* interaction, since a simple causal relation of the sort described above would suffice for knowledge of coexistence (pp. 272–273).

its own place in time, it cannot determine that event e_1 , which causes event e_2 at t_1 , occurs at t_1 . Yet nothing else could determine event e_1 at t_1 either, since the defining feature of the complex event-event model was its separation of cause aspects from effect aspects. Accordingly, complex event-event models of causality fail in the context of the argument of the Third Analogy, just as simple event-event models did.

GROUNDS, CAUSAL POWERS, AND DETERMINATIONS

Indeterminate Grounds, Joint Determination, and Causal Powers

If Kant's model of causality cannot be explained solely in terms of events, what other options are open to him? In particular, since Kant explicitly identifies the effect with an event (as a change of state), the decisive question must be: What does Kant think that a cause is? The First Analogy of Experience might seem to provide an immediate and simple answer to this question, since it contains an argument for phenomenal substances, and it might seem obvious that a cause must be a substance: What causes the motion of the second billiard ball? The first billiard ball, which is simply a spatial substance. While there is a nontrivial sense in which such an answer is correct, it is important to recognize that it can be only a small part of Kant's full answer. For this position is subject to a series of pressing philosophical questions, questions that one cannot answer by appealing simply to the notion of a substance as such: How can the *mere existence* of a thing bring about an effect? How is one to understand that such a cause would bring about its effect *at any one time* rather than at any other, if it is supposed to be the substance rather than its state at any given time that is the cause? How does the mere existence of *one* thing explain an effect in *another*? Is it not the *state* of the thing at a particular time (as opposed to the *thing* per se) that could explain the effect?¹⁸

But note that the argument of the previous section has ruled out that the cause could be what these questions might suggest, namely a determinate state of the substance. In terms of our concrete example, if these questions show that the cause of the motion of the second ball cannot simply be the first billiard ball per se, Kant's argument in the Third Analogy

¹⁸ These questions are simply reformulated versions of the kinds of questions that proponents of event causation (such as Donald Davidson and C. D. Broad) pose to advocates of agent causation (such as Reid, Taylor, and Chisholm). It is clear that Kant is aware of such questions as early as 1755 in his *Nova Dilucidatio*, where he explicitly distinguishes between the existence of a substance and its causal relations with other substances.

shows that it cannot be the motion of the first billiard ball, either. But if the cause cannot be simply a substance (the first billiard ball) nor a determinate event in it (its motion), one faces, once again, the question of what it can be. It is helpful to note here that Kant sometimes uses the phrase “the causality of the cause” and, on several occasions, explicitly distinguishes between the cause (i.e., the substance) and “the causality of the cause.”¹⁹ This suggests that it is Kant’s notion of the causality of the cause that is crucial to understanding his model of causality. But how is “the causality of the cause” to be understood? To answer this question, let us first attend to Kant’s notion of a ground and then consider how his model of causality can be explained in terms of such a notion.

The basic idea of a model of causality for which grounds are central is that one substance determines the successive states of another by means of an unchanging ground that is part of its essential nature. Since a ground both acts in accordance with essential features of the substance and is a source of change (insofar as it determines the successive states that constitute change), it cannot itself change from one determinate state to another (because that would entail an infinite regress). As a result, a ground is not temporally determinate in the way in which the effect is, since the effect, unlike the cause, has one determinate state at one moment in time at its beginning and another such state at its end.²⁰

One can find three distinct lines of support for understanding Kant’s model of causality in terms of grounds. First, understanding Kant’s model of causality in terms of grounds allows one to avoid the problems that the various event-based models faced in explaining how mutual interaction is to be possible. Second, the structure of Kant’s explanation of motion in terms of grounds in his pre-Critical period is analogous in fundamental respects to what is needed to account for knowledge of simultaneity discussed in the Third Analogy. Finally, on closer inspection, one can find unambiguous textual evidence in the Second Analogy that Kant adopts precisely this notion of ground. In addition, if one links grounds with

¹⁹ See, for example, his *Metaphysics Mrongovius* lectures: “Causality is the determination of a cause by which it becomes a cause, or the determination of the relation of a thing as cause to a determined effect. Thus the cause is always to be distinguished from the causality” (29:893), and the L2 lectures, where he notes: “When the cause has been posited, the effect is posited <posita causa ponitur effectus> already flows from the above. But when the cause has been cancelled, the effect is cancelled <sublata causa tollitur effectus> is just as certain; when the effect has been cancelled, the cause is cancelled <sublato effectu tollitur causa> is not certain, but rather the causality of the cause is cancelled <tollitur causalitas causae>” (28:573).

²⁰ The reason why grounds are not temporally determinate is discussed below.

causal powers, one can see further textual support for the claim that Kant's model of causality involves grounds and also understand why Kant would not have thought it necessary to emphasize his divergence from (Humean) event-event models.

By considering how grounds differ from Humean events, one can see how understanding Kant's model of causality in terms of them can avoid the problems that the various event-based models faced. Kant's grounds are distinct from Humean events in several respects. First, unchanging grounds endure throughout the change of states that they cause, whereas Hume's instantaneous events pop into and out of existence. Second, since (real) grounds determine changing states, they can be responsible for some kind of necessity between it and its effects and they are thus not distinct from their effects in the same way that events are from each other. Finally, whereas events are temporally determinate (since they occur at a determinate instant in time), grounds are temporally indeterminate, given that they do not change from one determinate state to another.²¹

Now recall the various difficulties that event-based models encountered in attempting to account for simultaneity by means of mutual interaction. First, for simple event-event models, the cause and the effect depended on each other ontologically. Second, for complex event-based models (which distinguish between cause and effect aspects of a substance), the cause and effect aspects of the one substance could diverge temporally from, and thus not be simultaneous with, the cause and effect aspects of the other substance. The cause of this potential lack of simultaneity was that the two causal relations that were to constitute mutual interaction could obtain independently of each other. The independence of these two causal relations was implied, in turn, by the fact that the cause-effect relationship used to construct mutual interaction invoked nothing more than events, and events, as Hume argued, are necessarily distinct from each other. Third, because of this second difficulty, it appeared that event-based models of causality could be used to explain simultaneity only if a substance could determine its own place in time (given that the second difficulty eliminated the possibility that other substances could do so), but this principle contradicts a fundamental assumption of the argument of the Third Analogy (since if a substance could determine its own place in time, then *mutual* interaction would not be necessary).

²¹ The exact senses in which grounds are temporally indeterminate are developed in note 40.

If Kant's model of causality is based on grounds rather than events, it can avoid the problems encountered by the different versions of event-event models as follows. First, since the ground of one substance that determines the (successive or changing) states of another substance does not, in turn, depend on those successive states, this model is not committed to reciprocal existential dependencies of the sort that plagued simple event-event models. Second, if the way that the grounds of one substance determine the successive states of another is not independent of the way that the grounds of the second substance determine the successive states of the first substance, then it is possible that these grounds jointly determine their respective states. Since simultaneity is simply a particular instance of the joint determination of the states of substances, a model of causality based on grounds is not immediately barred from providing an explanation of simultaneity. Finally, since grounds are temporally indeterminate, there is no need to assume the temporal determinacy of the cause in the first place and one is thus not in danger of violating any fundamental assumptions of Kant's argument in the Third Analogy. Nor does such an understanding of grounds either make mutual interaction unnecessary to account for simultaneity or lead to an infinite regress of grounds.

Understanding Kant's model of causality in terms of grounds also allows one to make use of parallels it has to his pre-Critical account of causality. As we saw in our discussion of the *Nova dilucidatio* in Chapter 2, Kant held that mutual interaction is required for mutual changes of substances, and the case of motion nicely illustrated why he might have thought that to be the case. Suppose one body is changing its motion with respect to another. According to Kant's principle of determining ground, there must be a ground that causes this change (or the successive determinations that constitute it). Yet because a substance's grounds act according to its unchanging essence, the grounds in the first body cannot cause a change in its own determinations. As a result, grounds in the second body must be the cause of the change in the first body. However, since motion is a reciprocal relational property, (a change in) the motion of the first body toward the second necessarily implies (a change in) the reciprocal motion of the second body toward the first. By reason of parity, if the grounds in the first body cannot cause its own (change of) motion, then the grounds in the second body cannot be the cause of its own (change of) motion, either. As a result, grounds in the first body must be the cause of (the change of) motion of the second body, just as grounds in the second body must be the cause of (the change of) motion

of the first body. That is, the mutual change of state of two substances is possible only if they stand in mutual interaction (i.e., if each one grounds the motion of the other).

However, describing the situation in this way could be misleading, for it might suggest that there are two completely unrelated events, the motion of the first body and the motion of the second, which require two independent causes, the ground in the second body and the ground in the first body. More specifically, if this description were complete, one could raise the objection that this model does not guarantee simultaneity, since it fails to connect the grounds, and such a failure, as we saw above, produces the potential lack of simultaneity of the states determined by those grounds. To avoid this misleading impression, what must be emphasized about this example is that neither the motions of the two bodies nor their grounds are independent. First, the motion of the one body toward the second both logically implies and is logically implied by the motion of the second body toward the first. Second, since a substance cannot contain the grounds of its own changes, one must attribute the ground of the change that each substance undergoes to the other substance. Yet due to the reciprocal relations between the effects, there must be reciprocal relations between each of the grounds so that they are not independent of each other. In short, there are two grounds that *jointly* determine the reciprocal motion of the two bodies.²²

To understand how grounds can jointly determine the states of substances, it is important to distinguish between grounds *simpliciter* and the way in which such grounds *actively* determine the states of a substance *in a specific situation*. Grounds *simpliciter* involve essential features of the substances that contain them and are therefore as independent from each other as self-subsistent substances are. However, they are not essential features *insofar* as they *actively determine* the states of another substance. As we saw in Chapter 2, as early as 1762 Kant understood that (causal) relations between substances cannot be logically necessary, since that is incompatible with the independence that defines substances as such. Accordingly, there must be a distinction between grounds that are necessary for a substance to exist at all and the particular way in which they bring about effects or determinations in other substances. As we saw in

²² This model of mutual interaction also explains how mutual interaction can be transitive, just as the weak interpretation of the Third Analogy requires. Since mutual interaction requires that both substances' states exist at the same time, no gap can arise between either of these substances (or their states) and any third substance that stands in direct mutual interaction with one of these substances.

Kant's pre-Critical account, for grounds to *determine* the states of other substances (as opposed to simply constituting the subsistence of their own substances), they must be placed in some position, situation, or context that allows them to bring about specific determinations.²³ To revert to the case of bodies, whether two bodies are placed two or three feet away from each other makes a difference not in the essential grounds of the substances so placed, but rather in the determinations that they will bring about in each other (e.g., in how strongly they attract each other). But given that one and the same substance could be placed in different relations to different substances, there are different ways in which its grounds can be effective. That is, depending on how two substances are situated with respect to each other, their grounds can determine different states in them.

This description of how Kant's pre-Critical account of causality in the *Nova dilucidatio* invokes mutual interaction to account for motion allows us to see how Kant understands mutual interaction in his Critical period such that it is necessary for knowledge of simultaneity. First, it is clear that simultaneity is a reciprocal relational property, just as motion is. If the state of one substance is simultaneous with the state of a second substance, then it is impossible that that state of the second substance not be simultaneous with the state of the first. Second, if changes in a substance require the causal efficacy of a distinct substance on the grounds that a substance cannot act on itself so as to change itself, then it is analogous to claim that since a substance cannot determine its own place in time, one must attribute the source of such determination to another substance. Third, just as one must appeal to both the masses of the bodies and their relations (e.g., distance) to determine what motions they will cause, so too one must consider how both the grounds and the relations between the substances that contain the grounds determine whether two states are simultaneous or not. Fourth, since motion, as a reciprocal relational property, must be determined by two separate but jointly determining grounds in both bodies due to the fact that no single substance can determine motion by itself, it follows, by reason of parity, that two grounds must jointly determine the simultaneous states of these substances as well. Finally, if one must draw a distinction between grounds *simpliciter* and grounds insofar as they actively determine the states of the substances to which they are related, it becomes more understandable

²³ Kant's way of putting this would be to say that substances must stand in some relation to each other if they are to belong to one and the same world.

what Kant means when he distinguishes between a substance per se and that aspect of a substance by means of which it can be a cause, that is, between the mere existence of a substance and the "causality of the cause" (i.e., the causality of the substance that is the cause). As a result, the analogies between Kant's explanation of the motion of bodies in his pre-Critical period and the basic elements of his Critical model of causality help us to see, at least in rough outline, what these fundamental features must be like.

One might, however, still be concerned that this new model of mutual interaction is not ultimately able to avoid the problems faced by the event-event models. For if it is the case that Kant's notion of mutual interaction requires that the grounds of two substances jointly determine their simultaneous states, it might seem to follow that these grounds must be simultaneous just as their states are, and if that is right, then two serious questions, or rather objections, arise for this model. First, if the grounds must be simultaneous, would Kant's argument not beg the question? That is, would this model not render the inference to simultaneity trivial by assuming simultaneity in the first place and thus simply push the crucial issue of objective time-determination back one stage, since the question immediately arises as to how we know that mutual interaction is occurring? Second, if the grounds must be simultaneous, does it not follow that they would therefore also be temporally determinate and does that not entail that this model, despite the best of intentions, would encounter the very same difficulties that the simple and complex event-event models did?

Regarding the first question, it is important to recall what was argued in Chapter 3, namely that the Analogies are concerned with the metaphysical presuppositions of our knowledge of the temporal relations of objects, not exclusively with its epistemological presuppositions. As a result, insofar as the Analogies are concerned with establishing the *necessity* of substance, causality, and mutual interaction, Kant is not primarily concerned with how we can know empirically when each of these things is present. However, even if this is correct as an interpretation of Kant's intentions in the Analogies, the question itself is not obviously illegitimate and is thus deserving of an answer despite the fact that Kant does not explicitly develop it as such in this context.

Kant's most promising resources for developing an answer to this question about the notion of mutual interaction stem, I would suggest, from the *Metaphysical Foundations of Natural Science*. In the *Mechanics*, for example, Kant argues that any instance of the communication of motion

(either in impact or in gravitational attraction between two bodies) requires that the moved bodies stand in mutual interaction insofar as their action and reaction on each other must be equal. Accordingly, if we can know that two bodies communicate motion to each other, then we can know that they stand in mutual interaction, and if we know that they stand in mutual interaction, then we know that their states (e.g., at impact) are simultaneous. Because knowledge that two bodies communicate motion to each other does not presuppose knowledge that any determinate states of these bodies coexist, presupposing knowledge of the communication of motion does not immediately beg the question in attempting to show that their states coexist.²⁴

Even if the first question can be answered in this way, the second question might seem all the more pressing, since it could be strengthened by drawing on one crucial line of argument from the *Nova dilucidatio*. There Kant had attempted to show that (1) in light of the principle of determining ground, changing determinations require changing grounds and (2) since grounds cannot change, change within an isolated substance is impossible. Given this line of argument, one might think that just as changing determinations require changing grounds, so, too, simultaneous determinations would require simultaneous grounds, and if grounds are simultaneous, they must also be temporally determinate.

But Kant ultimately rejects the inference from the determinacy of (simultaneous) states to the determinacy of (simultaneous) grounds in the Critical period. If the grounds of simultaneous states were simultaneous and thus temporally determinate, then, given that the problem of time-determination would apply to the grounds of states as much as it applies to the states themselves, further grounds would have to determine them. But if grounds were to require further grounds in order to be determinate, an infinite regress would arise, since the former could be determinate only if the latter were, and they, in turn, could be determinate only if there were yet a further set of grounds, and so on. As a result, Kant must reject the idea that grounds are temporally determinate.

Moreover, Kant's rejection of the temporal determinacy of grounds in the Critical period fits in with his adoption of Transcendental Idealism. According to Transcendental Idealism, one must distinguish between things in themselves, which are completely determinate (but atemporal),

²⁴ For an interesting discussion of this issue and its contemporary ramifications, see Martin Carrier, "How to Tell Causes from Effects: Kant's Causal Theory of Time and Modern Approaches," *Studies in History and Philosophy of Science* 34 (2003): 59–71.

and appearances, which are not completely determinate, but rather become determined through application of our understanding's categories to objects given through spatio-temporal forms of intuition. As a result, unlike things in themselves, appearances do not already exist with their spatio-temporal properties from the start. Instead, any determinate properties they have must arise through a process of determination. Accordingly, even if the grounds of appearances are themselves appearances, there is no reason to think that they are temporally determinate, unless one can cite some ground in virtue of which such temporal determination has been established. Transcendental Idealism thus makes room for Kant to assert that grounds need not be temporally determinate.

Finally, lest one think that the Critical Kant entirely rejects (as dogmatic) his pre-Critical notion of an unchanging ground, unambiguous textual evidence supports the view that Kant's position should be understood in this way. For Kant explicitly endorses several of the central features of this notion in a relatively neglected passage from the Second Analogy that argues: "according to the principle of causality actions are always the primary *ground* of all change of appearances, and *therefore cannot lie in a subject that itself changes*, since otherwise *further actions and another subject*, which determines this change, would be required" (A205/B250, emphasis added). That is, Kant argues that every change requires a ground, and that the ground of a change cannot itself change on pain of infinite regress. Moreover, it is neither the mere existence of a substance nor any determinate event in it, but rather *action* or *activity* that serves as the *ground* of the changes that objects undergo. The parallels between this passage and the *Nova dilucidatio* are striking. Just as Kant holds in the *Nova dilucidatio* that invoking changing grounds to account for changing determinations leads to an infinite regress in the explanation of change and that change must rather be explained on the basis of changing relations between immutable grounds, he explicitly points out in the *Critique* that the ground of change must itself be an unchanging activity.

But notice where reconstructing Kant's model of mutual interaction in such a way that it can fulfill the various conditions placed on it by the Analogies has led us. The crucial notion is that of an unchanging ground that brings about determinate change in other objects. That notion is nothing other than the traditional notion of a causal power or force (*Kraft*), as understood in its most *basic* form by, for example, Aristotle, Leibniz, and Locke. While the notion of causal powers thus plays a crucial role in Kant's understanding of causality, it must be admitted that he does

not exactly highlight the notion in the first *Critique*. However, two of his remarks in the Second Analogy explain why he would have thought there to be no need to do so. First, “where there is action, consequently activity and force [*Kraft*], there is also substance” (A204/B250). Second,

[t]his causality leads to the concept of action, this to the concept of force [*Kraft*], and thereby to the concept of substance. Since I will not crowd my critical project . . . with analyses that address merely the elucidation (not amplification) of concepts, I leave the detailed discussion of these concepts to a future system of pure reason – especially since one can already find such an analysis in rich measure even in the familiar textbooks of this sort. (A204/B249)

We return below to a discussion of Kant’s rich and detailed comments on such “familiar” analyses, since they reveal quite distinctive features of his understanding of causal powers. For now, it is important merely to see that in the Second Analogy Kant *takes for granted* that we can appeal to the notion of a *Kraft*, that is, force or causal power, and that his intended audience would already be familiar with the broadest outlines of this kind of model of causality from standard metaphysics textbooks. Because virtually all philosophers in the modern period accepted causal powers, Kant would feel no immediate need even to provide a basic analysis of the notion, especially since more pressing issues are on his agenda in the *Critique*.

Simultaneous Causation and the Principle of Continuity

Before we turn to Kant’s distinctive conception of the traditional notion of causal powers as described in passages outside the *Critique*, it is important to discuss two passages from the Second Analogy that do not form part of Kant’s explicit argument for the necessity of causality to our knowledge of succession but that nevertheless supply important insights about the framework within which Kant is thinking about causality. One addresses the so-called problem of simultaneous causation, while the other concerns the principle of continuity.

Kant explicitly begins his discussion of simultaneous causation (at A202–203/B247–248) by noting that there is a reservation (*Bedenklichkeit*) about his treatment of causality up to that point that must be resolved.²⁵ The reservation stems from the fact that the principle of the

²⁵ The German text reads: “Hier äußert sich aber noch eine Bedenklichkeit, die gehoben werden muß” (A202/B247), which Guyer and Wood render as “Here, however, there is a reservation that must be raised.” I am inclined to read the passage slightly differently.

Second Analogy is designed to account for instances of succession, when the very same kind of causal principle that applies to such cases also applies to cases in which the cause and the effect are simultaneous. To illustrate this point, Kant describes a room heated by a stove as a case to which the principle of causation applies despite the fact that there is no succession, that is, the cause and the effect are simultaneous. He then explains how this is possible as follows:

The majority of efficient causes in nature are simultaneous with their effects, and the temporal sequence of the latter is occasioned only by the fact that the cause cannot achieve its entire effect in one instant. But in the instant in which the effect first arises, it is always simultaneous with the causality of its cause, since if the cause had ceased to be an instant before, then the effect would never have arisen. Here one must note that it is the order of time and not its lapse that is taken account of; the relation remains even if no time has elapsed. The time between the causality of the cause and its immediate effect can be vanishing (they can therefore be simultaneous), but the relation of the one to the other still remains determinable according to time.²⁶ If I consider a ball that lies on a stuffed pillow and makes a dent in it as a cause, it is simultaneous with its effect. Yet I still distinguish the two by means of the temporal relation of the dynamical connection. The temporal sequence is accordingly the only empirical criterion of the effect in relation to the causality of the cause that precedes it. (A203/B248)

Kant's explanation here, the "fuss about vanishing times" in particular, has appeared convoluted to many.²⁷ However, against the background of the basic model of causality sketched above, there is a way of understanding this passage that renders it fairly straightforward and intelligible. One can immediately note that this passage reinforces the argument of the previous section, as it makes clear that Kant cannot accept Hume's model of causality. For Hume (at least in the *Treatise*), it is an analytic truth that a

I take "äußert" to correspond to the idea that there is a reservation or objection of some sort, and "gehoben" as an indication that the objection must be responded to rather than "raised," since taking it to mean "raised" makes "äußert" redundant. So, the point is not that one must raise an objection, but rather that there is an objection that must be answered.

²⁶ Guyer and Wood translate this part of sentence as: "the temporal relation of the one to the other still remains determinable." I prefer the following: "the relation of the one to the other still remains determinable according to time." The German text reads as follows: "das Verhältnis der einen zur anderen bleibt doch immer, der Zeit nach, bestimmbar." Grammatically, it could be read as Guyer and Wood do, but if one holds that the cause is temporally indeterminate and must always be such insofar as grounds can never become determinate, then the temporal relation is not determinable in that sense and an alternative translation is required. Fortunately, "der Zeit nach" can be read not as modifying the relation, but rather as the means by which the relation is determined.

²⁷ Guyer, *Kant and the Claims to Knowledge*, p. 262.

cause must precede its effect in time. If Kant agreed with Hume about this principle, then the problem of simultaneous causation would be more than a reservation or a problem; it would be a blatant contradiction. But it is clear that Kant does not think that the *Bedenklichkeit* he is attempting to resolve in this passage amounts to a contradiction. Rather, it is a question of how a causal principle that is designed to apply to succession could also apply to cases where there is no succession.

In this context, one can see Kant addressing two questions. First, if a cause must be simultaneous with its effect, does that not *rule out* succession? For if the effect exists at the same time as its cause, it might seem that no new determination could arise. This question might seem to be especially pressing since, as we saw above, Kant clearly states (at A205/B250) that the cause, as the ground of change, cannot itself change. Kant's explicit, though also quite concise, answer is that a cause cannot necessarily bring about its *entire* effect at the first moment of its activity. The case of bodies in motion can illustrate this idea, which, as we saw in Chapter 2, Kant had expressed as early as the *True Estimation*. If a body has a certain amount of motive force, it may not be able to achieve its entire effect in an instant, as, for example, an arrow in motion toward a target does not lose all of its motion immediately, but rather does so only (somewhat) gradually as it moves through various resisting media of the target and becomes lodged some way into it. Yet as we saw in our discussion of the *Nova dilucidatio*, the idea can be explained in purely metaphysical terms as well. As a substance changes its relations to the substances on which it acts, so, too, its effects will change until equilibrium has been reached when the unchanging grounds of these substances relate to each other in such a way that they do not posit any new determinations.

Second, if cause and effect are simultaneous and the cause can immediately produce its effect in such a way that there is no succession in the effect, how can one distinguish the cause and the effect? Kant's answer to this question is straightforward, even if the terminology he uses might mislead. Philosophically, his point is simply that even if a cause and its effect are simultaneous, an effect depends on its cause for its existence, whereas a cause does not depend on its effect for its existence. Kant invokes the potentially confusing idea of vanishing times merely to show that there is no metaphysically significant difference between cases of causality where it *is* possible to distinguish between the effect and its cause on the basis of succession in the effect and those cases where it is *not*, that is, where the succession of states in the effect diminishes to

zero in a state of equilibrium and hence vanishes such that no change occurs in the effect. In other words, whether there is succession in the effect or whether it has vanished, the dependence relationship of the effect on the cause remains. As Kant points out in conclusion, however, there is still an epistemological difference between these two cases. Since causal dependence is not directly perceived, succession can serve as an empirical criterion by which one can determine in a particular case what is the cause and what the effect.

The second passage in the Second Analogy that provides helpful information about the framework Kant is using in thinking about causality concerns the principle or law of continuity, sometimes referred to as the principle of no leap. Kant repeatedly and explicitly endorses this principle in his Critical period: in the first *Critique*, in the *Metaphysical Foundations of Natural Science*, and throughout his metaphysics lectures.²⁸ Though Kant distinguishes different versions of the principle of continuity (especially in the metaphysics transcripts), the version that is relevant for our purposes concerns alteration, and states simply that all alteration is continuous. Kant's preferred way of illustrating the principle is with the example of the motion of a body: "[W]hen one body transfers from one point to another, then it must go through infinitely many intermediate spaces, it must go through all intermediate locations lying between the one point in the line and the other" (28:201). In several passages (e.g., 29:864), however, Kant makes it clear that the principle is relevant to any kind of alteration, not just the motion of a body.

In the Second Analogy Kant explains this version of the principle of continuity as follows:

The question therefore arises, how a thing passes from one state = *a* into another one = *b*. Between two instants there is always a time, and between two states in those instants there is always a difference that has a magnitude. . . . Thus every transition from one state into another happens in a time that is contained between two instants, of which the former determines the state from which the thing proceeds and the latter the state at which it arrives. Both are therefore boundaries of the time of an alteration, consequently of the intermediate state between two states, and as such they belong to the whole alteration. Now every alteration has a

²⁸ See A208–209/B253–254, A228–9/B281, 28:41, 28:199–205, 28:662–663, 29:862, 29:1006, and 4:552–553. While one might initially think that the principle of continuity is a dogmatic remainder from Kant's pre-Critical period that he ought to have left out of his Critical works, he refers (approvingly) to the principle of continuity so often that one ought to reconsider how such a "dogmatic" sounding principle might in fact be compatible with Kant's Critical turn. See my "Kant on Rational Cosmology" in *Kant and the Sciences*, ed. E. Watkins (New York: Oxford University Press, 2001), pp. 70–89.

cause, which manifests its causality in the entire time during which the alteration proceeds. Thus, this cause does not produce its alteration suddenly (all at once or in an instant), but rather in a time, so that as the time increases from the initial instant a to its completion in b , the magnitude of the reality ($b-a$) is also generated through all the smaller degrees that are contained between the first and the last. All alteration is therefore possible only through a continuous action of causality, which, insofar as it is uniform, is called a moment. The alteration does not consist of these moments, but it is generated through them as their effect. (A208–209/B253–254)

This rich passage brings out a number of important points about Kant's model of causality with unusual clarity. First, Kant thinks of an alteration as a transition from one determinate boundary state to another, with an infinite number of states in between the boundary states. Second, the transition from one determinate boundary state to another occurs not in an instant, but over time. More specifically, the alteration from one state to another occurs not in discrete "jumps," but rather in a continuous flow, and as a result only the boundary states can be described in temporally determinate terms. Third, Kant describes the effect as being brought about by means of a "continuous action of causality." In other words, the continuous change of states is brought about by a continuous activity of the cause. Finally, alteration does not consist of the continuous action of causality, but is rather generated through such action.²⁹

These points suggest that we specify Kant's model of causality more precisely as follows. When a cause brings about its effect, it acts uniformly, thereby generating a continuous flow of states in an object from one determinate boundary state to another. Further, it is crucial to note that, in this picture, only the boundary states of the effect are temporally determinate, because (1) all that is specified about the intermediate states is that the object must pass through each and every one of them but not when, that is, at what determinate time each one will do so, and (2) the causality of the cause, despite its activity, is uniform, that is, does not

²⁹ This passage confirms that Kant is not thinking of the causality of the cause as being noumenal. For this passage asserts that the causality of the cause can be continuous and uniform, which clearly implies temporality. In addition, as we saw above, Kant talks about the cause not being able to bring about its effect all at once. However, because Kant seems to accept (at A203/B248) the principle that a cause must exist in order to bring about its effect and the effect occupies a determinate span of time, the causal activity must continue to exist even after it has started to bring about its effect in order to bring about the remaining part of its effect. If it is appropriate to say (as Kant does) that the causality of the cause *continues*, then it must be temporal and thus cannot be noumenal.

change from one determinate state to another. Moreover, since temporal determinacy comes about only as the result of the causal activity (i.e., as a result of its act of determining the boundary states), the “causality of the cause” or “the continuous action of causality” is not itself determinate. For reasons that will be made clearer shortly, nor could it ever become determinate through the causal activity of a distinct cause.³⁰ For the only kind of entity that can be determinate is a *state* of an object that lies between determinate boundaries; by contrast, the causality of the cause is not and cannot be a determinate state, but rather must be a continuously efficacious *activity*. The ontology Kant invokes in the context of filling out his model of causality is thus very different from Hume's ontology of distinct, instantaneous events.

Causal Powers, Asymmetry, and Activity

Kant's distinctive understanding of causal powers is articulated further in transcripts from his lectures on metaphysics from 1782–1783, called the *Metaphysics Mrongovius* lectures. While these lecture transcripts often contain mere explications of Baumgarten's position, it is not unusual for them also to contain critical remarks that Kant is directing at Baumgarten's views from the perspective of his own Critical position. In the present context, they represent further clear textual evidence that Kant is committed to a causal powers model of causality, which thereby complements the philosophical and textual argument presented above against event-event models of causality. More importantly, however, these transcripts contain materials that lay out Kant's distinctive understanding of causal powers insofar as they illustrate that for Kant (1) whenever they are exercised, causal powers form an irreducible and asymmetrical relationship between the substances that have them, and the determinations

³⁰ It is important to be especially clear about this point. One can of course determine some state of the object that was the cause, but this determination in no way implies that one has thereby shown that *that state* was the cause of the effect in question. Sometimes Kant does speak loosely about determining the cause, but that is consistent with the *causality of the cause* not being determinable in the same way. Often, when Kant speaks about determining cause-effect relationships, his real intent is to show that determinism is true for determinate phenomenal events. To this end, Kant has to be committed to saying not only that a cause is responsible for the change from one state to another, but also that there is another cause that is responsible for the change of state that brings about the initial state of the effect from an even earlier state (which is, in turn, brought about by yet another cause, etc.).

that they produce and (2) the asymmetrical aspect of this relationship must be understood in terms of an active-passive dichotomy.

Kant's first task concerning causality in the *Metaphysics Mrongovius* lecture transcripts is to defend the idea, familiar in one respect from his pre-Critical thought, that causality is based on a *real* or ontological ground-consequence relationship as opposed to Baumgarten's purely *logical* ground-consequence relationship. Kant's generic definition of grounds and consequences is uncontroversial, since it asserts that they are entities conjoined by the relation of connection (rather than that of opposition). If the one is posited, then the other is necessarily posited as well, and vice versa. If the connection is analytic (i.e., according to the principle of identity), then the ground-consequence relationship is logical and a priori, whereas if it is synthetic (i.e., if reason cannot comprehend the connection), then it is real and a posteriori. Because the generic definition of ground and consequence entails that they are necessarily coexistent and completely symmetrical, Kant recognizes that one "cannot distinguish ground and consequence by [this generic] definition" (29:808). To capture the crucial asymmetry that holds between ground and consequence (namely that the consequence is posited *only because of* the ground and not vice versa), he refines his definition by means of a notion of determinacy as follows:

The ground is that by which, having been posited, another thing is posited determinately, the consequence is that which is not posited unless something else is posited . . . for if there is a consequence, there must likewise always be a ground, and if something is a ground, there must likewise always be a consequence, but in the first case it is indeterminate, in the other determinate. (29:808)

Accordingly, the ground determines the consequence according to a rule, but the consequence does not determine the ground in precisely the same way, since, in principle, a given consequence could follow from any one of a number of different grounds, that is, the rule according to which it follows is not determined. Yet since the cause-effect relationship is described as an instance of the real ground-consequence relationship, which is asymmetrical precisely in virtue of the consequence being determinate and the ground being indeterminate, the effect must likewise be determinate and the cause indeterminate. This view is precisely what was required of Kant's model of causality by his argument in the Analogies.

Does Kant's discussion of this asymmetry between the indeterminacy of the cause and the determinacy of the effect add to his notion of a causal power and the model of causality explicated in terms of it so far?

Later in the ontology section of the Metaphysics Mrongovius lectures, Kant discusses the relational categories of substance-accidence (subject), cause-effect (principle), and active-passive (interaction). In describing the way in which accidents inhere in substances, Kant describes accidents as positive determinations of substances. His primary concern is to show that negative and logical predicates are not accidents in the same sense in which positive predicates are and that a substance does not carry its accidents in the same way that a bookcase supports its books. Yet he also clarifies the relationship between substance, accident, and determination by claiming that “insofar as a thing is determined positively, accidents inhere in it” (29:770). In short, accidents are simply positive determinations. Moreover, a substance cannot itself be a determination. For Kant holds that a determination must be a determination of something and a substance cannot be a determination of something if it is to be defined as “that which exists without being the determination of another” (29:770). So far, the picture is clear enough. The effect is determinate insofar as the cause brings about a change of determinate states, or accidents, in a substance. The cause, insofar as it is to be identified with substance, must not be a determination, that is, a determinate state, because a substance, by definition, cannot be the determination of another.

But how do causal powers fit into this ontological framework? Kant continues in the Mrongovius lectures by relating his discussion of substances, accidents, and determinations to causality as follows: “With a substance we can have two relations: in relation to accidents it has power insofar as it is the ground of their inherence; and in relation to the first subject without any accidents, that is the substantial. Power is thus not a new accident, but rather the accidents are effects produced by the power” (29:770). Here Kant clearly asserts that a power is not an accident or determination of a substance.

If Kant has thus stated clearly that a power is not an accident, then the temptation is to identify it with substance, as Baumgarten does. Again, Kant's discussion of Baumgarten's position in the Metaphysics Mrongovius lectures is instructive. While Kant agrees with Baumgarten that a (causal) power (*Kraft*) is “that which contains the ground of the inherence of the accidents” (29:771), he rejects Baumgarten's claim that power is identical to substance. “Since accidents inhere in each substance, he [Baumgarten] concludes that every substance is power. That is contrary to all rules of usage: I do not say that substance is a power, but rather that it has power, power is the relation of the substance to the accidents, insofar as it contains the ground of their actuality” (29:771). In this passage

Kant clearly denies that a power is itself a substance.³¹ But, as we saw above, it could not be an accident either and it must thus be in between (and irreducible to) substance and accident. Moreover, Kant is clearly aware of this implication and explicitly embraces it. For he continues as follows: “We thus have something [power] that is not substance, yet also not accident” (29:771). Accordingly, Kant clearly rejects the assumption that everything is either a substance or an accident and asserts that a power is in between its substance and the determinate states it produces.³²

Does Kant have any motivation that extends beyond linguistic usage for understanding causal powers in this way? The first point to note is that one is committed to ontological entities that are neither substances nor accidents as soon as one accepts into one’s ontology substances, accidents, and an inherence relationship between the two. This commitment can be illustrated by considering how one might attempt to reduce the inherence relationship to accidents and substances. Is the inherence relationship an accident of the substance or a substance? If it is an accident, then it obtains only in virtue of the fact that it itself inheres in the substance. But of course this second inherence relationship would be an accident too and would obtain only if it stood in a third inherence relationship with the substance, *ad infinitum*. At the same time, it is equally problematic to claim that the inherence relationship is itself a substance, since if it were, then one would need an inherence relationship between it and the accident that is supposed to inhere in the first substance, but this inherence relationship, which would be a substance as well, would require yet another inherence relationship, and so on. In this case, too, an infinite regress seems unavoidable. It seems that the only way to avoid an infinite regress lies in accepting inherence as something indeterminate “in between” substances and their accidents.³³ That Kant understands this

³¹ In the Second Analogy (A204/B250) Kant asserts that activity is a criterion of substance, but such an assertion does not obviously entail that they are identical.

³² This kind of issue arises in discussions about realism and nominalism in the medieval period by Scotus, Ockham, and Suarez, since they, too, were concerned with how to understand the exact status of the relations between fundamental ontological entities, such as the union of (substantial) form and (prime) matter and the inherence of an accident in a substance (which could, in principle, be distinct from each other really, conceptually, formally, or modally).

³³ The same point holds true of objects, properties, and the exemplification relationship that holds between objects and their properties. Is exemplification a relational property? If so, it would entail an infinite regress of exemplification relationships, which would always be just further properties of which it would be unclear whether or how they related to the object itself. At the same time, it is difficult to see how a thing could literally be an exemplification.

point is evident from the fact that he explicitly includes the substance-accident-inherence relationship as a primitive relational concept of the understanding in his table of categories.

But notice that the substance-accident inherence relationship is structurally analogous to the relationship between a cause and its effect as expressed by “the causality of the cause” or by the way in which a causal power grounds new determinations in another substance. For according to Kant, both substances and causes ground their consequences, namely the inherence of accidents. The primary difference for Kant is that a substance is an *inner* sufficient ground of its own accidents, whereas a cause is an *outer* sufficient ground of the accidents that are its effect.³⁴ That is, a substance is an inner ground of the inherence of its own accidents, while a substance is a cause insofar as it is an outer ground, that is, a ground for the inherence of accidents in another substance. Accordingly, both substances and causes are real grounds and the main difference concerns whether the consequence, namely the inherence of accidents in a substance, is internal or external to the substance that contains the ground. Yet this difference does not immediately affect the status of grounding. If grounding is in between a substance and its accidents in the one case, it will be in between a substance and its accidents in the other as well. Again, Kant seems to be quite aware of this point when he declares the cause-effect relationship to be a category, that is, a primitive concept.

Another reason Kant would seem to have for not identifying power with substance, as Baumgarten does, is that it would force one to take what Kant thinks of as an overly restrictive view of faculties. In particular, as we saw in Chapter 1, Crusius thought that human beings are endowed with understanding and will as faculties that are not reducible either to each other or to any more primitive power. Though Kant seems to follow Crusius in thinking of the understanding and the will as distinct faculties, he is quite clear that reason (as active) and sensibility (as passive) are necessarily distinct (i.e., irreducible) faculties, for he famously criticizes both empiricists such as Locke and rationalists such as Leibniz for attempting to reduce the one to the other (at A271/B327). If faculties are simply epistemic powers, then the identification of power with substance would commit Kant to viewing human beings as composites of several substances. Though Kant may want to retain a kind of agnosticism about what we ultimately are (at the noumenal level), it is clear that he would not want to commit himself to the view that knowledge can

³⁴ See 28:51–53 for Kant's most explicit discussion of this point in the Herder transcripts.

occur only by means of interaction within two distinct substances. As a result, Kant would have significant epistemological reasons for resisting the identification of powers with substances, at least in the case of human beings.³⁵

If a causal power must thus be something indeterminate in between a substance and the accidents that the substance brings about qua cause, one might still ask why a causal power could not become determinate through some further act. The crucial point here lies in Kant's understanding of the notions of activity and passivity involved in causality. Recall that for Kant a cause is not only constantly conjoined or even necessarily connected with its consequent effect, but also *brings about* or *produces* its effect by *actively determining* the boundary states of an object. The object so determined thus does not exist fully formed from the start with this determination, but rather passively receives it.³⁶ Now, if one were to try to determine the cause (by some further causal connection), the result would, by parity of reasoning, be a passive determination of the substance that is the cause, not the activity essential to the causality of the cause. In other words, one could determine in this way a state of the cause, but not what is at issue, namely the activity by which the cause brings about its effect.³⁷

³⁵ As we see below, Kant thinks that in at least some cases substances should be thought of as containing only one basic power. There are several (compatible) ways of resolving this tension. First, one can distinguish between different kinds of substances, for example, those that are endowed with mental powers (i.e., faculties) and those that are not, and then hold that only certain kinds of substances can contain no more than one power. Second, one can maintain that even the claim that a substance can *have* only one power does not entail the *identification* of power with substance. Finally, one might think that the passages indicating Kant's commitment to the possibility of only one basic power may simply express his belief that, for methodological reasons, the identity conditions of the substances that underlie matter do not diverge from those of powers and that there would be no way to distinguish substances that would not also distinguish their powers.

³⁶ It is true that Kant holds that both finite substances are active in any causal interaction. For example, in his *Metaphysics* Mrongovius lectures, Kant states: "We can never be merely passive, but rather every passion is at the same time action" (29:823). But the fact that both substances are active in the production of an accident does not detract from the point that all accidents or determinations are themselves passive, since Kant clearly distinguishes the action by which an accident is brought about and the accident itself.

³⁷ In principle, this point was already present in the Third Analogy. For what would make the circularity of mutual interaction vicious in the case where the causality exercised by each substance is understood to be a determinate state is the fact that each determinate state is supposed to produce the other. Without the determinacy and thus without the activity and correlative passivity, there would be nothing problematic about two states being related to each other.

The crux of Kant's argument for the claim that the causality of the cause cannot become determinate is thus his idea that determinations are necessarily passive states that are brought about by activities of causal powers.³⁸ In this respect, Kant's notion of a determination is quite unlike contemporary notions of determinate states of affairs or properties, which would not normally be characterized as passive and which would thus not require a corresponding activity. What is Kant's justification for thinking of determinations as passive states that require activity of a causal power? The most prominent justification for thinking of determinations in this way is found in Kant's Analogies of Experience. As we saw in Chapter 3, what drives the argument of the Analogies is the problem of time-determination, that is, the presupposition that the determinate temporal properties of objects are not immediately given along with the subjective order of our representations, but rather result only from a process of determination involving the relational categories.³⁹ More specifically, the point is not that the objects already have determinate temporal properties and we simply must use the categories to determine the intuitions that are given to us in sensibility in order to *discover* these properties. Rather it is that the objects do not already have determinate temporal properties.⁴⁰ The objects must be temporally determined by something

³⁸ While it is standard to refer to substances as active and passive, there is a derivative sense in which one can speak of the active and passive "aspects" of substances in each case. The active "aspect" of a substance is the exercise of its causal power, while the passive "aspect" of a substance is the state (or determination) for which the exercise of causal powers is responsible.

³⁹ Kant's discovery of the problem of time-determination in this form (in the context of the development of Transcendental Idealism) is what distinguishes his Critical conception of a ground (according to which it is temporally indeterminate) from his pre-Critical conception (according to which it is temporally determinate).

⁴⁰ It may be helpful to clarify the senses in which the activity involved in the causality of the cause is and is not indeterminate. Such an activity is not indeterminate in the sense of being random or arbitrary. Nor is it for that reason atemporal, since, as we saw above, this activity is uniform and unchanging and thus unrestricted in its duration. Rather, it is temporally indeterminate in the sense that there is no intrinsic state of the substance that is responsible for the effect *and* to which one can assign a specific temporal index, since qua activity it cannot be a state that has been determined, and hence made possible, by an activity in the first place. (While Kant's position would stand in stark contrast to Hume's even without understanding causal activity as temporally indeterminate in this sense, it does seem that he must in fact be committed to this further claim.)

One might attempt to distinguish between the *temporal* indeterminacy of the activity whereby a substance causes a change of state and the indeterminacy in the *kind* of cause it is and then object that granting the latter does not entail acceptance of the former, since uncertainty about what kind of cause a substance is does not entail that the substance does not exercise its causal powers (whatever they may be) at a determinate

else, something active (e.g., a determining ground), to have temporally determinate properties.⁴¹ All three Analogies posit something active, namely a ground, to bring about or determine the temporal features of objects, whether the activity takes the form of an inner sufficient ground for substance or an outer sufficient ground in the case of causality and mutual interaction.

Further evidence that Kant understands causality in this way can be found in his discussion of the task of natural science. Kant asserts that “all natural philosophy occupies itself with the reduction of powers to a single basic power, which we cannot explain, namely that because something is, something else thereby follows. All basic powers must be given through experience” (29:772). Later in the same paragraph, Kant continues: “In natural science one has good reason to regard the attracting and repelling powers as primitive powers. Can there be in one substance many or only one basic power? For our reason there must be several because we cannot reduce everything to one, but the unity of each substance requires that there be only one basic power” (29:773–822).⁴² Though we return to this point below, Kant seems to think that in nature what we perceive are the effects of powers rather than the powers themselves. For if one could perceive powers directly, one would not need to attempt to reduce powers to a single basic power. As a result, we must infer what kinds of powers there are from the kinds of observable effects they produce. But since we seem

time. However, the inference in this case is not from our ignorance about the nature of a (fully determinate) cause to the indeterminacy of the temporal index of the cause, but rather from Kant’s view that appearances have no determinate features independently of our knowledge of them and from his claim that temporal determinacy is derivative on causal determination (i.e., the claim that one cannot attribute a specific temporal index to a substance except by attaching it to a state of the substance, which can occur only if there is a cause that brings that state about at a particular time).

⁴¹ See 29:818–819 for a clear statement of this point. Now one might think that it is not at all clear how the notion of a temporally indeterminate ground can be the cause of a temporally determinate change of state in an object, perhaps because the notion of a temporally indeterminate entity is unclear or because it is unclear how something temporally indeterminate can generate something temporally determinate. Of course, once one has accepted atemporal entities in the noumenal world, it should not be so difficult to accept the idea that something else is merely temporally indeterminate. The more difficult point concerns how something temporally indeterminate can bring about something temporally determinate. The underlying issue here, however, is familiar from discussions of the causal theory of time, according to which temporally determinate relations are derivative on causal relations. This is not to say that such a position is uncontroversial, but only that it is not unique to Kant’s position.

⁴² Kant makes a similar point in the *Metaphysical Foundations of Natural Science* at 4:498–499, where he argues that there can be no more than two kinds of primitive motive forces.

to have knowledge of powers in natural science only from their effects, and since powers are indeterminate given that, qua grounds, they are known only through their effects (i.e., their determinate consequences), we see, once again, that a causal power is necessarily indeterminate.⁴³

Kant's reflections on causality as reflected in his *Metaphysics Mrongovius* lectures thus offer helpful explanations of important features of his model of causality in general and his notion of causal powers in particular. In addition to supporting the idea that Kant understands causality not in terms of events (or even determinations of substances), but rather in terms of causal powers, they reveal that the relationship between cause and effect is irreducible and asymmetrical, that the asymmetry implies the determinacy of the effect and the indeterminacy of the cause, and that the asymmetry is derivative of an active-passive dichotomy, since both the determinacy of the effect and the indeterminacy of the cause stems from the activity of a causal power.

MAKING SENSE OF ACTIVITY, OR THE "CAUSALITY OF THE CAUSE"

Before we turn to several consequences that follow from understanding Kant's model of causality in this way, it may be helpful to reflect on the notion of activity that is central to and, in fact, distinctive of his model. For one might be concerned that very little positive content can be attached to the activity of a causal power or, as Kant sometimes puts it, "the causality of the cause." In particular, since, as was suggested above, one can, it seems, directly experience only the determinate effects of causal powers and neither the essential grounds of substances nor the activity of grounding by which those effects are brought about, an empiricist would naturally object that the activity of a causal power, along with any model of causality of which it is a part, is unintelligible or at least obscure. Accordingly, if it were possible to explain this notion of activity in terms that an empiricist such as Hume would find meaningful, Kant's model of causality would not be subject to one important line of

⁴³ While this argument is purely epistemological in nature, it can be strengthened in one of two ways. First, if it is conjoined with the premise that appearances are not determinate unless something makes them so, then powers will not be determinate according to the arguments given above and this argument simply complements them. Second, if one accepts an epistemological reading of Transcendental Idealism, then the fact that powers are indeterminate in an epistemological sense is all that one could possibly need to establish in the first place.

objection. In attempting to render Kant's notion of activity intelligible, several strategies are worth pursuing, namely those appealing to considerations internal to Kant's own architectonic system, to scientific instances of causality, and to certain features of self-consciousness.

Architectonic Considerations

If the Analogies of Experience are first and foremost arguments that attempt to establish the necessity of the relational categories for knowledge of objective temporal relations, one might attempt to render intelligible the notion of activity that these arguments establish by attending to the fact that this notion must be represented by categories, that is, pure or nonempirical concepts of the understanding. Investigating such nonempirical concepts might seem especially appropriate for several reasons. First, if Hume is right that we do not have a sensory impression of causality – a view Kant incorporates into his own position by maintaining that we do not perceive causal relations directly as given through sensible intuitions – then the notion of activity that Kant thinks of as central to causality should not be one that we could perceive empirically in sensible intuition, but rather must be represented in some other way, and Kant's distinction between intuition and concept strongly suggests that the categories, as nonempirical concepts, must be the appropriate representational means for depicting it. Second, since the categories are derived from the logical form of judgments, it is clear that they must be nonempirical, just as the notion of activity is. Moreover, since the categories are ontological (as opposed to purely logical) by virtue of representing objects, they would seem to be ontological in just the sense that activities are intended to be. Finally, if the Analogies presuppose the relational categories from the Schematism and Metaphysical Deduction, it would not be unreasonable to consider their origins in these texts in attempting to ascertain their precise meaning.

Unfortunately, appealing to the Metaphysical Deduction and the Schematism to determine the precise meaning of causal activity does not provide a satisfactory response to the empiricist charge. Although it is the specific task of the Metaphysical Deduction and the Schematism to derive the table of categories and to provide spatio-temporal meanings for them, it turns out that neither one can provide additional clarification of the notion of the activity Kant describes as “the causality of the cause,” even if they can explain other aspects of causality (e.g., the temporal nature of the effect). The Metaphysical Deduction derives the

category of causality from the logical form of hypothetical judgment, which states that the truth of one proposition depends logically on the truth of another. Logical dependence is not, however, the same as causal dependence, a point Kant makes explicitly both in his pre-Critical period and in the *Metaphysics* Mrongovious lectures by distinguishing between logical and real ground-consequence relationships. Nor does logical dependence capture what is potentially obscure in Kant's model of causality, namely the *activity* involved in the production of an effect, since logic is, in that sense, static. Kant himself stresses that there is a significant difference between the table of judgments, which belongs to formal logic, and the table of categories, which belongs to transcendental logic. The former "abstracts from all content of cognition," whereas the latter deals with "the manifold of sensibility that lies before it *a priori*" (A76/B102). More specifically, Kant claims that the table of categories (and, in particular, its concern with objects) introduces "a transcendental content" (A79/B105) into content-less forms of judgment. It is clear that the activity of the cause is at least part of this "transcendental content," but what the transcendental content is and how it is added to the notion of logical dependence expressed in hypothetical judgment are no clearer than is the notion of activity involved in causality in the first place and thus cannot be used to explain it.

Nor does the Schematism help on this point.⁴⁴ It is responsible for attributing a spatio-temporal meaning to the category of causality as follows: "The schema of the cause and of the causality of a thing in general is the real upon which, whenever it is posited, something else always follows. It therefore consists in the succession of the manifold insofar as it is subject to a rule" (A144/B183). In short, causality entails that the one state follows the other state (i.e., the effect occurs) according to a causal rule. However, by focusing on succession, which, as we have seen above, is simply the effect brought about by the cause, it does not clarify the activity of the cause. "Follows" and "subject to a rule," which would seem to stem from the understanding rather than intuition, are too vague to clarify

⁴⁴ One might attempt to emphasize the importance of the Schematism (and downplay the *Metaphysical Deduction*) by citing various passages in the *Critique* where Kant seems to say that the categories have no meaning independently of their application to sensible intuition. For these passages might be read as indicating that insofar as the table of categories is derived from the table of judgments (in the *Metaphysical Deduction*) they are meaningless, in contrast to the schematized categories, which alone have meaning. For an argument against such a reading, see my "Transcendental Idealism and the Categories," *History of Philosophy Quarterly* 19 (2002): 191–215.

what the causality of the cause is. Thus, Kant cannot easily explain the “causality of the cause” by appealing to the fact that causality is represented through a (schematized) category; he needs to appeal to other means.

Scientific Considerations

To determine whether scientific considerations might be of help in clarifying the notion of activity involved in Kant’s abstract model of causality, it might be useful to ask the following question first: Very generally, given that this model has been described at such a high level of abstraction, can it be consistently applied to the concrete examples of causality that Kant explicitly developed? On the basis of our answer to that question we can then ask the question with which we started this section: Can we really make adequate sense of, or attach enough specific content to, the generic notion of activity by means of which an effect is brought about according to Kant’s abstract model of causality? To address the first question, I consider briefly how Kant fills out his abstract model of causality in his account of attractive and repulsive forces in physics.

In his *Metaphysical Foundations of Natural Science* Kant devotes considerable attention to specific instances of causality in the realm of physics by discussing Newtonian attractive and repulsive forces at length. Specifically, in the Dynamics Kant argues that attractive and repulsive forces are necessary for a body to fill a determinate region of space, while in the Mechanics he argues that they are necessary for the communication of motion. In both the Dynamics and the Mechanics Kant holds that the respective masses of bodies are relevant to how strongly they attract and repel each other. Further, in the Mechanics (4:542–543) Kant argues that the masses of bodies cannot change in the communication of motion.⁴⁵ Though Kant does not explicitly argue for the analogous point in the Dynamics (i.e., that mass cannot change in the mutual attempts of bodies to penetrate and to resist the penetration of each other), a parallel argument can be constructed, since penetration and the resistance of penetration are simply special instances of the communication of motion (as Kant remarks at 4:536–537).

⁴⁵ An important consideration that supports this point is Kant’s claim in the First Analogy of Experience that “in all change of appearances substance persists, and its quantum is neither increased nor diminished in nature” (B224). While identifying the quantum of substance with mass may require argument, it is clear that Kant does accept such an identification.

Kant's Newtonian theory of attractive and repulsive forces can thus be seen as filling out his abstract model of causality as follows. The first point to note is that Kant identifies bodies with spatial substances so that one can, speaking loosely, say that bodies are causes just as it is permissible to say that substances are. However, just as it is not, strictly speaking, substances per se that actually bring about an effect, but rather the causal powers that they have, so, too, it is not the bodies as such that are the cause of changes, but rather the attractive and repulsive forces that they have. Moreover, the effect, which Kant understands generally as a change from one determinate state to another, is a change of motion (a spatial state of a body) in the case of physics. It is true that attractive forces bring bodies closer together, while repulsive forces move them farther apart, but this difference clearly pertains solely to the *direction* in which bodies are caused to move with respect to each other. Except for this point, their similarities dominate, since they both cause changes in the determinate spatial position of each other. Further, because bodies' masses do not change while determining the spatial properties of bodies, it is plausible to think that they form part of the essential nature of these bodies. Accordingly, what Kant says about attractive and repulsive forces fits in well with many of the general features of his abstract model of causality.

But it is striking that the parallels do not stop there. In Kant's account of attractive and repulsive forces one can find a plausible physical counterpart to the notion of activity that is central to his abstract model of causality. For at the level of physics, it is plausible to view the activity that produces a determinate change of state as the *exercise* of attractive and repulsive forces; whatever else an exercise of a force might be, it must be active. Moreover, the notion of the exercise of forces can even illustrate the three specific features that were attributed to Kant's notion of activity above, namely that it be irreducible, asymmetrical, and indeterminate. The exercise of these forces is irreducible because it cannot, so Kant believes, be reduced either to the motions that they produce or to the bodies in which they inhere. It is asymmetrical because the *relata* are very different; on the one side is a dynamical body (qua seat of forces) and its mass (i.e., an active spatial substance along with its causal powers and essential nature), whereas on the other is, for example, a change in motion (i.e., a determinate spatial event) that (passively) occurs. Finally, the exercise of a force is clearly not a determinate event or state of affairs that we could observe in the same way that we can a body's motion. One can see this last point even more clearly by considering that the exercise

of a force is merely an “attempt” at bringing about a change of state. If two bodies attract and repel each other with equal force such that they are in a state of equilibrium and thus do not change their state with respect to each other, their forces continue to be exercised, that is, each body is still attempting to change the state of the other, despite the fact that these attempts do not result in a change of state due to the exercise of the counterbalancing force in the opposite direction. Because the exercise of attractive and repulsive forces is an irreducible, asymmetrical, and indeterminate activity, it displays the same structure as did the notion of activity involved in Kant’s abstract model of causality.

To avoid misunderstanding how Kant’s account of forces in physics might illustrate his abstract model of causality in this way, it is important to recognize that the term “force” is often used in a variety of ways, not all of which are consistent with Kant’s view, if they are taken literally. For example, sometimes it is said that a body can transfer its force to another or that a body might exercise its force at one moment in time and not at another. From Kant’s perspective, these ways of speaking require reinterpretation in line with his metaphysical account. If substances act continuously according to their unchanging natures, then, due to the analogy developed above between Kant’s abstract model of causality and his account of forces in physics, it follows that bodies must exercise their forces continuously and uniformly in accordance with their mass. What changes is thus not the exercise of the forces themselves, but rather what *effects* the exercise of those forces will have. That is, a given substance acts in the same way at all times, but this activity can nonetheless cause different things to happen because the circumstances of the substance can be different (i.e., because different substances can stand in different relations to each other). As those circumstances change, so will the effects that the substance brings about. Thus, two bodies located three feet apart will cause a greater reciprocal acceleration toward each other than those very same bodies separated by three light years. Their attractive forces are no different in the two cases; the bodies have exactly the same intrinsic and essential properties. What is different is the effect that those attractive forces have, and that difference is due to the different relations in which the bodies stand to each other. One might think of the constant and immutable “sphere of activity” that a substance has as being akin to a dynamical force field that spreads out over a certain region of space in a specific way and that causes objects to behave in different ways only to the extent different objects come to occupy different positions within the field. Accordingly, as long as one is careful to understand our various

ways of speaking about forces in the appropriate way, one can identify a coherent interpretation of attractive and repulsive forces in physics in terms of Kant's abstract model of causality.

But even if we can provide a consistent interpretation of Kant's abstract model of causality in terms of the concrete example of attractive and repulsive forces in physics, it still leaves open the question that was raised above, namely whether we can really make proper sense of the generic notion of activity invoked in his abstract model. Though Kant reinterprets Newtonian attractive and repulsive forces in terms of his own metaphysics, he follows Newton (and other empiricists) in agreeing that the notion of activity that is expressed in the notion of an exercise is not something that we can directly observe (in intuition alone). Just as one does not literally see "the causality of the cause" or one billiard ball imparting motion to another, one does not see the exercise of attractive and repulsive forces. This point can be seen especially clearly in the case described above, where the forces of the two bodies counterbalance each other so that no change of state occurs, since in such a case there is no change to be seen and one must therefore infer the fact that forces are being exercised from other considerations. Since one can see nothing beyond the effects of these forces, there is no empirical content to physical forces per se (as opposed to their effects, which contain all such content).⁴⁶

If the objection arises for Newton just as it does for Kant, one might think that Kant can respond to the objection in the same way that Newton does. Unfortunately (though perhaps not surprisingly), Newtonian responses to the objection are ultimately unavailable to Kant. Newton himself, at least in the *Principia*, seems to skirt the issue, since he claims to be stating nothing more than the mathematical principles of natural philosophy and thus is committed only to formulating a mathematically

⁴⁶ Kant is clearly aware of this point. As early as 1747 (in the opening paragraphs of the *True Thoughts on the Estimation of Living Forces*) Kant sees the dilemma one faces in describing forces. He explicitly criticizes those who describe forces, in this case moving forces, exclusively in terms of their empirically observable effects, namely the ability to cause motion, since such an explanation can appear vacuous. However, he is also aware that empiricists may not find forces intelligible. In fact, he objects to Leibniz's attempt to explain force in terms of "entelechia" for precisely that reason. Kant's solution in the *True Estimation* is to describe forces as being "active." Thus, Kant has already adopted the central feature of his Critical account in his first pre-Critical publication, though at this point, he has no detailed explanation of what "activity" is, except in terms of its obvious contrast with passivity. For further discussion of these issues, see my "Kant on Force and Extension: Critical Appropriations of Leibniz and Newton," in *Between Leibniz, Newton and Kant: Philosophy and Science in the 18th Century*, ed. W. Lefevre (Dordrecht: Kluwer Academic, 2001), pp. 111–127.

consistent interpretation of forces. For example, his second law of motion: “The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed,” nowadays better known as $F=ma$, need not give any direct empirical content to forces as long as forces have an appropriate mathematical meaning.⁴⁷ In fact, the purpose of the second law seems to be to define the notion of force in mathematical terms that, unlike the metaphysical concept of force, have an immediately observable empirical content. Thus, every use of the term “force” in the *Principia* can be replaced by “*ma*.” In light of the fact that Newton’s project is limited to establishing mathematical principles (which leaves completely open the question of whether nature really contains entities that correspond to the fundamental terms of the mathematical principles), this kind of reductionist strategy is possible. Kant, by contrast, is interested in the *metaphysical* rather than the *mathematical* principles of natural philosophy and for that reason he cannot put the question off as Newton does.⁴⁸ As a result, if Kant is to be able to respond to this objection, he will have to go beyond Newtonian attractive and repulsive forces and find some other means for rendering “the causality of the cause” and “the exercise of forces” intelligible.

Self-consciousness

At this point, it is helpful to consider how Kant’s abstract model of causality can be illustrated not only by physical forces, but also by a specific kind of consciousness. For careful observation will reveal that what distinguishes activity in consciousness from the activity of forces in physics is the fact that we do have an immediate awareness of the self’s synthetic

⁴⁷ Isaac Newton, *Mathematical Principles of Natural Philosophy and His System of the World*, trans. A. Mott, revised by F. Cajori (Berkeley: University of California Press, 1934), p. 13.

⁴⁸ It is worth noting that, like Kant, later Newtonians, such as Euler and Maupertuis (and perhaps even Newton himself, in contexts that extend beyond the scope of the *Principia*), are interested in determining what nature is really like and whether it contains the attractive and repulsive forces posited only mathematically in the *Principia*. Yet Maupertuis, for example, recognizes that, given the epistemological resources available to an empiricist, no empirical content could be ascribed to forces and thus one cannot attribute them to nature (see, e.g., Pierre Maupertuis, *Recherche des loix du mouvement* (Berlin, 1746)). As a result, Kant cannot respond to the empiricist objection to forces as Newton or Newtonians do, since, insofar as they address the issue, they *deny* such forces. (Technically speaking, Maupertuis does not rule out the possibility that such forces exist. Rather, following Locke, he takes a skeptical epistemological position according to which human beings simply have no access to whatever “real essences” might exist.)

activities, whereas we have no direct awareness of the exercise of Newtonian forces. As a result, our awareness of this specific synthetic activity will allow us to render intelligible Kant's generic notion of activity as it is employed in his abstract model of causality. Since providing an accurate description of the structure of self-consciousness is notoriously difficult, we must begin by carefully distinguishing different aspects of self-consciousness and pay special attention to the argumentative contexts in which Kant discusses self-consciousness, since they determine what aspect of self-consciousness is of interest to him in each case.

One must first distinguish between inner sense and apperception (i.e., self-consciousness). Though Kant does not always express his views on inner sense with perfect consistency, its primary systematic importance lies in its similarities and differences with outer sense.⁴⁹ Kant's general idea is that we can intuit external or spatial objects by means of outer sense and internal or nonspatial objects by means of inner sense. The result in either case is empirical knowledge (of either spatial or nonspatial objects), since both inner and outer sense provide the material for empirical intuitions, which are required to justify empirical knowledge. However, inner sense must also be distinguished from self-consciousness, because inner sense provides us with *knowledge* of the self, whereas apperception, which is not a source of intuitions, does not amount to knowledge ("*Erkenntnis*") per se, but rather merely an *awareness* ("*Kenntnis*") of the self.

In §24 of the second edition Transcendental Deduction Kant addresses the complex relationship between inner sense and self-consciousness indirectly by exploring how to resolve a paradox that arises from the very idea of knowledge of the self, namely that we must (paradoxically) be both active and passive with respect to ourselves in knowing ourselves. On Kant's general account of knowledge, if we are to have knowledge at all, then the object of knowledge must affect us in order to be *given* to us in intuition. That is, knowledge requires that we must be passive in some way, at least with respect to the object of our knowledge. This requirement poses no special difficulty in the case of knowledge of objects that exist externally to us, since it is plausible that they can affect or act on us causally so that we can passively receive sensory information from them. However, in the case of knowledge of the self, matters are more complicated. Since the object of knowledge in such a case is not some

⁴⁹ For a discussion of some of the inconsistencies in Kant's doctrine of inner sense, see Henry Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1982), pp. 255–271.

external object, but rather the self, one is forced to admit that it is the self that acts so as to deliver the relevant intuition. But this seems to entail that to have knowledge of the self, the self must both act on and be acted on by itself, which can seem paradoxical.

Kant argues that this paradox can be resolved by means of his distinction between inner sense and apperception. As we saw above, inner sense is the passive faculty through which objects affect us. The understanding, by contrast, is an active faculty in us by means of which inner sense can be acted on or determined. Because inner sense and apperception are distinct from each other, one need not hold that self-knowledge requires one and the same thing to be both active and passive in the same respect. Accordingly, Kant's distinction between inner sense and understanding provides the basic framework for resolving the paradox of self-knowledge.

What is relevant about this resolution for current purposes is that in the course of explaining how the understanding can be an active faculty that determines or affects inner sense, Kant explicitly claims that the self can be aware of its activity, for example, in syntheses. Kant clearly states that the understanding "exercises that action on the **passive** subject, whose **faculty** it is, about which we rightly say that the inner sense is thereby affected," and its "synthesis, considered in itself alone, is nothing other than the unity of the action *of which it is conscious as such* even without sensibility" (B153, italics added). As we see in more detail below, earlier in the second edition Transcendental Deduction Kant also claims that a specific feature of *self-consciousness* requires that we be "conscious of synthesis" (B133). Since Hume consistently denies that he has any internal impression of a necessary connection or causal activity within himself, it is important to investigate this point with great care.⁵⁰

In §24, immediately after introducing the distinction between understanding (or apperception) and inner sense and claiming that we can be aware of the understanding's synthetic activities, Kant provides several illustrations in support of this claim:

We cannot think of a line without **drawing** it in thought, we cannot think of a circle without **describing** it, we cannot represent the three dimensions of space at all without **placing** three lines perpendicular to each other at the same point, and

⁵⁰ While Hume is quite explicit in this denial, it is significant that his second definition of causality in the *Treatise* requires that the idea of the cause "determines the mind to form the idea" of the effect (p. 170). If "determines" is a causal notion and if one can be aware of it occurring, then Hume is implicitly committed to the thesis that we have an impression of causality in self-consciousness. As a result, Hume must understand "determines" in some other sense.

we cannot even represent time without, in **drawing** a straight line . . . , attending merely to the action of the synthesis of the manifold through which we successively determine the inner sense, and thereby attending to the succession of this determination in inner sense. Motion, as action of the subject (not as determination of an object), consequently the synthesis of the manifold in space, if we abstract from this manifold in space and attend solely to the action in accordance with which we determine the form of **inner sense**, first produces the concept of succession at all. The understanding therefore does not **find** some sort of combination of the manifold already in inner sense, but **produces** it, by **affecting** inner sense. (B154–155)

All of the examples Kant describes here are cases in which one can be immediately aware of the understanding's activity in synthesis. While Kant explicitly casts doubt on whether one could represent a line at all without first drawing it, regardless of whether one follows him on this point one can still see him asserting a clear and tangible difference between (1) drawing a line in thought and (2) a case where one first becomes aware of a point and then watches as that point moves across one's visual field in such a way that it seems to leave behind a series of points forming a continuous line. In short, in the first case, one is drawing a line, whereas in the second, one is merely watching as a line comes into existence in thought. It is plausible to interpret these cases as follows. In drawing a line in thought the understanding is *actively producing* the line and one can be immediately aware of this activity, whereas when one is watching a line being formed in thought, one is passive insofar as one is not directly aware of the creation of the line as depending on the understanding's activity (even should it be true that the line is created by such an activity).

The same kind of point is also present in Kant's distinction between motion as the determination of an object and motion as the action of a subject.⁵¹ In the one case, there is an empirical object that changes its position in space and we represent that change as occurring in the object. In the other, there is the activity of the subject by means of which our representation of a change of position as such is produced, an activity that does not necessarily require (even the representation of) an externally existing object in motion. In a footnote meant to clarify this example, Kant interprets the latter case as involving the "**description** of a space" (B155), rather than a determination of an object. Whether one can determine that an object is in motion without also describing space, it is clear that there are two distinct points at issue. One concerns the

⁵¹ Michael Friedman helpfully draws attention to this example in *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992), pp. 40, 131, and 200–201.

attribution of motion to an object, while the other concerns what activity the subject is engaged in when it represents motion in space at all. It is in this latter case that we can be directly aware of our own activity in producing a representation of motion (just as we are when we draw a line in thought).

Further, Kant suggests that we are immediately aware of the understanding's activity in a wide range of cases and not just in what might appear to be the special cases described above. In a footnote to his discussion of the paradox of self-knowledge, he suggests that every "act of attention" (B156) is an example of us affecting inner sense. Again, the philosophical basis for Kant's point here can be made quite clear. It is one thing to have, or be conscious of, a certain representation. It is quite another to focus one's attention on it (or some aspect of it). Paying attention to a specific feature of our intuition is clearly distinct from simply having that intuition given through (inner or outer) sense, since one can have an intuition without paying attention to it at all. That is, it is plausible to think that the way in which one can, apparently at will, focus one's attention on one's own mental states should be understood as an act that we ourselves perform and can be aware of performing. Thus, whether one is drawing a line in thought or simply focusing one's attention on one's own mental states, it is clear that we are immediately aware of the understanding's activity, and this specific kind of self-awareness allows us to render intelligible the generic notion of activity that Kant employs in his abstract model of causality.

However, Kant's interest in our awareness of the understanding's activity extends beyond the value these particular examples have for clarifying a notion that is central to his model of causality. For he attempts to build a more robust account of *transcendental self-consciousness* on the empirical awareness of activity illustrated by these examples. Kant's most famous discussion of transcendental self-consciousness occurs near the beginning of the second edition Transcendental Deduction where he addresses the issue of how we can explain the fact that various representations are mine. To this end, he describes apperception as "that self-consciousness which, because it produces the representation **I think**, which must be able to accompany all others . . . , cannot be accompanied by any further representation" (B132). A few lines later, he argues that the identity of apperception "contains a synthesis of the representations, and is possible only through the consciousness of this synthesis" (B133). The basic idea behind Kant's argument is that I can know that representations are mine only if I know that one and the same I has each one, but this can be

known only if (1) I connect them (given that they do not come into my consciousness already connected) and (2) if I am aware that I am connecting them (since only my awareness of my connecting them allows me to know that each representation is being had by one and the same self).⁵² Further on in the Transcendental Deduction Kant attempts to show that these connections must be represented by categories and that there is a tight argumentative link between such connections and knowledge of objects. In this way, Kant hopes to prove the objective reality of the categories, that is, that the categories are necessary for knowledge of objects.

While there is little consensus about much of Kant's argument in the Transcendental Deduction, there is widespread agreement about the idea that Kant is attempting to give an account of the fact that my representations are mine. And on this point, Kant appears to enjoy a significant advantage over Hume, who faces a serious dilemma.⁵³ Hume's empiricist principles seem to commit him to the claim that either we can know the self directly through an impression (just as we know any other object, whether external or internal) or we cannot know it at all. Hume avoids the former horn of the dilemma (by rightly noting that we do not have an impression of the self on a par with our impressions of other objects) only to succumb to the latter by asserting that the self is "a bundle" of perceptions without being in a position to explain what a bundle is and why it is not a fiction in the way that bodies are.⁵⁴ Hume himself famously admits (in the Appendix to the *Treatise*) that his bundle theory of the self cannot account for "the principles, that unite our successive perceptions in our thought or consciousness."⁵⁵

Kant agrees with Hume that the first horn of the dilemma should be avoided, since he not only shares Hume's insight that we do not have an impression of a single, enduring self that remains the same throughout all the changes in our perceptions, but also develops this critical insight in detail in the context of his discussion of traditional metaphysical arguments. Kant expresses the point in his own words by noting (against

⁵² For a helpful reconstruction of this strand of Kant's argument, see Henry Allison, *Kant's Transcendental Idealism*, pp. 133–172, esp. 142–143.

⁵³ For discussion of Kant's reply to Hume on the issue of self-consciousness, see Patricia Kitcher, "Kant on Self-Identity," *Philosophical Review* 91 (1982): 41–72.

⁵⁴ The self cannot be a fiction in *precisely* the same way in which bodies are because the explanation of why bodies are fictions depends on the self (specifically, on the imagination's principles).

⁵⁵ Hume, *A Treatise of Human Nature*, p. 636.

rational psychologists such as Descartes and Leibniz) that we do not have an intuition of the self as a thinking substance (or subject). Yet Kant disagrees with Hume that one must therefore accept the second horn of the dilemma, for, as we have seen above, Kant thinks that we can be aware of the self and its identity *indirectly*, that is, by being aware of the activity of the self when it connects its various representations and by then inferring that it is one and the same self that does the connecting. By suggesting that the self can become aware of its identity not directly as an object of consciousness, but rather indirectly as the subject of activities of which we can be conscious, Kant is attempting to resolve the dilemma Hume faces.

At the same time, Kant does not beg the question against Hume by appealing to something that Hume could not in principle accept. It is true that Hume denies having any impression of necessary connection on the grounds that he has neither an external nor an internal impression of such a necessary connection. But if Hume's denial that we have an internal impression of necessary connection is based on the expectation that our internal impressions have to be exactly analogous to our external impressions (and then the fact that our internal impressions do not reveal anything similar enough), then Kant can be seen as pointing out that Hume's expectation is unfounded. Specifically, the analogy with external impressions may have led Hume to think that internal impressions would still be impressions of objects, and thus not to have been attentive enough to the features of consciousness to which Kant wants to draw our attention. As a result, Hume could have noted the distinction Kant is making between a subject and an object of consciousness and the way in which Kant explains how the activity of the self is necessary for one to become aware of the identity of the self in self-consciousness.⁵⁶

Whatever merits Kant's account of self-consciousness might have with respect to the dilemma Hume faces, it is relevant to note that the structure of self-consciousness illustrates various features of Kant's abstract model of causality. The first point to note is that it is clear that the syntheses the self is aware of in self-consciousness are activities. Without these activities, there would be no connections between our representations. Specifically, these activities would seem to be instances of a particular kind of activity, namely an activity whereby a connection between representations is brought about as its effect. (Kant also states quite clearly

⁵⁶ Had Hume understood self-consciousness in this way, he might well have revised other parts of his position, too (his epistemology, his general account of causation, etc.).

that apperception *produces* "I think", which is obviously an activity of the self.) Moreover, the effect is determinate insofar as the particular kind of connection required by self-consciousness, namely a connection represented by the categories, *determines* our representations of an object (and its change of states; cf. B128). In light of the fact, however, that determinacy is the result, product, or effect of these synthetic activities, the synthetic activities cannot themselves be determinate.⁵⁷ We can thus discern a structure to self-consciousness that is similar in several respects to what we saw in Kant's abstract model of causality.

There is yet another aspect of self-consciousness, distinct from the synthetic activity required to explain how I can know that my representations are mine, that is also relevant, though in a different way, to understanding Kant's model of causality.⁵⁸ In the Paralogisms, where Kant rejects various arguments that rational psychologists attempt to offer on the basis of self-consciousness alone, he explores the peculiar structure of self-consciousness in detail, arguing for an asymmetry similar in certain respects to the asymmetry that a cause bears to its effect. In this context, Kant not only distinguishes between the object and subject of consciousness, but also notes that our awareness of the subject in self-consciousness is radically different from our awareness of objects.⁵⁹ It is one thing to say that we do not see the self in the same way as we do spatio-temporal

⁵⁷ One might object (e.g., on behalf of Leibniz) that because Kant must rely on consciousness to render intelligible the notion of activity, he cannot obviously or immediately claim that this same notion of activity is also at work in physical cases, especially if Kant admits to important differences between the two cases. Perhaps, the objection continues, physical cases must be unintelligible or perhaps they stand in need of a deeper metaphysical explanation (e.g., in idealistic terms). However, the activity that we are aware of in our own consciousness is supposed to be simply a *specific* instance of a *generic* notion of activity that Kant wants to employ in his general model of causality. That is, the objection provides no reason to think that Kant *cannot abstract* from those particular features of consciousness that happen to attend the activity we can be directly aware of in self-consciousness.

⁵⁸ For helpful discussions of Kant's account of self-consciousness, see Manfred Frank, "Is Subjectivity a Non-Thing, an Absurdity [Unding]? On Some Difficulties in Naturalistic Reductions of Self-Consciousness," in *The Modern Subject: Conceptions of the Self in Classical German Philosophy*, ed. K. Ameriks and D. Sturma (Albany: SUNY Press, 1995), pp. 177–197, and Karl Ameriks, "From Kant to Frank: The Ineliminable Subject," in *ibid.*, pp. 217–230. See also chap. 2 of David Carr, *The Paradox of Subjectivity* (New York: Oxford University Press, 1999).

⁵⁹ Berkeley encounters a similar difficulty on this point, since for him ideas are necessarily passive, yet the self is supposed to be active. He solves the problem by claiming that he has not an idea, but rather a "notion" of the self. While introducing a notion of the self in this fashion might initially seem to be ad hoc, Berkeley may be responding to an independent philosophical concern. For an excellent discussion of Berkeley's position,

objects. It is another thing altogether to say that we cannot even *describe* it in any positive way by using the predicates that we use in describing objects, as it makes the use of such predicates possible in the first place.⁶⁰ On Hume's view, there is presumably nothing special about the self that precludes our seeing it. Rather, it simply does not exist (as a single, identical entity), and its nonexistence accounts for the fact that we do not see it (as such). On Kant's account, there is something very special about the self or at least about our mode of conscious access to it. Since it is not a directly observable object, but rather, insofar as we can be aware of it at all, issues in an activity rather than a determinate state of an object, it is not something that can be described by using the same concepts that apply to the determinate states of external objects that we observe in exactly the same way. According to Kant, therefore, Hume mistakes our inability to describe the self in typical object-language terms with an inability to be aware of it at all.

Kant also points out that, as a consequence of this fact, the self or the synthetic activity that we are aware of in self-consciousness is neither a determinate phenomenon nor a noumenon. Like the "causality of the cause," it requires a distinct explanatory level. In a detailed footnote added to the second edition, which is worth quoting at length, Kant describes several of these points as follows:

The "I think" is, as has already been said, an empirical proposition, and contains within itself the proposition "I exist." But I cannot say "Everything that thinks, exists"; for then the property of thinking would make all beings possessing it into necessary beings. Hence my existence also cannot be regarded as inferred from the proposition "I think," as Descartes held (for otherwise the major premise, "Everything that thinks, exists" would have to precede it), but is rather identical to it. It expresses an indeterminate empirical intuition, i.e., a perception . . . , but it precedes the experience that is to determine the object of perception through the category in regard to time; and here existence is not yet a category, which is not related to an indeterminately given object, but rather to an object of which one has a concept, and about which one wants to know whether or not it is posited outside this concept. An indeterminate perception here signifies only

see Robert Adams, "Berkeley's 'Notion' of Spiritual Substance," *Archiv für Geschichte der Philosophie* 55 (1973): 47–69.

⁶⁰ For perhaps different reasons, Shoemaker endorses both of these points in "Self-reference and Self-awareness," *Journal of Philosophy* 65 (1968): 555–567, esp. 563, when he says: "I think that the main source of trouble here is the tendency to think of awareness as a kind of perception, i.e., to think of it on the model of sense-perception," and p. 564, when he notes that if the use of first-person pronouns as a subject "were not possible then there would be much else, and much that we take for granted, that would also not be possible."

something real, which was given, and indeed only to thinking in general, thus not as appearance, and also not as thing in itself (a noumenon), but rather as something that in fact exists and is indicated as an existing thing in the proposition "I think." For it is to be noted that if I have called the proposition "I think" an empirical proposition, I would not say by this that the I in this proposition is an empirical representation; for it is rather purely intellectual, because it belongs to thinking in general. Only without any empirical representation, which provides the material for thinking, the act I think would not take place, and the empirical is only the condition of the application, or use, of the pure intellectual faculty. (B422-423 n.)

In this passage Kant is ostensibly clarifying what can and cannot be accepted about Descartes's cogito argument (as he interprets it). The "I think" is an empirical proposition, but a very unusual one insofar as the representation of the self (or "I") contained in the proposition is not itself empirical. It is empirical in the sense that it depends on something empirical being given to thought, since only if material is given for thought can the self (qua intellectual faculty) combine the manifold contained in that material into a unified representation and only by that means can one be immediately aware of the truth of the proposition. At the same time, because it precedes the categories and thus the determination of any object of perception, it is not a determinate empirical object, that is, a phenomenon, and cannot be described by conceptual or discursive means. Still, Kant wants to insist that we can be aware of it, namely indirectly through our awareness of our intellectual activities. In this way Kant shows how self-consciousness reveals the very same structure that is explicit in his model of causality. Moreover, in light of these parallels and the fact that Kant's account of self-consciousness has distinct advantages over Hume's, one can also come to understand in an intuitive way the potentially most obscure aspect of Kant's model, namely the notion of activity employed in "the causality of the cause."

Accordingly, we now have an adequate answer to one important question that one might raise about Kant's model of causality, namely: How is one to understand the notion of activity that it employs? We have seen that neither Kant's *Metaphysical Deduction* nor his *Schematism* can clarify how activity is to be understood, since their focus is on deriving causal from logical dependence and on introducing temporal meanings into purely ontological concepts, respectively. And while Kant's account of attractive and repulsive forces in physics illustrated Kant's abstract model of causality in metaphysics in several respects, epistemological limitations to our access to such forces stood in the way of its providing additional

clarification of that notion. Instead, Kant's distinctive account of self-consciousness and the awareness of our intellectual activities that makes it possible provided the clarification that was called for on this issue.

IMPLICATIONS FOR THE SECOND AND THIRD ANALOGIES
AND FOR THE "CRITICAL TURN"

If Kant's model of causality can be understood in this way, two questions naturally arise about the consequences that this model might have for our understanding of the *Critique*. First, how does this understanding of Kant's model of causality affect our understanding of the arguments of the Second and Third Analogies as presented in Chapter 3? Second, in light of the particular ways in which Kant is drawing on elements from his pre-Critical model, what *specific* implications does that have for understanding the relationship between Kant's pre-Critical and Critical periods?

Implications of Kant's Model for the Second and Third Analogies

In the second edition formulation of the principle of the Third Analogy, Kant asserts that mutual interaction is necessary for knowledge of the coexistence of substances only if those substances are *spatial*. This restriction to spatial substances might appear puzzling. Although Kant added a few new paragraphs to the second edition, they do not pertain to spatiality per se. Moreover, the arguments Kant develops do not invoke any explicit spatiality in their premises. That is, the arguments seem to hold for any kind of substance and are not necessarily restricted only to spatial substances. So, why the restriction?

In response, one might appeal to commitments that are made explicit in the Refutation of Idealism. The claim of the Refutation of Idealism is that knowledge of the temporal order of my own mental states requires a spatial substance that is distinct from me. Its argument runs roughly as follows. According to the First Analogy, temporal ordering requires a permanent substance. The distinctive move in the Refutation of Idealism is that *I* cannot be the permanent substance that would be required for (my knowledge of) the temporal ordering of my own mental states, because I cannot know that I am an immaterial substance, a view Kant provides some support for in the First and Second Paralogisms. As a result, there must be some permanent substance distinct from me that can serve to make possible (my knowledge of) the ordering of my mental states. The only substances (I know of) that are distinct from me are spatial. Therefore,

knowledge of the order of my own mental states presupposes (knowledge of) spatial substances. If the argument of the Refutation of Idealism requires that all phenomenal substances are spatial, then the fact that Kant restricts the principle of the Third Analogy to spatial substances would be of no consequence.

While many details of the argument of the Refutation of Idealism are controversial, it is clear that it does not establish that all substances must be spatial, as this reply requires. At best, it establishes that the only substances that could serve to allow my knowledge of the temporal order of my own mental states must be spatial. It rules out the possibility, therefore, not of nonspatial substances, but rather only of ones that might fulfill a particular epistemic role. Nor do the First and Second Paralogisms, which are invoked in that reply, establish that all phenomenal substances must be spatial. They argue that one cannot establish that the soul is an immaterial substance on the basis of arguments of rational psychology, that is, on arguments that employ only the mere concept "I think." But it would be a considerable leap to assert, on the basis of the failure of one particular kind of *argument* for a certain claim, that the claim itself is false. Moreover, even if it could be established that we do not (or even cannot) *know* that there are nonspatial substances, it does not immediately follow that there cannot *be* such substances. As a result, neither the Refutation of Idealism nor the Paralogisms supports this explanation of the restriction of mutual interaction to spatial substances. Hence the restriction puzzle remains.

The model of causality developed above, however, suggests a different and more satisfying explanation, one that clarifies in an important way the relation between the Second and Third Analogies. In thinking through how mutual interaction could explain simultaneity, the crucial idea, which stemmed from Kant's pre-Critical model, was that simultaneity is a reciprocal and symmetrical relation. It would not be possible to determine that A is simultaneous with B without also determining that B is simultaneous with A. The example of motion, which Kant himself emphasizes in a footnote in the Second Analogy (A207/B252 n.), is analogous. For the motion of A toward B is reciprocal and symmetrical too. Moreover, the analogy also holds between the *grounds* of such reciprocal and symmetrical relations. Just as whatever causes the motion of one body toward another cannot be independent of what causes the motion of the second body toward the first, so, too, whatever determines the place in time of the one cannot be independent of what determines the place in time of the other, if they are to be determined as simultaneous states.

It is the realization of this point that forces Kant to restrict the Third Analogy to spatial substances. For if there were an immaterial substance, there would be no guarantee that it could interact with a spatial substance with the kind of reciprocal symmetry that is required for the reciprocal determination of their simultaneous states. But in the case of spatial substances, it is clear that they do interact in the requisite way. Moreover, as one can see from the way in which Kant explains causal interactions between spatial substances in the *Metaphysical Foundations*, the attractive and repulsive forces that bodies exercise on each other in generating the communication of motion and the filling of determinate regions in space display precisely the kind of reciprocal symmetry that is required for mutual interaction. As a result, in the second edition of the *Critique* in 1787, after having developed an explicit account of attractive and repulsive forces in the *Metaphysical Foundations* that was modeled on his abstract account of causality in the first edition of the *Critique*, Kant may have been much clearer that spatial substances can act on each other in ways that are consistent with mutual interaction, whereas the same cannot be said of causal interaction between spatial substances and any nonspatial substances that might exist.⁶¹

This explanation makes explicit, however, that there could be instances of causation that are not also instances of mutual interaction. Does this mean that Kant has two distinct models of causality, one for the Third Analogy, which has been developed above, and a completely different one for the Second Analogy, which would be designed to explain a completely different phenomenon, namely succession rather than coexistence?⁶² If that were the case, could one not then return to the standard view, namely that Kant accepts Hume's event-event model of causality at least for the Second Analogy? While it is true that it was primarily the symmetry considerations involved in the Third Analogy's notion of mutual interaction that entailed the rejection of an event-event model of causality and that the Second Analogy does not require the same sort of symmetry, it is clear that Kant does not think that the notion of causality invoked in the Second Analogy is radically different from that of the Third Analogy. If

⁶¹ Another possible explanation is that Kant may think that mental substances (or substances that might be alive) operate according to principles different from those that are essentially only spatial. However, I am aware of no passages from the relevant time period that would support such an interpretation.

⁶² I would like to thank Desmond Hogan for raising this possibility in a particularly pressing way.

mutual interaction is a two-way causal relation by means of which substances contain the grounds for determining simultaneous states, the notion of causality invoked by the Second Analogy should be understood in terms of the same *kind* of causal relation that is invoked in mutual interaction, except that it need not be symmetrical, or “two-way,” but can rather be “unidirectional,” or “one-way.” In other words, the Second Analogy is committed to the idea that a cause brings about its effect when an immutable ground in a substance determines the successive states of (another) substance, and it is natural to think that the notion of grounding in this case is identical to the basic notion of grounding in the case of mutual interaction.

Moreover, there is ample textual evidence for such an interpretation. For example, in the Second Analogy, not only does Kant accept the possibility of simultaneous causation (something he would not be able to do easily if the Second Analogy were committed to a model of causality that differed from that of the Third), but he also commits himself to the idea that the principle or ground of change (i.e., the cause) cannot itself change, another crucial element of the model of causality that is associated with mutual interaction. Kant's description of the respective schemata of causality and mutual interaction also makes it clear that the Second Analogy does not appeal to notions that are foreign to what Kant is committed to in the Third Analogy, since both invoke “causality,” whether in the guise of “*Kausalität*” or “*wechselseitige Kausalität*.”

In fact, even the considerations raised above ultimately reinforce the idea that the causal notions in the Second and Third Analogies are not radically different. According to Kant's official explanation of the table of categories, the third category under each heading is formed by “combining” (B110) the first two, though Kant is also clear that the third category in each case is not reducible to the first two. Applied to the category of mutual interaction, this means that mutual interaction would be formed from the combination of substantiality and causality. Accordingly, the notion of causality that Kant is utilizing in the notion of mutual interaction must be fundamentally the same as that used in the Second Analogy, even if it must also contain additional elements that are not reducible to that of substantiality and causality. As we have seen, the concept of mutual interaction requires considerations of reciprocity and symmetry that go beyond the notions involved in the first two categories under that heading, namely the idea of a substance exercising its causal power to determine the state of another substance. Therefore, we do not have two

fundamentally different notions of causality at work in the Second and Third Analogies, but rather just one primitive notion of causal power that they are using in different ways in light of their respective goals.

A second implication deriving from the model of causality described above directly concerns the claim and argument of the Second Analogy of Experience. As we saw in Chapter 3, there are two ways to interpret the claim of the Second Analogy – the weak and the strong readings – according to whether it is supposed to establish that every event has a cause (the every-event some-cause principle) or whether there are causal laws that necessarily hold in the future just as they have in the past (i.e., the same-cause same-effect causal principle). As reconstructed above, the argument appeared to be able to establish only the weaker reading. For it concluded that a cause was needed in order to determine the succession of two states, and that principle would not seem to involve anything more than the determination of this specific instance. While one might attempt to argue that the very notion of a causal *rule* might be taken to imply generality and thus at least the existence of causal laws, it is unclear how such a notion could be capable of doing such significant philosophical work. If the argument of the Second Analogy requires a causal rule only to determine in a particular instance that determination B follows (rather than precedes) determination A, why, one might ask, should the causal rule have to cover more than this particular instance? Granted, in order to account for the unity of time, there must be other successive determinations that follow determination B, but, according to the Second Analogy, that simply means that for each event that is required for the unity of time there must be a cause that determines it as successive. The argument of the Second Analogy, so understood, does not seem to require that the cause of any event be identical to the cause of any other and thus be viewed as particular instances of a more general causal law.

Given this understanding of the argument of the Second Analogy, it is tempting to rest content with the weaker reading. There is no shame in establishing merely that every event must have a cause, since that would still be a significant result, one famously called into doubt in Hume's *Treatise*. And given the apparent power of Hume's argument, one might be skeptical of any argument that would be able to establish that the future *must be* like the past, given that all of our empirical evidence does derive only from the past, as Hume's argument presupposes. Thus, philosophically, it might appear to be a good thing if Kant's argument establishes only the weaker reading.

At the same time, there is a textual motivation for the stronger reading that does not sit well with the weaker reading. For in the Second Analogy Kant repeatedly uses terms such as “universality,” “always,” and “invariably,” all of which strongly suggest that Kant has in mind causal laws that would hold over time. Moreover, Kant seems to slide back and forth between the weak and strong meanings of the principle without explicitly acknowledging the considerable philosophical difference between them. Since both of these “facts” are incompatible with the weaker reading, we should take them as an opportunity to see whether Kant’s model of causality might not provide us with materials that could (1) suggest a reasonable argument for the stronger principle and (2) explain why Kant slides back and forth between the two meanings with such ease (or at least without any discernible discomfort).⁶³

Consider, again, the argument of the Second Analogy and how it might naturally be understood in terms of the model of causality described above. The basic thrust of the Second Analogy is to argue that causality is necessary for the determination of successive states in an object, since objects do not have temporal determinacy independently of such causal determination. But if the model of causality described above dictates that causal determination is to be understood in terms of grounds, then the point of the Second Analogy is that change is to be explained in terms of grounds. Yet since the ground that causes change cannot itself be changing at the same time – at least not if one is to avoid an infinite regress – change presupposes an unchanging ground and this might be thought adequate to support causal laws, since the grounds cannot change in the future and thus must, it would seem, bring about the same effects as before. As we just saw, there seems to be a significant weakness in such an argument. For the argument of the Second Analogy, so understood, establishes only that a ground cannot change *while it is determining successive states in an object*, and not that a ground cannot change *at all*, which is what would be required in order to establish causal laws.

⁶³ Michael Friedman reconstructs a provocative argument designed to establish causal laws in “Causal Laws and the Foundations of Natural Science,” in *The Cambridge Companion to Kant*, ed. P. Guyer (New York: Cambridge University Press, 1992), pp. 161–199. However, since Friedman’s primary concern is to establish a close relationship between the notion of law-governedness contained in the Second Analogy and particular empirical causal laws, he presupposes what a Humean can deny, namely that Kant has not established any legitimate notion of law-governedness in the Second Analogy. Below, we see how Kant’s model of causality can address this concern, which Friedman’s interpretation would be free to accept as part of a more general story.

However, Kant might have thought that this weakness could be remedied by pursuing several related lines of thought. First, Kant might take recourse to the Inaugural Dissertation's idea that substances' natures are general in order to establish that the grounds that constitute them cannot change. The idea would be that if a substance's nature is truly general, then it will hold not only for any substances that are part of the same world to which that substance happens to belong, but also for all times, that is, for all states of all such substances. Accordingly, if a ground were to change at some point between t_1 and t_2 , then the generality of the nature would be compromised. As a result, the generality of natures (or of the grounds that form them) might entail unchanging causal laws.

Second, Kant might have thought that the First Analogy in conjunction with his pre-Critical conception of grounds could support causal laws as well. For if the First Analogy can establish that a substance must be permanent and if it is clear from Kant's pre-Critical account of substances and grounds that a substance is constituted by grounds that are immutable, then it follows that the grounds of a substance must be both immutable and permanent. And if the grounds of a substance that bring about change are immutable and permanent, then it stands to reason that such grounds entail causal laws (insofar as immutable and permanent grounds would be the foundation for causal laws).

Finally, if grounds are to be understood as involving indeterminate activities rather than determinate states, then Kant might also have thought that grounds are not capable of change (at least not in the way that states are). For Kant understands change as a change of determinations, and if grounding activities are not determinations, then whatever might happen to them could not be represented as a change of determinations. The obvious objection to this line of thought is that it seems arbitrary to restrict change to change of determinations. Why not allow change of determining in addition to change of determinations? In light of the indeterminacy of the activity involved in grounding, it is unclear how such change is to be understood other than in terms of changes in the determinate states that the grounds bring about. And if a change of grounds can be understood only in terms of changes in the determinations they bring about, then we would be faced with the issue of what the identity conditions of grounds are. Given Kant's conception of grounds, one can see why it would not be tempting to view them as capable of change. If grounds cannot be directly perceived, then the primary basis for asserting that they are changing is removed as well, at least as long as a different

interpretation is available. For if one accepted changing grounds, then it would force one to undertake an impossible task, namely explaining why grounds changed in precisely this way at precisely this time, and whatever explanation one gave, it would, so it seems, have to be in terms of further grounds that either changed or did not, in which case no real improvement would have been made. Thus, instead of saying that a ground has changed, it can seem more attractive to assert that a different ground is active in bringing about different effects.

These three lines of thought and the idea that they would support the stronger, "causal law" reading of the Second Analogy can be illustrated further by Kant's understanding of the attractive and repulsive forces of bodies and the concept of mass associated with them. Spatial substances exercise their attractive and repulsive forces either in filling a determinate region of space or in the communication of motion, and they do so in accordance with their particular mass, which is part of the ground of the way in which these forces are exercised. Moreover, even if the effects of attractive and repulsive forces change depending on the changing relations between bodies, the attractive and repulsive forces themselves, along with the mass associated with them, do not change and are therefore the foundation for causal laws that describe the different kinds of effects that they have under different conditions. Further, drawing on the principle of continuity, it is clear that such forces produce their effects by means of a continuous and indeterminate activity. Since the mass of a body does not refer to other particular substances (or states thereof) at any particular time, it is clear that the nature of mass is general and, one might think, unchanging. The First Analogy of Experience (A182/B224) and the First Law of Mechanics (4:541) assert that not only substance, but also the quantity of substance does not change, which reinforces the idea that the mass of bodies must be immutable and could thus ground causal laws. Finally, insofar as the exercise of attractive and repulsive forces are indeterminate activities based on the immutable mass of spatial substance, Kant would never think that they could be altered. What can be altered are determinate states of motion or the determinate filling of a specific region of space, but not the mass nor the attractive and repulsive forces themselves.

Even if Kant's account of physics nicely illustrates why he might think that the grounds that bring about the determinate states of substances are immutable in such a way that they could provide the foundation for causal laws, it is important to note that it does not actually add any argumentative support to Kant's claim. Since his account of physics depends on the

metaphysical framework he is developing in the *Critique*, it cannot be used as an argument in support of that framework. It does, however, still play an important role for us insofar as it confirms our *interpretation* of Kant's model of causality and of the claims he actually makes in the Second Analogy.

Does this reconstruction of Kant's argument not prove too much? That is, would the argument, if successful, not establish that the sun will necessarily rise tomorrow, and would this not be objectionable given that we could imagine events that would prevent this from happening? In numerous passages Kant makes it clear that, unlike transcendental laws, empirical laws can be determined only on the basis of experience (and that regulative principles may be indispensable in discovering such laws). But if I see that one determination follows another in a certain set of circumstances, does that not immediately establish a causal law stating that the one determination follows another in such circumstances, such as the sun rising in the morning? The point to Kant's claim is that his complex notion of grounds and the model of causality based on it merely supply a *formal ontological framework* that must be filled in with empirical content. Accordingly, this framework entails only that whatever grounds and causal laws have held in the past will not change in the future. Thus, even if Kant were to establish the metaphysical necessity of causal laws for the determination of the changes that occur in the world, the epistemological question of ascertaining what grounds exist in the world has not been addressed at all. As Kant clearly indicates, this epistemological question can be answered only by consulting experience and can presumably never be established with absolute certainty, since one can never rule out the possibility that future evidence might require a revision in our understanding of what grounds there are in nature. For knowing that there are immutable grounds does not at all resolve the question of what empirical content they have.⁶⁴

⁶⁴ Even if this interpretation does accurately represent Kant's intentions, it is unclear that Kant's arguments can carry the weight of the strong reading of the Second Analogy at a metaphysical level. For even if the natures and grounds of substances *might* be general in the full-blooded sense that would be required to support causal laws, Kant has given no argument for thinking that they *must* be general in such a rich sense. Also, however plausible it may (or may not) seem to assert that permanent grounds cannot change, Kant has given no argument for the identity conditions of grounds that would definitively determine that permanent but changing grounds are impossible. Finally, even if one were to grant both that indeterminate activities cannot change in determinate ways and that we have no cognitive means by which indeterminate changes might be known, that still does not provide proper justification for claiming that indeterminate activities could not

By taking Kant's model of causality into account, we can thus make sense of important features of Kant's Second and Third Analogies that were otherwise puzzling. First, by understanding the reciprocal nature of mutual interaction, which was an important feature of Kant's pre-Critical model of causality, it became clearer why Kant thought that the Third Analogy should be restricted to spatial substances. Second, and perhaps more important, if grounds are understood as immutable, just as they were in Kant's pre-Critical period, we can understand why Kant would feel justified in claiming that causal laws are necessary in the Second Analogy. While this second result, in particular, might seem to constitute precisely the kind of refutation of Hume that many had hoped but failed to find in Kant, we see below in Chapter 6 that Kant's reply to Hume is more complex than is suggested by the idea that Kant intends to refute Hume.

The "Critical Turn" and the Limits of Metaphysics

In discussing Kant's Critical arguments for causality in Chapter 3 and his Critical model of causality (in this chapter), we have, at various points, appealed to Kant's pre-Critical account of causality. Although we had already sketched a speculative general interpretation of the nature of Kant's "Critical turn," we can now see that two specific aspects of his Critical model of causality that were taken over from his pre-Critical account have implications for our understanding of the "Critical turn." First, in the Third Analogy, Kant's reason for holding that a substance cannot determine its own place in time was ultimately derivative from his view expressed in the *Nova dilucidatio* that a substance cannot act on itself so as to change its own state. Second, Kant's model of causality is based not on Humean events, but rather on the concept of a ground that he first developed in his pre-Critical period in the context of an argument against Wolff and his followers.

There are, however, further significant parallels. To cite just one, Kant's pre-Critical argument that substances must stand in mutual interaction in order to form a real rather than merely ideal world resonates with several passages from the Critical period. Specifically, Kant discusses this

change nonetheless. In short, although it is now intelligible why Kant may have been tempted to think that his model of causality would entail that not only causality, but also causal laws would be required in order to have knowledge of objective succession, it is also clear that he does not explicitly develop clear-cut arguments that would in fact suffice to establish the stronger reading of the Second Analogy.

particular pre-Critical argument in a series of passages from the *Metaphysics Mrongovius* transcripts:

The form of the world is a real connection because it is a real whole. For if we have a multitude of substances, then these must also stand in connection with each other, otherwise they would be isolated. Isolated substances, however, never constitute a whole. If the substances are together, thus a whole, then they must also be a real whole. For were they ideal, then indeed they could be represented in thought as a whole, or the representations of them would constitute a whole; but the things in themselves would still not constitute a whole on that account. (29:851)

Later in the *Mrongovius* lectures Kant makes it clear that he is not simply reiterating his pre-Critical argument, but rather considering that argument from a Critical perspective:

But this proof holds only for the noumenal world. In the phenomenal world, we do not need it, for it is nothing in itself. Here, everything is in interaction due to space. The systems of occasional and pre-established harmony take place only in the sensible world.⁶⁵ For here the question is whether God so established the harmony in the beginning or in the course of time, and this presupposes time. Physical influx is called [so] in the cruder original sense, insofar as God now effects nothing in it; it is taken in a more subtle or derivative sense, when the possibility of the influence still arises from God as its prior origin. Physical influx occurs according to general laws, both systems of ideal connection do not. (29:868)

In addition to his endorsement of the *Inaugural Dissertation's* way of demarcating physical influx from occasionalism and pre-established harmony, this passage makes it quite clear that Kant recognized that his argument applies in different ways to the noumenal and phenomenal worlds. He seems to think that the argument continues to work for substances insofar as they are now understood as things in themselves, but – and here we encounter an important modification – that the argument is not even *needed* for phenomenal substances. In yet another passage from the *Metaphysics Mrongovius*, Kant reiterates this point: “Real influence presupposes a passion which, however, is at the same time action as well. I cannot at all derive from the concept of substance how this is possible. It is possible in the phenomenal world from the mere existence of substance in space. For space connects them all” (29:865). In other words, just as

⁶⁵ I interpret this statement to mean only that pre-established harmony and occasionalism have a chance of obtaining only in the sensible world. One cannot take this passage literally because pre-established harmony and occasionalism are incompatible and thus could not both “take place in the sensible world.”

Kant argued for the necessity of God in the *Nova dilucidatio* in order to make mutual interaction between substances possible, so Kant argues in this passage that we cannot understand just on the basis of our understanding of the concept of noumenal substances how they could interact with each other since through their concepts they are understood merely to be self-sufficient and thus not as standing in real relations with each other.

If Kant thus endorses one pre-Critical line of argument for the noumenal realm, these passages raise in an especially pressing way the question of how his pre-Critical views bear on the phenomenal realm. While our general conception of the "Critical turn" suggested that Kant would retain as much as possible of his pre-Critical view, incorporate it into an epistemological context, and eliminate only those elements that are flatly inconsistent with his Critical views, here we are repeatedly presented with assertions that might seem to entail that neither proof nor explanation of causal interaction is needed in the phenomenal world given the spatiality of phenomenal substances.

At the same time, such assertions require careful interpretation. First, what is actually asserted is not that a proof of causal interaction between phenomenal substances is not needed, but rather that a proof of causal interaction that is based on the idea that substances must interact to form a real whole is not needed. This particular proof is not needed, either because common sense naturally presupposes without question that everyday objects interact with each other or because another proof is already at hand, namely one based on the spatiality of substances.⁶⁶ Recall that Kant explicitly distinguished in the Third Analogy between spatial *communio* and causal *commercium*, arguing that the former can be cognized only by virtue of the latter. That is, these statements are consistent with the idea that the argument of the Third Analogy has already established causal interaction between phenomenal substances and that such causal interaction is apparent from the spatial community of such substances. Accordingly, these passages can be read as implicitly appealing to the argument of the Third Analogy by referring to the spatiality of phenomenal substances.

Second, the framework of the Analogies of Experience is flatly inconsistent with the idea that one could dispense with the need for any ground of causal interaction. As became clear in Kant's argument in the Second

⁶⁶ The "common sense" interpretation is suggested by Karl Ameriks in "Kant on Science and Common Knowledge," in *Kant and the Sciences*, pp. 31–52.

Analogy, what made causality necessary for knowledge of objective succession was the necessity of a ground that would determine the successive states of an object. For without such a ground, there would be no reason for the second state to follow the state that the object was already in. The same sort of idea was present in the Third Analogy as well. Since the simultaneity of states of two substances is not immediately given along with the mere existence of substances, mutual interaction is required to ground that relation. In this way, one can reconcile the primary content of Kant's statements in the *Metaphysics Mrongovius* with Kant's explicit argument in the Analogies.

However, if the argument of the Analogies presupposes that *temporal* relations need a ground, does it not follow that *causal* relations would need a ground as well? And if so, what could that ground be? To see the driving force behind this question, recall Kant's position in the *Nova dilucidatio*. The principle of succession claimed that causal interaction is necessary for change to occur in substances, while the principle of coexistence argued that causal interaction is possible only by means of the divine schema of the understanding. While the arguments for both principles depend on the need for grounds that would posit the determinations of substances, the argument for the principle of coexistence in particular turns on the idea that only God could relate substances to each other in such a way that they could interact causally. This is now relevant as follows. The principle of coexistence was necessary because the causal relation shown to be necessary by the principle of succession required a ground and that ground could not be contained entirely in the substances so related (or in their "mere existence"). As a result, the argument concluded that God must be the ground of that relation. Viewed in this light, we can see that a similar question arises about the ground of the causal relation between phenomenal substances.

If the question about the ground of the causal relations between substances is already clearly both posed and answered in the pre-Critical period, could Kant not simply appropriate his pre-Critical answer and could we not view this particular issue as representing yet one more continuity between his pre-Critical and Critical periods? Given that Kant provides detailed objections to the three traditional theistic proofs and explicitly argues that God is not an object of possible experience, one might think that he could not invoke God as the ground of causal relations in the Critical period and might infer that continuity on the issue of causality across the board cannot be maintained. However, one need not immediately infer such a complete and radical break. For the fact that we cannot

know that God exists does not imply that God could not *be* the ground of the causal relations between substances. Rather, in line with the (theoretical) agnosticism about things in themselves required by Transcendental Idealism, one can simply admit that we cannot know what the ultimate ground of causal interaction is, even if we may form beliefs that stem from the presuppositions of practical reason (since we may believe that God must exist in order to coordinate virtue and happiness in conformity with the highest good).

A more complicated interpretation is intimated, however, when Kant makes the following remark in the *Metaphysics Mrongovius* transcripts. "The concept of space achieves in the sensible world what divine omnipresence does in the noumenal world, and one can thus call it [space] a phenomenon of divine omnipresence" (29:866).⁶⁷ While it is possible that God is the immediate ground of the causal relations between phenomenal substances, this passage suggests that God might be the mediate ground of such relations, with the noumenal world functioning as an intermediate ground between God and phenomenal substances. According to this passage, God is the immediate ground of the causal relations between noumenal substances, to which he is omnipresent, and the noumenal substances, in turn, ground causal relations between phenomenal substances, for which reason space is the counterpart in the phenomenal realm to divine omnipresence in the noumenal realm.

Granted, the textual evidence in favor of such a reading is limited, but perhaps that is exactly what one ought to expect given Kant's commitment to Transcendental Idealism.⁶⁸ That is, although it is clear that there must be sufficient grounds (or, as he sometimes puts it, conditions) for the various things that we know to be the case in the phenomenal realm, it does not immediately follow that these grounds are also entities in the phenomenal realm and knowable. Sometimes they are, as is the case when causality and mutual interaction function as grounds of temporal relations. However, sometimes they are not, as happens with the ground of causal relations (and with the existence of substances). Accordingly, it is impossible for us to know whether God is a mediate or an immediate ground of the causal relations between phenomenal substances (even

⁶⁷ Kant also strongly suggests this picture in the *Critique of Practical Reason* (5:102).

⁶⁸ We have reason to return to what I call the "grounding thesis" in Chapter 5. Here we can simply note in advance that considerable textual support can be adduced for thinking of noumenal substances as the ground of phenomenal substances.

if we can know that there must be such a ground). In fact, one cannot even establish the existence of God, given that one cannot rule out the possibility that noumenal substances alone could function as the grounds of such relations. That Kant sees this connection between Transcendental Idealism and the unknowability of the ground of causal interaction is suggested by his claim in the second edition version of the Paralogisms that the question of “how in general a community of substances is possible . . . lies outside the field of possible experience” (B428).

CONCLUSION

Commentators on Kant’s Second Analogy of Experience have assumed that Kant accepts the fundamental features of Hume’s model of causality, namely that one event causes another event, in order to refute Hume’s claim that there is no necessary connection between cause and effect and, in some, though not all, cases, that there is no justification for assuming that there are causal laws (that would necessarily hold in the future just as they have in the past). By attending to Kant’s notion of mutual interaction in the Third Analogy and to various nonargumentative passages in the Second Analogy, we have seen that Kant neither does nor can accept Hume’s event-event model of causality. Rather, with the benefit of an awareness of Kant’s pre-Critical account of causality, we saw that Kant’s texts and arguments commit him to a model of causality that involves substances exercising their causal powers so as to determine each other’s states.

While accepting causal powers is quite traditional in the context of early modern philosophy, Kant develops his notion of causal powers in detail and incorporates it into his broader account of (phenomenal) metaphysics. More specifically, the exercise of a causal power or, as Kant sometimes puts it, “the causality of the cause,” is not a determinate event, but rather an asymmetrical and indeterminate activity that brings about passive determinations in a distinct substance. Since one might charge that Kant’s distinctive notion of activity is obscure, it became imperative to see how to clarify it further. While one could see that it is consistent with the notion of causality Kant derives in central argumentative passages in the *Critique* (such as the Metaphysical Deduction and the Schematism) and with particular instances of causality that Kant explicitly develops in the *Metaphysical Foundations* (such as the notions of attractive and repulsive forces in Newtonian physics), it also became apparent that the notion is best clarified by means of the notion of activity that Kant invokes in his

distinctive account of self-consciousness, an account that also offers him other significant advantages over Hume.

To fill out our understanding of Kant's model of causality further, it was then important to consider both the specific ways in which this model relates to his pre-Critical account and the consequences that this model entails for his arguments in the Second and Third Analogies. Regarding the former issue, we discovered that despite considerable continuity between his pre-Critical and the Critical periods both in general and on a host of specific points, with respect to causal grounding Kant comes to think that some grounds are phenomenal and knowable, whereas others are noumenal and unknowable, which represents an interesting way of reconciling the major elements of his pre-Critical account with Transcendental Idealism. Regarding the latter issue, we came to see why Kant would restrict the conclusion of the Third Analogy to spatial substances and, perhaps more importantly, how Kant might have thought that the Second Analogy is in a position to justify not only the necessity of causality, but also causal laws.

PART THREE

CAUSALITY AND CONSEQUENCES

If we have provided an adequate description of Kant's Critical account of causality, what follows? More specifically, is his general model of causality as described in Part II connected in systematic ways to any of his other central Critical doctrines? Does his general model of causality have any implications for what his reply to Hume might be? The final two chapters develop answers to these questions. Chapter 5 ("The Metaphysics of Freedom") argues that Kant's metaphysical account of natural causality has implications for understanding the metaphysics of freedom, though his account of freedom also helps to clarify several aspects of his general model of natural causality as well. Chapter 6 ("Kant's Reply to Hume: Historical and Contemporary Considerations") considers what Kant's Critical account of causality reveals about the nature of his ultimate reply to Hume and several ways in which Kant's account can be relevant in contemporary philosophical contexts.

The Metaphysics of Freedom

INTRODUCTION

As we have seen in Chapters 3 and 4, the issue of causality occupies a prominent place within the Transcendental Analytic's Analogies of Experience. However, causality is also a major topic of the Transcendental Dialectic. While the Paralogisms and Ideal of Pure Reason touch on the issue in various ways (e.g., in addressing interaction between mind and body and in discussing the cosmological argument for the existence of God, respectively), it is one of Kant's central concerns in the Antinomies of Pure Reason, in the Third Antinomy in particular, where it arises in the context of the problem of free will and determinism. Thus, we continue our discussion of Kant's views on causality by considering his treatment of the problem of free will and determinism.

The problem of free will and determinism is, of course, a notoriously difficult one, and our discussion of it in this chapter in no way attempts to resolve all of its complexities, not even all of those that Kant himself draws our attention to throughout the *Critique of Pure Reason*. Rather, the primary aim of this chapter is merely to come to a better understanding of how the general model of causality that was described and argued for in Chapters 3 and 4 can both clarify and be clarified by Kant's views on freedom. In particular, we see that in solving certain aspects of the problem of free will and determinism, Kant appeals to many of the same notions that he invoked in his general model of causality. Thus, just as natural causality is to be understood not in terms of determinate events alone, but rather in terms of a substance determining the states of another substance by means of an exercise of causal powers in accordance with

its nature, so freedom is to be understood not in terms of desires alone (which are simply one kind of determinate mental event), but rather in terms of an agent or, more metaphysically, substance determining its actions according to its character. As a result of the structural similarities between these two accounts of causality and freedom, we can use the features of the one to clarify initially obscure aspects of the other and vice versa.

To undertake a comparison of Kant's general model of causality and his account of freedom, it is necessary to begin, in a first section, by describing the core problem of free will and determinism as Kant sees it. To this end, we present the main elements of the Third Antinomy: the Thesis and Antithesis arguments and its Resolution. This presentation indicates the way in which Transcendental Realism is essential to the Thesis and Antithesis arguments of the Third Antinomy and highlights the fundamental features of Kant's solution to the problem. For Transcendental Idealism draws a distinction between things in themselves and appearances that makes it possible for him to hold that determinism is true at the level of appearances, while freedom may be possible for things in themselves.

Since Kant admits that invoking only general features of Transcendental Idealism provides an explanation that "must appear extremely subtle and obscure" (A₅₃₇/B₅₆₅), it is necessary to consider two more specific aspects of Transcendental Idealism in a second section. One is the much-discussed question of whether Transcendental Idealism is an exclusively epistemological doctrine or whether it also has an ontological dimension. Rather than trying to decide between these two interpretations directly – a task that is complicated by the fact that the relevant textual evidence as well as various purely philosophical considerations pull in both directions – I suggest that Kant is not sufficiently clear about these different versions of Transcendental Idealism and that this lack of clarity is caused, at least in part, by his failure to understand precisely enough what kind of relation holds between the level of ultimate reality and that of phenomenal bodies.

The second issue concerns the meaning of Kant's oft-repeated claim that things in themselves are required to ground appearances. Clarifying this claim reveals that Kant's resolution of the modal conflict implicit in the problem of freedom and determinism hinges primarily on the idea that which laws of nature are actual and thus what is determined as necessary in accordance with them may depend on our freely made and thus contingent choices of our own natures or characters. Although

these choices are also responsible for the creation of what one might call “personal facts” – that a certain event occurring in the world is also to be viewed as *my action* – the creation of these facts by itself does not resolve the modal conflict between the necessity of determinism and the contingency of free will. Still, since this “personal” dimension is frequently neglected in discussions of the problem of free will and determinism, it is not an unimportant result that Kant can provide a plausible explanation of this dimension of our actions.

In the third and final section, we can then turn to comparing Kant’s general model of causality and his account of freedom. Such a comparison reveals that Kant’s account of freedom sheds light on several aspects of his general model of causality. First, establishing that Kant’s views on these two different topics are in fact consistent turns out not to be as trivial as one might think insofar as not all accounts of causation are consistent with Kant’s distinctive solution to the problem of free will and determinism. Specifically, since Kant’s solution involves the selection of laws of nature, which depend, in turn, on the natures of things, it is inconsistent with models of causality that, unlike Kant’s, do not involve natures. Establishing consistency thus helps to clarify the conception of laws of nature that is bound up with Kant’s general model of causality.

Second, Kant’s account of freedom can help us to make sense of several passages that seem to contradict Kant’s general model of causality by asserting that causes must be events (rather than indeterminate activities). The central point here is that one of the crucial aspects of Kant’s solution to the problem of free will and determinism – Transcendental Idealism’s distinction between what is fully determinate and what must be, to some degree, indeterminate – allows one to understand not just appearances in general, but also the activity of causal determination in particular as essentially indeterminate. Kant’s reflections on freedom thus clarify how the notion of determination that is a part of his general model of causality can be understood properly in the context of Transcendental Idealism.

Third, by appreciating the relevance of Transcendental Idealism to Kant’s account of causality one might attempt to reinterpret the entities that stand in mutual interaction on Kant’s general model of causation as noumenal rather than phenomenal substances, especially in light of difficulties that one might have with his conception of phenomenal substance in general. However, not only does one not need to dispense with phenomenal substances in this way – the difficulties just mentioned can be adequately resolved by paying close attention to Kant’s different

conceptions of substance – one also *cannot* dispense with them. For, as the primary bearers of changing phenomenal properties, they must be temporal, a function that noumena cannot take over in light of the requirements they are subject to in allowing for the possibility of freedom. Accordingly, Kant's views on freedom help to clarify several aspects of his general model of causality.

Yet clarification goes in the other direction as well. First, Kant's general model of causality supplies several of the basic concepts that he wants to use in his account of freedom. In particular, he draws on the concepts of substance, determination, activity, and natures to describe how freedom might be possible, though he recognizes that these concepts must be modified in certain respects to be appropriate to their new domain of application. As a result, Kant can think of freedom as a noumenal substance's determination of its character by means of the activity of its will. Second, he can also draw on his general model of causality to solve several problems that traditionally arise in discussions of free will and determinism. For example, to stop the regress that arises if one attempts to explain, solely on the basis of desires, why a certain action was performed, Kant can appeal to the same notion of a determining ground that was indispensable to his general model of causality. Also, in response to objections that arise if one tries to locate freedom in desires, which are themselves nothing other than a certain kind of event, Kant can understand free will in terms of an agent, which is simply a substance endowed with a special kind of causality in virtue of the rational choice that it can make according to its conception of the good. Kant's general model of causality can thus serve as an important resource for clarifying his account of freedom. As a result, Kant's model of natural causality and his account of freedom illuminate each other in several significant ways.

THE THIRD ANTINOMY

Kant presents the problem of free will and determinism as one of four Antinomies that arise with respect to the concept of the world. While the First and Second Antinomies concern the spatio-temporal size of the world (whether there is or is not a beginning of the world in time and an outer boundary to it in space) and the constitution of its parts (whether simple or infinitely divisible) and the Fourth Antinomy focuses on modality (whether or not any being in the world is absolutely necessary), the Third Antinomy addresses the problem of free will and

determinism.¹ Kant presents this problem in a format similar to that of the other Antinomies, where a given issue is developed by means of a Thesis and Antithesis – which, on the presupposition of Transcendental Realism, argue for contradictory positions – and a Resolution – which shows how Transcendental Idealism alone is in a position to dissolve the conflict generated by the Thesis and Antithesis arguments. Thus, in the case of the Third Antinomy, the Thesis and Antithesis argue for and against freedom in the world, while the Resolution contends that only Transcendental Idealism can keep this conflict from arising. Accordingly, to grasp the basic structure of Kant's understanding of the issue of free will and determinism, we must start with the arguments he presents in the Thesis and Antithesis for freedom and determinism and then turn to the most general features of his Resolution.

The Thesis and Antithesis Arguments

The Thesis and Antithesis arguments are based on the assumption of Transcendental Realism. Transcendental Realism is the view that identifies appearances and things in themselves, that is, it maintains that what appears to us in space and time is also what really exists independently of us and that what exists independently of us also appears to us in (or is at least subject to the conditions of) space and time. If Transcendental Realism takes appearances and things in themselves to be identical, any principles that apply to the one must also apply to the other so that, e.g., ontological principles must apply to whatever objects appear to us in space and time. In particular, since things in themselves have determinations only if the grounds exist that would be sufficient to posit those determinations, the transcendental realist must also be committed to the view that the sufficient conditions for appearances must be satisfied for them to exist. Also, since things in themselves must be fully determinate with respect to every possible predicate pair (A or $\sim A$), appearances must be fully determinate as well. As we see below in more detail, Kant holds that these two particular aspects of Transcendental Realism play an important role in generating the contradiction that arises in the Antinomies.

The Thesis of the Third Antinomy asserts that appearances require freedom, since causality in accordance with the laws of nature is insufficient to account for appearances. The basic idea behind the argument

¹ For a general account of central sections of the Antinomies, see my "Kant's Antinomies: Sections 3–8," in *Kooperativer Kommentar zu Kants Kritik der reinen Vernunft*, ed. G. Mohr and M. Willaschek (Berlin: Akademie Verlag, 1998), pp. 445–462.

is as follows: Since everything that happens in the world requires a sufficient reason, but caused causes are not sufficient, there must be uncaused causes or freedom. As is the case for all of the Antinomies, the argument proceeds by a *reductio ad absurdum*, which can be reconstructed as follows:

- P1 Suppose there were no freedom and that all causality occurs in accordance with the laws of nature.
- P2 If all causality occurs in accordance with the laws of nature, then, for every event that happens, there must be a previous state from which it follows in accordance with the laws of nature.
- C1 For every event that happens, there must be a previous state from which it follows in accordance with the laws of nature. (from P1 and P2)
- P3 If the state from which an event follows in accordance with the laws of nature had existed forever (i.e., were not an event, and thus did not come into existence and require a previous state from which it followed), then it could not have brought forth the event that is supposed to follow from it in accordance with the laws of nature. (Kant remarks: “since if it [the state] had been at every time, then its consequence could not have just arisen, but would always have been” (A444/B472).)
- C2 For any event that happens, the state from which it follows in accordance with the laws of nature is itself an event. (from C1 and P3)
- P4 If every event presupposes a preceding event from which it follows in accordance with the laws of nature, then there is never an absolutely first causal event and thus “no completeness of the series [of events] on the side of the causes descending from one another” (A446/B474).
- C3 There is no completeness of causes for any event. (from C2 and P4)
- P5 If there is no completeness of causes for an event, then that event happens “without a cause sufficiently determined a priori” (A446/B474).
- C4 Every event happens “without a cause sufficiently determined a priori.” (from C3 and P5)
- P6 C4 is false; no event happens “without a cause sufficiently determined a priori.”
- C5 P1 is false; there must be a kind of causality distinct from causality in accordance with the laws of nature, that is, one that occurs without its cause being determined by another, previous cause – “an **absolute** causal **spontaneity** beginning **from itself**” (A446/B474) called transcendental freedom. (from C4 and P6)

For current purposes, the crucial steps in this argument are P5 and P6.² For their plausibility depends on an interpretation of the principle of

² Kant treats P2 as an analytic truth insofar as what it means for events in time to be caused according to the laws of nature is for them to follow from a state according to a rule. Like P2, P4 is an analytic truth. The status of P3 might be less clear. However, as Kant points out in his remarks on the Thesis argument, what is essential to P3 is not the temporal

sufficient reason that is based on Transcendental Realism. Specifically, they assert not just that no event can happen without a sufficient cause, but also that the cause of an event can be a sufficient reason for that event only if there is a completeness in its causes (by which it is thus “sufficiently determined”). That is, they require not that every event be caused immediately by an uncaused cause, but rather that every event be caused by a cause that either is itself uncaused or else ultimately terminates in a cause that is uncaused, since any caused cause contains conditions *all of which* must be satisfied for the cause to be truly efficacious or sufficient.³

P5 and P6 rely on Transcendental Realism because the demand for completeness of conditions of an event (i.e., for taking the Principle of Sufficient Reason “in its unlimited universality” A446/B474) is justified only if the complete set of conditions is fully determinate, and we are justified in assuming that the complete set of conditions is fully determinate only if those conditions are satisfied by independently existing things, given that independently existing things are fully what they are in light of the grounds that posit their determinations. In other words, P5 and P6 involve a particularly strong reading of the Principle of Sufficient Reason by requiring that only a complete set of conditions will constitute a sufficient reason of an event, and such a reading of the Principle of Sufficient Reason is warranted by the assumption of Transcendental Realism, since Transcendental Realism’s assertion that appearances are fully real and thus *completely* determinate entities makes it possible to

nature of the series of causes, but rather the causal dependency. Kant draws the distinction as follows: “For here we are talking of an absolute beginning, not as far as time is concerned, but as far as causality is concerned” (A450/B478). And after illustrating this point with an example, he speaks of an event “which indeed follows *upon* [folgt *auf*] that series [of causally unrelated, but temporally prior events], but does not follow *from* it [aber *daraus* nicht *erfolgt*]” (emphasis added).

³ Jonathan Bennett, *Kant’s Dialectic* (New York: Cambridge University Press, 1974), calls our attention to the potentially odd-sounding phrase “a cause sufficiently determined a priori,” but dismisses the suggestion made above on the grounds that the principle of sufficient reason pertains to *effects*, not *causes* (p. 185). However, this diagnosis forces Bennett to admit, “I do not know what that [phrase] means” (p. 185), to acknowledge, “I do not have an interpretation of the Thesis-argument of the third antinomy” (p. 186), and to “confess defeat” (p. 187). What Kant is asserting, however, is simply the idea that if a cause is to be a sufficient reason for its effect, then there must be a sufficient reason for the cause as well. (Paul Guyer seems to recognize this point in *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987), p. 412.) Bennett may reject such an interpretation on purely philosophical grounds, but it does have considerable textual support, especially if one takes seriously the idea that a transcendental realist in cosmology will insist that we consider the world *as a totality*.

require that *all* of the conditions for any event we might experience must be given.

The Antithesis of the Third Antinomy, by contrast, asserts that there is no freedom in the world, since freedom is alleged to be incompatible with the requirements of the natural world and thus turns out to be nothing more than “an empty thought-entity” (A445/B473). As was the case with the Thesis, the argument of the Antithesis begins by provisionally assuming what is ultimately to be rejected as inconsistent, which, in this case, is freedom, and proceed from there:

- P1 Suppose there were freedom, that is, a spontaneous (or uncaused) cause of the (absolute) beginning of a series of events.
- P2 If a series of events was caused by a free or spontaneous cause, the spontaneous cause would not be caused by any previous state (or event) to be the cause of that series of events, that is, “the determination of this spontaneity itself to produce the series . . . will begin absolutely” (A445/B473).
- C1 A spontaneous cause is not caused by a previous state to be the cause of the series of events it causes. (from P1 and P2)
- P3 For everything (or for every event) that happens, there must be a previous state from which it follows in accordance with the laws of nature.
- C2 If a spontaneous cause happens (or begins to act), there must be a previous state from which it follows (causally) in accordance with the laws of nature. (from P3)
- C3 A spontaneous cause is caused by a previous state in accordance with the laws of nature. (from P1 and C2)
- C4 C1 and C3 are contradictory. P1 must be false; there can be no freedom in the world.

The argument of the Antithesis should not be especially controversial in this context. For the central assumption of the Antithesis argument – P3’s assumption that causality is required as a sufficient reason of events occurring in time – is a causal principle that would seem to be unproblematic given the assumption of Transcendental Realism. The primary difference between the Thesis and Antithesis arguments is that the Thesis argument takes this causal principle “in its unlimited universality,” whereas the Antithesis argument applies this principle to the case of spontaneous causality or freedom.

One might object that the Antithesis argument falsely assumes (in the inference from P3 to C2) that a spontaneous cause must be an event in time, which then generates the contradiction that there must be a temporally prior cause of a spontaneous cause. However, this objection is incompatible with Transcendental Realism. Because the transcendental realist identifies appearances, which are temporal, with things in themselves,

spontaneous causes must be a part of the same series of causes that is formed by natural causes, that is, they must be temporally determinate events just as natural causes are. If they were not, then, *pace* Transcendental Realism, there would be a distinction between things in themselves and appearances. Kant makes this point explicit in the preamble to his resolution of the dynamic-transcendental ideas when he draws a distinction between what lies within a series (e.g., what is sensible) and what lies outside the series (the intelligible), a distinction that is available only to the transcendental idealist (A530/B558).

This point, the consequences of which are discussed further below, reveals how the reconstruction presented above represents a departure from one prevalent line of interpretation of the Antithesis argument found in the secondary literature. Henry Allison and Hud Hudson, whose views are otherwise quite different, both suggest that Kant's argument in the Antithesis is essentially that of the Second Analogy.⁴ Paul Guyer, who differs from Allison and Hudson in rejecting rather than recommending the argument, likewise understands it as "explicitly epistemological" and as having the First and Second Analogies at its foundation.⁵ Presumably, these commentators took the fact that Kant refers to spontaneous causality as being inconsistent with the "unity of experience" (A447/B475) as evidence in favor of this line of interpretation (in addition to the fact that both are causal principles).

However, one should immediately note that the textual evidence does not unequivocally support such a reading. For it is possible to read Kant's reference to the unity of experience not as a premise in his argument, but rather as what follows as a conclusion. "Thus transcendental freedom is contrary to the causal law, and is a combination between the successive states of efficient causes in accordance with which no unity of experience is possible, which thus cannot be encountered in any experience, and hence is an empty thought-entity" (A445-447/B473-475). That is, on the present reading it is because spontaneous causality is inconsistent with natural causality that we cannot meet with it in experience (rather than vice versa). Accordingly, although we can form a concept of spontaneous

⁴ Henry Allison, *Kant's Transcendental Idealism: An Interpretation and Defense* (New Haven: Yale University Press, 1982), remarks that the argument of the Antithesis "rests almost entirely on the argument of the Analytic, specifically, the First and Second Analogies" (p. 312). See Hud Hudson, *Kant's Compatibilism* (Ithaca: Cornell University Press, 1994). This line of interpretation has also been developed recently by Wolfgang Malzkorn, *Kants Kosmologie-Kritik* (Berlin: De Gruyter, 1999), p. 214, n. 293.

⁵ *Kant and the Claims of Knowledge*, pp. 411-412.

causality, that concept can never be determined to have a referent and thus remains an “empty thought-entity.”

Moreover, the following systematic or architectonic point is even more problematic for the reading espoused by Allison, Hudson, and Guyer. Since both the Thesis and the Antithesis arguments must presuppose Transcendental Realism and since Transcendental Realism is inconsistent with the Second Analogy of Experience – regardless of whether one interprets the Second Analogy as presupposing Transcendental Idealism at the start – the Thesis and Antithesis arguments cannot be based on the Second Analogy. Put the other way around, if the arguments of the Third Antinomy did not presuppose Transcendental Realism, then Transcendental Idealism could not be required as a necessary part of any resolution of the contradiction that is generated by its Thesis and Antithesis arguments.⁶ And if the arguments of the Third Antinomy do presuppose Transcendental Realism, then they cannot be identical to arguments, such as that of the Second Analogy, that require Transcendental Idealism.

If the arguments Kant presents in the Thesis and Antithesis are reconstructed in this manner, two points about the basic structure of Kant’s understanding of the problem of free will and determinism are clear at this point. First, the Thesis and the Antithesis do explicitly contradict each other insofar as one asserts, whereas the other denies, spontaneous causality or freedom in the world. Second, since both arguments are based on the assumption of Transcendental Realism, it is clear that Transcendental Realism must be rejected to avoid the contradiction and resolve the Antinomy. To see how rejecting Transcendental Realism can remove

⁶ While it is obviously true that not every premise in the Thesis or Antithesis arguments must admit of a transcendently realistic interpretation, it is clear that the causal principles they invoke are their crucial premises and thus must involve Transcendental Realism in some way. Henry Allison, *Kant’s Theory of Freedom* (New York: Cambridge University Press, 1990), objects that “even transcendental realism affirms the validity of the analogies within experience (although it cannot account for this validity). The point here is that the issue between transcendental realism and transcendental idealism does not arise at the empirical level but only when, as in the final step of the antithesis argument, these principles are extended beyond the limits of possible experience” (p. 21). In response, I would argue that Allison’s reading threatens to make the argument invalid, since the sense in which the premises are to be understood differs from the sense in which the conclusion is intended, whereas the reading proposed above has it that the argument is valid insofar as the causal principles that are at the heart of the Antithesis argument are to be taken as applying to objects *simpliciter* (since the transcendental realist has no distinction between appearances and things in themselves with which to understand the causal principles properly in the first place).

the explicit contradiction between freedom and determinism, we must turn to Kant's Resolution of the Third Antinomy.

The Resolution

While Kant's fullest solution to the problem of free will and determinism involves a variety of considerations at a number of different levels of generality, we can start by describing the basic framework that he adopts to resolve the conflict between the Thesis and Antithesis of the Third Antinomy. In the "Resolution of the cosmological idea of the totality of the derivation of occurrences in the world from their causes" (A532/B560) – after asserting that all causality occurs either according to nature or from freedom (A536/B564) – Kant sets up the question raised by the Third Antinomy as follows:

Thus the difficulty we encounter in the question about nature and freedom is only whether freedom is possible anywhere at all, and if it is, whether it can exist together with the universality of the natural law of causality, hence whether it is a correct disjunctive proposition that every effect in the world must arise **either** from nature **or** from freedom, or whether instead **both**, *each in a different relation*, might be able to take place simultaneously in one and the same occurrence. The correctness of the principle of the thoroughgoing connection of all occurrences in the world of sense according to invariable natural laws is already confirmed. . . . Thus the only question is whether, despite this, in regard to the very same effect that is determined by nature, freedom might not also take place, or is this entirely excluded through that inviolable rule? (A536/B564, italics added)

In short, if natural causality and freedom are inconsistent, then freedom must be rejected, since one cannot renounce natural causality, given that "the thoroughgoing connection of all appearances in one context of nature is an inexorable law" (A537/B565). However, if there were a way of understanding natural causality and freedom as consistent, that is, a way of rendering intelligible how both natural causality and freedom, "each in a different relation," could take place with respect to a given event, then at least one basic aspect of the problem of free will and determinism (in the guise of natural causality) could be resolved. Accordingly, the crucial issue for Kant's resolution lies in understanding what difference in what relations might allow natural causality and freedom to be consistent.

Unsurprisingly, the crucial issue turns out to rest squarely on Transcendental Idealism. As Kant clearly states: "if appearances are things in themselves, then freedom cannot be saved" (A536/B564). Accordingly,

it is necessary that one distinguish between appearances and things in themselves. What is this distinction and how does it help? Kant's preliminary explanation is as follows:

If, by contrast, appearances do not count for any more than they are in fact, namely, not for things in themselves but only for mere representations connected in accordance with empirical laws, then they themselves must have grounds that are not appearances. Such an intelligible cause, however, will not be determined in its causality by appearances, even though its effects appear and so can be determined through other appearances. Thus the intelligible cause, with its causality, is outside the series; its effects, on the contrary, are encountered in the series of empirical conditions. The effect can therefore be regarded as free in regard to its intelligible cause, and yet simultaneously, in regard to appearances, as their result according to the necessity of nature; this is a distinction which, if it is presented in general and entirely abstractly, must appear extremely subtle and obscure, but in its application it will be enlightening. (A537/B565)

While Kant is surely right that this solution to the problem of free will and determinism can appear "extremely subtle and obscure," we can still explain, at least at a certain level of generality, several central features of Transcendental Idealism and how they offer a solution to both the conflict that arises in the Antinomies and the problem of free will and determinism.

The basic idea underlying Transcendental Idealism is Kant's distinction between appearances – "mere representations" of objects that are given to us through our purely subjective forms of intuition and have "no existence outside our thoughts" (A491/B519) – and things (as they are) in themselves – things that are not given to us in intuition and that exist independently of us (more specifically, independently of our mind's forms of intuition). From the fact that things in themselves are not given to us, it follows, so Kant claims, (1) that they cannot be spatio-temporal (since space and time are only subjective forms through which objects are given to us) and (2) that we cannot have substantive (e.g., synthetic) knowledge of them (since objects must be given to us if we are to have such knowledge of them).⁷ By contrast, since appearances are defined in terms of being given to us through our forms of intuition, it is clear (1) that they are spatio-temporal (since space and time are our forms of intuition), and (2) we are in a position to have knowledge of them (since

⁷ This brief presentation of Transcendental Idealism should in no way be taken to suggest that Kant's *arguments* for these claims are simple, straightforward, or uncontroversial. See, for example, my "Transcendental Idealism and the Categories," *History of Philosophy Quarterly* 19 (2002): 191–215.

they satisfy the “formal” conditions of knowledge). While the distinction between things in themselves and appearances involves a series of further complications (to which we have reason to return below), this characterization of some of its basic features allows us to understand Kant’s most general explanation of his resolution of the Antinomies.

In section seven of the Antinomy of Pure Reason, where Kant presents the “Critical decision of the cosmological conflict of reason with itself,” he explicitly describes the fundamental difficulty of all of the Antinomies as follows: “The entire antinomy of pure reason rests on this dialectical argument: If the conditioned is given, then the whole series of all conditions for it is also given; now objects of the senses are given as conditioned; consequently, etc.” (A497/B525). Thus, in the case of each Antinomy some conditioned object is given in the world – whether it be the spatio-temporal magnitude of the world, the constitution of its objects, or the effects of causality in the world – and the Thesis and Antithesis present arguments showing that the series of conditions of the object that is given both can and cannot be complete in each instance.

Kant exposes the error of this dialectical argument by distinguishing between the ways in which the conditioned-condition relationship applies to things in themselves and appearances. “If the conditioned as well as its condition are things in themselves, then when the first is given, . . . the latter is thereby really already given along with it” (A498/B526). By contrast, “if I am dealing with appearances, . . . then I cannot say with the same meaning that if the conditioned is given, then all the conditions (as appearances) for it are also given. . . . For the **appearances**, in their apprehension, are themselves nothing other than an empirical synthesis (in space and time) and thus are given only **in this synthesis**” (A498–499/B527). Thus, for appearances, “if the conditioned is given, then through it a regress in the series of all conditions for it is **given** to us **as a problem**” (A497–498/B526). In short, whenever a conditioned object is given, then its conditions must also be given *if* the conditioned is a thing in itself, but *not* if it is an appearance. In the case of the latter, one must rather *search* for its conditions.

Why, one might ask, does the condition-conditioned relationship apply to things in themselves and appearances in such different ways? The difference arises, at least in part, from the fact that things in themselves are completely determinate, whereas appearances are not. As we saw above, things in themselves are completely determinate in the sense that for every pair of contradictory predicates one of them must be truly ascribed to each thing in itself (A571–576/B599–604). Appearances, by contrast,

are not completely determinate, since they, as “mere representations,” must be apprehended by us and undergo empirical synthesis (in space and time as well as conceptually) to be determined.⁸ But since things in themselves are completely determinate and the principle of sufficient reason applies to each determination, the reason or condition of every determination must also be determinately given. Otherwise, these things would not have the determinations they do. Since appearances can be indeterminate (and thus our knowledge of the sensible world incomplete), the conditions that ground the determinations of an appearance can be indeterminate, that is, may not be given to our forms of intuition. But since appearances are defined in terms of what is given to us, the condition-conditioned relationship entails that though we must allow that a condition for some appearance is not given, we must still proceed as if it existed, and reason – the faculty that strives for the conditions of all conditioned objects – must undertake the search for it and any other outstanding conditions.

This difference regarding determinacy between things in themselves and appearances is crucial not simply to Kant’s metaphysical reflections on the Antinomies in general, but also to their particular resolutions. In the case of the First Antinomy, for example, the contradiction generated by the Thesis and Antithesis is that the world can be proved to be both infinite and finite in its spatio-temporal magnitude. Its resolution, which distinguishes the spatio-temporal world of the senses from the non-spatio-temporal world of things in themselves, asserts that while the world of things in themselves must have a determinate (and, as it turns out, finite) size (29:856), the world of appearances is of indeterminate magnitude since we must always continue to look for (i.e., seek out intuitions of) further regions in space and earlier moments in time (A518/B546). The resolution of the Second Antinomy is similar (with things in themselves being determined as simple), except that spatio-temporal objects can be shown to be infinitely rather than merely indeterminately divisible. For, by contrast with the First Antinomy, all the intuitions we need to determine the constitution of an object are already

⁸ Kant defines an appearance as “the undetermined object of empirical intuition” (A20/B34), which contrasts with other possible definitions of appearances, for example, as being a representation that requires an object distinct from it of which it is the appearance. That Kant defines appearances in terms of being undetermined allows him to connect appearances with his account of reason as the faculty that searches for the unconditioned condition of conditioned objects, since undetermined objects will require determining grounds as their conditions.

given and it is merely the process of division that can never be completed (A523–524/B551–552).⁹

Given this understanding of the fundamental features of Transcendental Idealism and of the way that Kant uses it to parse the condition-conditioned relationship, and a brief sense of how he invokes it in resolving the First and Second Antinomies, we are now in a position to explain the basic structure of his resolution of the Third Antinomy. As we saw above, the question posed by the Third Antinomy was whether an event could be caused both according to nature and from freedom, and Kant's answer was in the affirmative, because nature and freedom could stand "in a different relation" to effect in question. In light of Transcendental Idealism, one can now understand, at the most general level, that a different relation obtains in each case, because the cause is an appearance in the case of nature and a thing in itself in the case of freedom. In short, Transcendental Idealism resolves the contradiction argued for in the Third Antinomy by holding that the Thesis's assertion of freedom is possible for things in themselves (but not for appearances), while the Antithesis's assertion of determinism is true for appearances (but may be false for things in themselves; A531–532/B559–560). In this way the contradiction between freedom and determinism is avoided.¹⁰ Since this resolution is still too abstract and general, we return to it below.

In the meantime, it is instructive to consider the following question: How does this resolution square with the particular arguments of the Thesis and Antithesis? It is one thing to suggest that these arguments rely on Transcendental Realism, quite another to show in detail which specific features of Transcendental Idealism render these arguments invalid. The Thesis's argument against determinism turns on the idea that causality in accordance with the laws of nature contradicts itself "when taken in its unlimited universality." Because Kant distinguishes between appearances and things in themselves in such a way that the sensible world of appearances is essentially indeterminate and incomplete, the laws of nature cannot be taken to apply to the sensible world with unlimited

⁹ Kant explicitly affirms the idea that things in themselves must be simple substances in the *Metaphysical Foundations of Natural Science*, 4:507 (as well as in various metaphysics transcripts).

¹⁰ Kant presents the same solution later in the *Critique of Practical Reason*: "Consequently, if one still wants to save it [i.e., freedom], no other path remains than to ascribe the existence of a thing so far as it is determinable in time . . . *only to appearance, and to ascribe freedom to the same being as a thing in itself*" (5:95).

universality. More specifically, the Thesis argument assumes (1) that no event can occur unless its cause is “sufficiently determined” (P6) and (2) that an event cannot have its cause be sufficiently determined if there is no completeness in its causes (P5). In light of Kant’s distinction between appearances and things in themselves, it is clear that P5 and P6 are true for things in themselves, which are completely determinate, yet false for appearances, since they are indeterminate or incomplete with respect to their causes. In other words, the Principle of Sufficient Reason applies to appearances and things in themselves in different ways because what counts as a *sufficient* reason differs depending on the determinacy or indeterminacy of what exists.

As we saw above, the Antithesis argument against freedom holds only if the cause is temporal (an idea made explicit in P3), since it is the temporal determinacy of the cause that entails a prior cause, which is incompatible with the idea (expressed in P2) that a spontaneous or free cause must be uncaused by any prior cause. Thus the argument holds for appearances, which are necessarily temporally determinate, but not for things in themselves, since they are, in some sense, atemporal and thus not temporally determinate, either. By thus distinguishing between temporal appearances and atemporal things in themselves, Transcendental Idealism creates room for the possibility that things in themselves could be free.¹¹

Kant’s general strategy for solving the problem of free will and determinism in the Third Antinomy is thus clear. By rejecting the transcendental realist’s identification of appearances and things in themselves, he can ascribe determinism to the world of appearances and the possibility of a spontaneous causality of freedom to the realm of things in themselves. Such ascriptions are possible, however, only because of the specific differences between appearances and things in themselves. Things in themselves are completely determinate things that are not spatio-temporal and not for that reason subject to determinism, whereas appearances are essentially spatio-temporal objects governed by prior conditions that are neither fully determinate nor immediately given in their totality, but rather set “as a problem.”

¹¹ It is important to keep in mind Kant’s significant restrictions on what exactly has been demonstrated about freedom. Thus, he explicitly rejects the idea that we have any positive insight into the possibility of freedom, just as we do not have any insight into the possibility of fundamental powers (A448/B476).

ASPECTS OF TRANSCENDENTAL IDEALISM

While commentators might agree on the main lines of Kant's statement of, and solution to, the problem of free will and determinism when it is presented at a high enough degree of generality (i.e., as presented above), there are several more specific issues pertaining to this topic that have been the subject of considerable controversy. A first issue involves the nature of the distinction between things in themselves and appearances, with some viewing the distinction as exclusively epistemological and others seeing it as having ontological import as well. A second issue concerns one particular claim that Kant makes about the relationship between things in themselves and appearances, namely that the former "ground" or "underlie" the latter. To develop a more specific and detailed understanding of Kant's solution to the problem of free will and determinism, we consider each of these issues in turn.

Things in Themselves and Appearances: Epistemological
or Ontological?

One central dispute concerning Transcendental Idealism concerns whether the distinction between appearances and things in themselves is epistemological (sometimes called methodological) or rather ontological. According to the epistemological line of interpretation, there is just one reality and the distinction between appearances and things in themselves is to be cashed out in terms of the different ways we have of *considering* that reality. One prominent example of this line of interpretation – the abstraction version – has it that objects are appearances insofar as we consider those objects as they are given in spatio-temporal intuition, whereas those very same objects are things in themselves insofar as we *abstract* from how they are given in intuition and consider them only as they can be thought through the understanding. Since the distinction is between two different ways in which one set of objects is being considered (namely as given to us in intuition or in abstraction from how they are given to us), that is, two different aspects of one reality, the epistemological view is sometimes referred to as the "two-aspect" interpretation. By contrast, the ontological interpretation asserts that the distinction between appearances and things in themselves is not merely a distinction between ways that we consider objects, but also a distinction between two different realities. For example, one particular

version of this line of interpretation maintains that appearances are the set of spatio-temporal objects while things in themselves are an ontologically distinct set of non-spatio-temporal objects. Since this version holds that appearances and things in themselves are numerically distinct sets of objects or “worlds,” it is often called the “two world” interpretation.¹²

Unfortunately, on the basis of textual evidence alone no consensus has been reached as to which of these two interpretations is correct.¹³ The two-aspect interpretation finds confirmation in passages such as the following: “Do freedom and natural necessity *in one and the same action* contradict each other?” (A557/B585, emphasis added). However, proponents of the two-world interpretation can marshal support from quotations such as the following: “[T]he antinomy that there is freedom and necessity in the world is removed, because it is shown that they are not actual opposites, because they apply to *different objects*, namely the first to the noumenal world and the other to the phenomenal world, and therefore both can take place at once” (29:924–925, emphasis added). Since determining which interpretation might have a preponderance of textual evidence in its favor has not resulted in a clear decision, it is natural to bring philosophical considerations to bear on the debate.

On that score, the epistemological interpretation can be motivated by several powerful philosophical arguments. First, the epistemological interpretation is in a position to provide a succinct argument for Kant’s controversial thesis that things in themselves are not spatio-temporal. For example, if things in themselves are defined in abstraction from how reality appears to us and reality appears to us in space and time, then it would seem to follow immediately that things in themselves are not spatio-temporal, because one is thinking of a thing in itself only if one has abstracted all spatio-temporal properties away from reality.¹⁴ Moreover, this argument is not necessarily tied to the version of the epistemological

¹² Not all versions of the ontological interpretation are instances of the two-worlds interpretation. For example, both Rae Langton, *Kantian Humility* (New York: Oxford University Press, 1998), and Daniel Warren, *Reality and Impenetrability in Kant’s Philosophy of Nature* (New York: Routledge, 2001), develop ontological interpretations that hold that things in themselves and appearances are different aspects of one reality, namely intrinsic versus relational properties. For discussion of these views, see Karl Ameriks, “Kant and Short Arguments to Humility,” in *Kant’s Legacy: Essays in Honor of Lewis White Beck*, ed. P. Cicovacki (Rochester: University of Rochester Press, 2001), pp. 167–194, and my review of Warren’s book in *Notre Dame Philosophical Reviews* (February 2002) (<http://ndpr.icaap.org/content/archives/2002/2/watkin-warren.html>).

¹³ See Gerold Prauss, *Kant und das Problem der Dinge an sich* (Berlin: de Gruyter, 1974).

¹⁴ Henry Allison presents such an argument in *Kant’s Transcendental Idealism*, pp. 113–114.

interpretation that invokes “abstraction” in its description of the thing-in-itself aspect of reality. A “standpoints” version of the epistemological interpretation – a version that interprets the distinction between appearances and things in themselves as involving differences that arise from our taking up either a theoretical/scientific or a practical/deliberative standpoint – can argue for the same conclusion.¹⁵ Objects that we investigate insofar as we are engaged in scientific inquiry are necessarily spatio-temporal, whereas maxims or potential courses of action that we are considering from the first-person standpoint of deliberation are understood not as spatio-temporal events caused by previous spatio-temporal events, but rather as principles that we might act on due to reasons that we accept as following from our conception of the good.

Second, the epistemological interpretation can propose that Kant’s “Critical turn” advances a radically new and provocative picture of reality and of our access to it. According to this interpretation, Kant is boldly rejecting the idea we can even talk meaningfully about a reality that would exist independently of our human standpoint, since the only way to talk about such a reality presupposes a God’s-eye point of view, a point of view that is simply inaccessible to us.¹⁶ Rather, reality must be understood in terms of cognitive capacities that are specifically human.¹⁷ According to the epistemological view, it is this insight – that metaphysics is essentially vacuous without, or at least heavily dependent on, epistemology – that is central to Kant’s “Critical turn,” a point that resonates in contemporary debates about antirealism in metaphysics, meta-ethics, philosophy of science, and philosophy of language.

Third, the epistemological interpretation can seem to provide an extremely attractive solution to the problem of free will and determinism. If things in themselves and appearances are simply different aspects of reality, or different standpoints from which reality is being considered, the problem of free will and determinism appears to admit of a breathtakingly

¹⁵ Henry Allison (*Kant’s Transcendental Idealism*) is a prominent defender of the abstraction version, while Christine Korsgaard, *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996), defends the standpoints version. However, in *Kant’s Theory of Freedom*, Allison seems to accept many aspects of the standpoint version as well, since the two are not mutually exclusive; it could be by abstracting from space and time that one adopts the deliberative standpoint or vice versa.

¹⁶ Once again, Henry Allison has presented this position forcefully in the context of Kant scholarship. Hilary Putnam’s internal realism represents this view in a contemporary context.

¹⁷ When this kind of interpretation of Kant is emphasized, he is typically viewed as a verificationist.

simple solution. Take the standpoints version of the epistemological interpretation. From the deliberative standpoint we must consider ourselves free, since the very act of deliberation that defines such a standpoint presupposes that the outcome has not already been determined. However, from the theoretical standpoint, scientists, who are simply investigating the causes of natural events in the world, are perfectly justified in considering our actions to be fully causally determined, since our actions are merely a subset of the events occurring in the world for which they hope to find the proper causes. By maintaining that the differences between these two standpoints preclude one from adopting them both simultaneously, the problem of free will and determinism is resolved, since there is no single standpoint from which a contradiction between their assertions can even be formulated.

The advantages that the epistemological interpretation might seem to enjoy have not, however, gone unchallenged. First, it is not clear that the easy argument for the non-spatio-temporality of things in themselves suggested by the epistemological interpretation is ultimately successful. For it may turn out that the argument is either too weak or too strong. It could be too weak insofar as it may not be able to exclude from things in themselves spatio-temporal-“like” properties, that is, properties that are not numerically identical, but rather qualitatively identical to spatio-temporal properties.¹⁸ However, it may, instead, be too strong. For it might rule out the possibility that things in themselves are spatio-temporal *analytically*. That is, if what it *means* to be a thing in-itself is to be non-spatio-temporal, then the argument works, but it does so trivially insofar as it does not rule out what Kant seems to view as a meaningful possibility, namely that space and time might be things in themselves or relations between things in themselves, just as Newton and Leibniz had thought.¹⁹

Second, while it can be tempting to emphasize the novelty of Kant’s position (in the guise of Kant’s “revolutionary” “Critical turn”), one must guard against exaggerating what is new in it. As a matter of historical fact, changes in philosophical frameworks are rarely as radical as one might (like to) think, since philosophers inevitably adopt at least part of the framework, terminology, and/or distinctions of their predecessors. Philosophically, even if antirealism represents an important option in several areas of contemporary philosophy, it is not undisputed. For even if

¹⁸ See Lorne Falkenstein, “Kant’s Argument for the Non-Spatiotemporality of Things in Themselves,” *Kant-Studien* 80 (1989): 265–283.

¹⁹ See, e.g., Langton, *Kantian Humility*, pp. 9–12.

one concedes that significant benefits can be gained by defining meaning (and truth) in terms of conditions that would be wholly accessible to us, one can still dispute that those benefits outweigh the costs associated with that view. In particular, one might argue that defining meaning and truth in terms of human cognitive abilities constitutes an overly anthropocentric conception of reality. One could also object that adopting a standpoint does not necessarily determine the metaphysical facts apprehended from within that standpoint, even if it may influence how those facts must be characterized. For example, adopting the practical standpoint may require that deliberation occurs and that it be oriented toward an individual's conception of the good, but it should not require that a given individual is committed to, say, one conception of the good rather than another. Thus, whether or not it can be established on the basis of available historical evidence to what extent Kant was groping toward a more radical conception of reality, there may be good philosophical reasons that would hold him (or us) back from embracing epistemologically oriented views, even if he (or we) might find them tempting in certain contexts.

Third, and most important for our current purposes, one might question whether the solution to the problem of free will and determinism suggested by the epistemological interpretation is ultimately satisfying. The crucial move in that solution is to assert that an action can be both free and determined only if these otherwise contradictory attributions are made from different standpoints, since the fact that the standpoints cannot be held at the same time entails that one is never in a position to state the contradiction coherently.

However, this move is susceptible to two serious questions. The first question is: What is the nature of the difference between the standpoints that precludes the possibility that they could be held at the same time?²⁰ While it might be tempting to describe this difference in terms of incommensurability, such a description seems to be inconsistent with two fundamental features of Kant's philosophical system. First, the difference between our epistemic access to appearances and things in themselves would seem to consist not in the fact that we use *incommensurable*

²⁰ For interesting discussions of this and related issues, see Dana Nelkin, "Two Standpoints and the Belief in Freedom," *Journal of Philosophy* 97 (2000): 564–576, and Fred Rauscher, "Kant's Two Priorities of Practical Reason," *British Journal of the History of Philosophy* 6 (1998): 397–419. Although neither Nelkin nor Rauscher discusses the difference between standpoints in terms of incommensurability, they do both raise considerations that would apply to the issue when framed in such terms.

cognitive resources in considering things in themselves and appearances, but rather simply in the fact that appearances would be defined by means of the *addition* of spatio-temporal forms of intuition to the categories (which are used for both things in themselves and appearances). If the cognitive resources appealed to are simply greater in the one case than in the other, it is not obvious that the one case would therefore be incommensurable with the other. Second, in deliberation we typically do acknowledge that our actions are restricted by what is physically possible for us, which strongly suggests that the deliberative or practical standpoint is not completely incommensurable with the theoretical. It seems to be the case rather that the former is able to incorporate at least some facts from the latter. This same point can be brought out by reflecting on Kant's principle that "ought" implies "can," since in some cases we readily admit that what we ought to do (on the basis of our rational deliberations) is restricted by what we can do (given what is physically possible for us). Thus, appearances and things in themselves do not seem to be incommensurable and it remains unclear why the two standpoints cannot be held at the same time.

The second question is: Regardless of whether or not the two standpoints can be held at the same time, can they both be true or must one of them be illusory? The idea is supposed to be that one is free only from the practical/deliberative standpoint and determined only from the theoretical/scientific standpoint, and as long as one cannot adopt the practical and theoretical standpoints at the same time, an action would never be both free and determined at the same time. However, since the answer to the question "Am I free in action A or not?" will depend on which standpoint one has adopted, one can now raise the following question: So which standpoint is accurate? That is, granted that we conceive of ourselves as free and as determined (albeit from different standpoints at different times), which of these conceptions contains a true description of how we are? At this point, the proponent of the epistemological interpretation could claim – in line with the antirealist position underlying the previous point – that this last set of questions is illegitimate, perhaps suggesting that one would have to adopt either a God's-eye viewpoint or a standpoint outside all standpoints so as to determine the accuracy or inaccuracy of each one. However, just as one might think that truth is not completely determined by the adoption of a standpoint, even if its characterization may be influenced by such an adoption, so one can reply that no further standpoint is required, since the accuracy of a standpoint is not determined by any putatively divine meta-standpoint,

but rather simply by the metaphysical facts of the matter. Therefore, the epistemological interpretation's solution to the problem of free will and determinism may not be able to sustain the advantages claimed for it.

If the epistemological interpretation runs the risk of not being able to resolve the problem of free will and determinism, one might be tempted to adopt its competitor, the ontological interpretation. For, whatever else may be true of the ontological interpretation, it can at least avoid the logical contradiction between freedom and determinism in a very straightforward way, namely by asserting that they pertain not to "different" aspects of one and the same thing but rather to distinct entities altogether. If contradictory predicates must be ascribed to one and the same entity to generate a contradiction, then the ontological interpretation is not pressured in the same way as the epistemological interpretation is, since the contradictory predicates are ascribed to numerically distinct entities.

Moreover, one might suspect that the ontological interpretation can derive an advantage from precisely those issues that presented the epistemological interpretation with its difficulties. First, if things in themselves are defined as numerically distinct from appearances rather than in terms of abstraction or standpoints, then it is clear that no "short" argument for the non-spatio-temporality of things in themselves will follow immediately or analytically from any considerations about appearances and that Kant must develop a "long" argument for this conclusion (one that requires all of the intricate twists and turns that we find in the *Critique*).²¹ Second, if the distinction between things in themselves and appearances is ontological rather than epistemological, then Kant would clearly have room for metaphysical facts about how things "really" are, even if one of his most fundamental goals in the *Critique* is to establish that we cannot attain knowledge of such matters, in spite of the fact that our reason is constituted in such a way that we must pose such questions about the ultimate character of reality. Third, because the ontological interpretation is ontological, the epistemological distinction between standpoints that we might or might not adopt need not carry any significant argumentative weight on this issue. In particular, since these standpoints need not be incommensurable, at least in principle one can allow that our practical deliberations might take into account restrictions stemming from our physical limitations.

²¹ See Karl Ameriks, "Kantian Idealism Today," *History of Philosophy Quarterly* 9 (1992): 329-342.

Even if the ontological interpretation enjoys the advantages just outlined, it no doubt encounters significant objections of its own. These objections typically attack either the coherence or the plausibility of ascribing atemporal causation to the free actions of things in themselves. For some, the difficulty lies with the very idea of atemporal causation; since every instance of causation that I know in an uncontroversial way occurs at a particular time, one might infer that causation is an essentially temporal notion and that atemporal causation is thus impossible. For those who would prefer to remain agnostic about whether atemporal causation is conceptually incoherent or at least uninstantiated in our world, the objection focuses on the implausibility of ascribing it to us in our free actions. If a free act is atemporal and thus completely outside time, it becomes tempting to ask: What does such a capacity have to do with *me*? After all, the actions I seem to perform freely (if I perform any freely at all) are the day-to-day actions that fill my life, which calls into question the idea that anything outside time is directly relevant to what I cherish more than most other things, namely my freedom.

The proponent of the ontological interpretation is not necessarily without resources to respond to these objections. For example, while empiricists might have grounds for objecting to notions of atemporal causality on the grounds that they cannot be derived from essentially temporal perceptions, it is clear that such an objection carries no weight for Kant. For the *unschematized* category of causality will not contain any temporal content, given that Kant quite clearly holds that the schemata are what *add* temporal content to the atemporal categories. Moreover, one could also argue that, contrary to initial appearances, our fundamental conception of ourselves must have a nonempirical and even nontemporal component.²² For insofar as I can imagine that I might have had a different body and lived and acted in a different place and time, I must be holding “fixed” some essential feature of myself that is not bound to the particular spatio-temporal properties that I have merely contingently, so that I can consider various counterfactual situations in which “I” might be found. If these rejoinders can be developed in detail, the ontological interpretation may be not only represented as a logically possible metaphysical position, but also even made to sound reasonably plausible.

²² See Allen Wood, “Kant’s Compatibilism,” in *Self and Nature in Kant’s Philosophy*, ed. A. Wood (Ithaca: Cornell University Press, 1984), pp. 73–101. The specific line of reasoning presented in the rest of the above paragraph is not, however, suggested by Wood. We have reason below to return to certain aspects of Wood’s discussion.

As a result, one cannot conclude at this point that either the ontological or the epistemological interpretation enjoys decisive philosophical support.

If neither the epistemological nor the ontological interpretation finds decisive textual or philosophical confirmation, what conclusion ought we to draw about the proper interpretation of Transcendental Idealism? The most plausible hypothesis to adopt in the current situation is that Kant was not always unambiguous about how to understand this doctrine in certain contexts. In fact, the historical context presented in Chapters 1 and 2 provides independent support for this hypothesis. For one can see how Kant might have been unclear about the two views by considering the equivalent question in Leibnizian metaphysics: Are bodies identical to the monads that constitute them or are they distinct entities?²³ Insofar as bodies are spatial and monads are not, it would seem that they must be ontologically distinct entities. Insofar as bodies are constituted by monads (e.g., by their confused perceptions) and every monad perceives the world from the point of view of *its own* body, it would seem that they must be identical to each other (in some sense). And what might clarify the appropriate answers to these questions – namely a detailed description of the relationship between bodies, along with their derivative forces, and monads endowed with primitive forces – was, as we saw in Chapter 1, unavailable in public philosophical discourse at the time.²⁴ As a result, given all of the evidence available, Kant may have been no clearer about this issue than were his predecessors, and this lack of clarity may have infected the development and deployment of Transcendental Idealism.

The Grounding Thesis

Despite the uncertainty and ambiguity that Kant's use of these two different understandings of Transcendental Idealism creates at a general level, we can still turn to a particular aspect of Transcendental Idealism that is fundamental to his understanding of freedom and determinism, namely the issue of "grounding."

²³ In defense of Leibniz and Kant, questions about whether or not composition is distinct from identity are notoriously difficult, as is evidenced by the current (and, by now, long-standing) debate about whether or not, for example, a statue is identical with the clay of which it is composed.

²⁴ For considerations that complement this point, see Robert Adams, "Things in Themselves," *Philosophy and Phenomenological Research* 57 (1997): 801–825, esp. 823–825.

Grounding and the Epistemological Interpretation

In various works, Kant repeatedly suggests that things in themselves, or the noumenal world, “grounds” or “underlies” appearances, or the sensible world. For example, in his *Metaphysics Mrongovius* transcripts, Kant asserts that the senses “show us merely the appearances of the things. But these are not the things themselves. They indeed underlie the appearances [*liegen . . . zum Grunde*], and I can therefore surely infer the actuality of the things from the appearances, but not the properties of the things themselves from the properties of the appearances” (29:857). In the third section of the *Groundwork*, we find a similar remark: “. . . we must admit and assume behind appearances something else that is not appearance, namely things in themselves, although, since we can never become acquainted with them but only with how they affect us, we resign ourselves to being unable to come any closer to them or ever to know what they are in themselves” (4:451). In short, despite significant epistemic limitations associated with the critical turn, we can still claim, Kant thinks, that things in themselves underlie or ground appearances.²⁵

In the *Critique*, the grounding relationship is described in more detail as follows.

If, on the other hand, appearances do not count for any more than they are in fact, namely not for things in themselves but only for mere representations connected in accordance with empirical laws, then they themselves must have grounds that are not appearances. Such an intelligible cause, however, will not be determined in its causality by appearances, even though its *effects* appear and so can be determined through other appearances. (A537/B565, emphasis added)

Here Kant explicitly asserts that the grounding relationship is causal, since he identifies appearances as the effects of an intelligible cause (or set of such causes). Moreover, this assertion is not an isolated remark on Kant’s part. In setting up his resolution of the problem of free will and determinism, he asks, in a suggestive tone, the following question: “Is it not rather possible that although for every effect in appearance there is required a connection with its cause in accordance with laws of empirical causality, this empirical causality itself . . . could nevertheless be an *effect of a causality* that is not empirical, but rather intelligible?” (A544/B572, emphasis added).

²⁵ The grounding thesis is also expressed in the *Metaphysical Foundations of Natural Science*, 4:507.

In *Reflexion* 5611 (tentatively dated 1778–1779) Kant adds: “[T]he actions of reason itself are not also appearances, but rather only its effects are [appearances]. . . . Now actions are in large part occasioned by sensibility, but not completely determined; for reason must give a complement of sufficiency” (18:252). This final passage not only confirms that Kant thinks of the grounding relationship between things in themselves and appearances as causal, but also reinforces the idea (developed in the course of the Antinomies) that appearances are not completely determined on their own and lack a complete sufficient reason. For this passage states that things in themselves are required precisely to provide the sufficient determination that appearances lack on their own.

If Kant thus thinks that things in themselves *cause* appearances and add the sufficient reason that they lack on their own, what else does he want to say about this relationship? On the one hand, Kant makes a series of epistemological remarks about this relationship. For example, in two of the passages just cited (29:857 and 4:451), immediately after asserting the grounding relationship, Kant clarifies how such an assertion does not run afoul of his view that we cannot have substantive knowledge of particular things in themselves. Even if we can know that things in themselves must exist in order to cause appearances, we cannot know any particular properties of individual things in themselves on the basis of the appearances that we perceive.

Nevertheless, Kant sometimes adds that the relationship is not completely inscrutable. In *Reflexion* 5612, Kant writes: “Actions here in the [phenomenal] world are mere schemata of the intelligible [world]” (18:253). In the *Critique* we find similar remarks: “In regard to the intelligible character, of which the empirical one is only the *sensible schema*, no **before** or **after** applies, and every action, irrespective of the temporal relation in which it stands to other appearances, is the immediate effect of the intelligible character of pure reason” (A553/B581, italics added); “the intelligible character, which is the transcendental cause of the former [i.e., the empirical character], is passed over as entirely unknown, except insofar as it is indicated through the empirical character as only its sensible sign” (A546/B574). In these passages Kant suggests that the empirical character is a (sensible) schema of the intelligible character, which seems to imply that it in some way “mediates” between the atemporal intelligible character and its various temporal realizations in the phenomenal world. While one might like to hear a more detailed story about how such a schema might work, Kant quite consistently refrains

from attempting to explain which particular intelligible character we might be able to impute to things in themselves on the basis of the empirical characters that cause the particular actions we perceive in the sensible world.²⁶

On the other hand, despite the epistemic limitations Kant places on what we can know about how *specific* features of things in themselves might ground appearances, Kant makes several *general* claims about grounding. For one, Kant makes clear that the grounding relationship is one-way and not reciprocal. As we saw above, he remarks that “such an intelligible cause, however, will not be determined in its causality by appearances, even though its effects appear and so can be determined through other appearances” (A537/B565). Similarly, “reason therefore acts freely, without being determined dynamically by external or internal grounds temporally preceding it in the change of natural causes” (A553/B581). Things in themselves ground appearances, but appearances do not ground things in themselves.

For another, in the third section of the *Groundwork* – in a crucial part of his argument for the claim that we should view the moral law as having normative force for us – Kant states:

But because *the world of understanding* [the noumenal world] *contains the ground of the world of sense and so too of its laws*, and is therefore immediately lawgiving with respect to my will . . . , it follows that I shall cognize myself as intelligence, though on the other side as a being belonging to the world of sense, as nevertheless subject to the law of the world of the understanding. (4:453)

Not only does the noumenal world of things in themselves cause the *existence* of appearances in the sensible world, but it is also responsible for the *laws* that govern appearances.

Though we have reason to return to these features of the grounding thesis, we should note that it does not sit well with the epistemological interpretation of Transcendental Idealism. The standpoints version of the epistemological interpretation seems to be incompatible with the grounding thesis, because the assertion that “things in themselves ground appearances” is a claim that cannot be made from either standpoint. Assertions about things in themselves can be made only from the practical or deliberative standpoint, while claims about appearances can be made

²⁶ For example, in a footnote to the Third Antinomy, Kant explicitly notes: “The real morality of actions (their merit and guilt), even that of our own conduct, therefore remains entirely hidden from us. Our imputations can be referred only to the empirical character” (A551/B579).

only from the theoretical or scientific standpoint. The abstraction version of the epistemological interpretation is in a similar situation. Assertions about things in themselves abstract from spatio-temporality, while assertions about appearances do not, so that the assertion “things in themselves ground appearances” would have both to abstract and not abstract from spatio-temporality.

The ontological, two-worlds interpretation of Transcendental Idealism, by contrast, has no difficulties with the grounding thesis. Things in themselves, which are members of one set of entities, cause appearances, which are members of another set of entities. In light of the interpretation suggested above, however, one should not think that this incompatibility amounts to a decisive objection to the epistemological interpretation; rather, it simply reveals that in certain contexts Kant sees the need to appeal to resources available only to an ontological understanding of Transcendental Idealism.

Grounding and Compatibilism

One central context in which the grounding thesis is relevant involves the question of compatibilism (as it is commonly understood in contemporary contexts). At the most general level, contemporary compatibilists hold that freedom and determinism are compatible because our actions can be free, despite the fact that they are also determined by natural causes. Present-day incompatibilists, by contrast, hold that freedom and determinism are incompatible because our actions cannot be free if they are determined by natural causes. Considerable discussion is then typically devoted to debating whether being free must entail that “I could have done otherwise,” whether being determined by natural causes implies that “no event could have been otherwise” (given the laws of nature and the initial conditions of the universe), and finally whether, if both of these claims are accepted, the contingency entailed by a free action is or is not compatible with the necessity of events that are determined by natural causes.

Is Kant a compatibilist in this sense or not? On the one hand, Kant clearly rejects as “a wretched subterfuge” (5:96) any version of compatibilism that is based on what he calls a comparative concept of freedom, that is, where the compatibility of free will and determinism derives from a difference in whether the action is caused by “determining grounds *within* the subject or *outside* him” (5:96), since he thinks that determining grounds, of whatever kind they may be, determine with

necessity.²⁷ In light of such an assertion, one might think that Kant is an incompatibilist. Plus, is it not the very incompatibility of freedom and causal determinism that is supposed to generate the contradiction expressed in the Thesis and Antithesis of the Third Antinomy? On the other hand, if there are different versions of compatibilism, Kant could easily reject one version in favor of another. Moreover, as we saw above, the main outlines of his solution to the problem of free will and determinism seemed to consist in the assertion that it is precisely Transcendental Idealism's distinction between things in themselves and appearances that allows one to hold freedom and causal determinism at the same time. Accordingly, Kant might ultimately seem to be a compatibilist, saddled with the compatibilist's difficulty of explaining how an event can be both free and determined with necessity by previous events according to the laws of nature. However, if one accepts the two-world, ontological interpretation of Transcendental Idealism, then it might seem that it is precisely because freedom and determinism are incompatible that one must divide things into two separate worlds and attribute (the possibility of) freedom to (some) members of the one world (noumenal selves) and determinism to members of the other (the phenomenal world) so that Kant ends up being an incompatibilist.

What's worse, this perhaps dizzying set of possible answers might seem to be supported by ample textual documentation. As we saw in our discussion of the Antithesis argument, Kant understands transcendental freedom as "an absolute causal spontaneity beginning from itself" (A446/B474), that is, as "a faculty of absolutely beginning a state" by which he means that "nothing precedes it through which this occurring action is determined in accordance with constant laws" (A445/B473). Whatever "absolute spontaneity" and other such phrases mean, Kant is clearly committed to the idea that such causality is independent of all natural causes.

But the phrase "independent of all natural causes" stands in need of clarification if we are to ascertain whether or not Kant accepts compatibilism (and in what sense). Does it mean that free actions are determined by something other than natural causes, for example, by a so-called noumenal self, though such a self would be, in a sense, "forced" to cause its actions in such a way that they *necessarily coincide* with those actions that are determined with necessity by natural causes, as the compatibilist would

²⁷ Kant is thus critical of his own very early pre-Critical solution to the problem of free will and determinism as presented in the *Nova dilucidatio* (1:400).

have it? Or does it mean that free actions are independent in the sense that they can act *contrary to* or *interrupt* what was (or would have been) determined as necessary by natural causes according to the laws of nature, as the incompatibilist would claim?

On the one hand, several passages forcefully uphold the necessity of natural causes (and thus seem to support a compatibilist interpretation). As we saw above, Kant proffers his solution to the problem of free will and determinism by asking: "Is it not rather possible that . . . this empirical causality itself, *without the least interruption* of its connection with natural causes, could nevertheless be an effect of a causality that is not empirical, but rather intelligible?" (A544/B572, emphasis added). Similarly,

all the actions of the human being in appearance are determined in accord with the order of nature by his empirical character and other cooperating causes; and if we could investigate all the appearances of his power of choice down to their basis, then there would be no human action that we could not predict with certainty, and recognize as *necessary* given its preceding conditions. (A549-550/B577-578, emphasis added)²⁸

According to such passages, human actions are like all natural events; if we could discover all of their natural causes, then not only could we predict each and every one of them with certainty, but we would know them to be necessary. Even more clearly, Kant explains:

We need the principle of the causality of appearances in order to be able to seek and specify the natural conditions . . . for natural occurrences. . . . Now this is *not in the least impaired*. . . . if one assumes that among natural causes there are also some that have a faculty that is only intelligible, in that its determination to action never rests on empirical conditions but on mere grounds of the understanding, *as long as the action in the appearance* of this cause *accords with all laws of empirical causality*. For in this way the acting subject, as *causa phaenomenon*, would have all its actions linked with inseparable dependence to the natural chain of causes, and only the *phaenomenon* of this subject . . . would contain certain conditions that, if one would ascend from empirical objects to transcendental ones, would have to be regarded as merely intelligible. For if we follow the rule of nature only in that which might be the cause among appearances, then we need not worry about what sort of ground is thought for these appearances and their connection in the transcendental subject, which is empirically unknown to us. This intelligible ground does not touch the empirical questions at all. (A544-545/B572-573, italics added)

²⁸ One finds a similar expression of necessity in the following passage: "Could reason's action then be called free even though in its empirical character . . . it is all precisely determined and necessary?" (A551/B579).

Even assuming that an intelligible cause determines appearances, there is no problem as long as this cause acts in accord with all laws of empirical causality. Accordingly, it seems clear that Kant is committed to natural necessity and that this necessity cannot be interrupted or impaired by intelligible free causes and is thus “untouchable.”

However, there are apparently “incompatibilist” passages as well. In the Resolution of the Third Antinomy Kant notes that “one regards the causality of reason not as a mere concurrence with other causes, but as complete in itself, *even if sensuous inclinations were not for it but were indeed entirely against it*” (A555/B583, emphasis added). And in explaining his notion of practical freedom – which is distinct from transcendental freedom insofar as it is defined negatively in terms of independence from necessitation by sensible inclinations in particular and positively in terms of autonomous determination by pure reason – Kant remarks that practical freedom presupposes that

although something has not happened, it nevertheless **ought** to have happened, and its cause in appearance was thus not so determining that there is not a causality in our power of choice such that, independently of those natural causes *and even opposed to their power and influence*, it might produce something determined in the temporal order in accord with empirical laws, and hence begin a series of occurrences **entirely from itself**. (A534/B562, italics added).

In *Reflexion* 5612 (tentatively dated at 1778–1779), Kant again seems to express the idea that a free action at the noumenal level presupposes the ability to have done otherwise and that it could do so by “taking the place” of an efficient cause at the phenomenal level:

We explain freely performed actions according to the laws of the nature of human beings, but we do not thereby cognize them as determined; otherwise we would not view them as contingent and require that they could have and should have occurred otherwise. In free actions reason is influential [*fließt ein*] not merely as a comprehending, but [also] as an efficient and driving [*treibendes*] principle. How it might not merely reason and judge, but [also] take the place of a natural cause we have no insight into. . . .²⁹ (18:253)

In light of this textual evidence, it might appear that Kant is quite confused about whether or not he wants to be a compatibilist. In some

²⁹ Kant may not literally mean “take the place of” since that would imply that something noumenal, namely the exercise of reason, would (nonsensically) be something phenomenal, namely an efficient natural cause. Similarly, the passage from A534/B562 may be read as pertaining not to the level of efficient causality, but rather to the idea that there may be a normative dimension in addition to what naturally occurs, and the opposition that Kant speaks of may not be an opposition of efficient causality.

passages, he seems to recognize that if he insists on the contingency implicit in one's free action, then he must also reject the necessity of causal determinism. In others, he appears to note that the necessity of determinism is an indispensable feature of the phenomenal world (in light of its essential temporality), but then fails to explain how our free actions could have been other than what appearances dictate. In short, in the passages considered so far, Kant *seems* unable to decide between the poles of necessity and contingency as embodied in the issue of free will and determinism.

In an excellent article, titled "Kant's Compatibilism," Allen Wood has suggested that this particular issue, which he calls the problem of indifference, can be solved by invoking the atemporality of our noumenally free choices:

How is the timelessness of the noumenal self supposed to safeguard our ability to do otherwise than we in fact do? . . . Events in time follow a necessary order, as determined by their natural causes. A particular timeless choice of my intelligible character affects the natural world by selecting a certain subset of possible worlds, namely those including a certain moral history for my empirical character, and determining that the actual world will be drawn from that subset of possibilities. For each such choice there is an almost endless variety of ways in which I might have chosen differently, and endless variety of possible empirical selves and personal moral histories I might have actualized. Of every one of my misdeeds it is true that I would have left it undone had I made a different timeless choice. Hence it is in my power to leave any misdeed undone, despite the fact that in the actual world it follows inescapably from what preceded it in time.³⁰

In light of this resolution to the problem of indifference (and other issues), Wood concludes that Kant "decided that the temporality of our agency is the necessary ransom that must be paid to the free will problem if our high vocation as moral agents is to be preserved."³¹

There is no doubt that Kant accepts the idea that our noumenal agency must be considered atemporal, and it is equally clear that this view is forced on him by his belief that temporality requires causal determination in such a way that temporal agency would be incompatible with the kind of spontaneous freedom he thinks is required for freedom and moral responsibility.³² However, the question at hand is how to understand Kant's views on the nature of the modal conflict between the necessity

³⁰ Wood, "Kant's Compatibilism," pp. 90–91.

³¹ *Ibid.*, pp. 100–101.

³² I also agree with Wood that Kant's notion of atemporal agency is crucial to the problem of fatalism, which he discusses in the last part of "Kant's Compatibilism."

of events determined by natural causes and the contingency implicit in freedom. With respect to that question, the matter of atemporality may not be the most crucial issue. After all, our current focus is on a conflict between different modalities, necessity and contingency, which seems to be an issue different from (even if closely related to) that of temporality.

Rather, what would seem to be in a position to accomplish significant work in the explanation Wood describes is, in fact, the grounding thesis.³³ Though the choices that selves make qua noumenal beings are atemporal, it is *what* they choose (not *when* they choose it) that is to resolve the modal conflict. On Wood's account what they choose is "a certain subset of possible worlds" where the principle of selection involves "moral histories" and "empirical characters." In light of our discussion in Chapter 4, it may be helpful (and still in line with the spirit of Wood's interpretation) to say not that they choose a certain subset of possible worlds directly, which might sound counterintuitive, but rather that they immediately choose their own natures and only indirectly the laws of nature that are based on them, which in turn determine (along with initial conditions) what is and is not necessary. Yet saying that there is something noumenal (a free choice) that at least indirectly causes (by selecting) the phenomenal laws (and their necessity as well) is simply to assert a particular version of the grounding thesis.³⁴

But how exactly does this version of the grounding thesis resolve the modal conflict that is explicitly raised in the issue of compatibilism? As we saw above, events in the world of appearances are determined to follow with necessity from the laws of nature, and that claim stands firm. However, since the laws of nature that govern the world of appearances depend on contingent features of things in themselves, the necessity of appearances is conditional rather than absolute. Given that natural necessity is conditional on contingent features of things in themselves (more specifically, on their free choices), Kant is justified in saying that events that occur in the phenomenal world (my actions included) could have been otherwise. For if things in themselves were different from what they are – in particular, if rational agents had freely chosen to act

³³ Wood discusses the grounding thesis just prior to addressing the problem of indifference, so he may well be relying on it as well in his discussion, in which case the interpretation developed below is consistent with his main line.

³⁴ For clear textual evidence in support of this point, see Kant's discussion of freedom and determinism in the *Prolegomena*, where he asserts not only that things in themselves cause appearances, but even that in at least some cases, "reason is the *cause of these natural laws* and is *therefore free*" (4:346, emphasis added).

differently – then events in the phenomenal world (my actions included) would have been different as well.

Kant expresses this point in *Reflexion* 5617 (tentatively dated 1778–1779): “Among appearances everything must be determined, but either according to pathological or moral laws. In the former case, its opposite is *still possible according to the laws of reason*, and thus man is free; in the latter case, the subject is free as well” (28:256, emphasis added).³⁵ That is, events determined by pathological laws (laws that hold for certain appearances) are not necessary *if* one takes into account the fact that these laws are, in turn, derived from the free (i.e., rational) choices of things in themselves. However, *given* that things in themselves are what they are, the laws that govern the appearances that depend on them are necessary. By having the laws of nature depend on things in themselves Kant can show “the compatibility of compatibilism and incompatibilism.”³⁶

Does the grounding thesis (in this particular form) necessarily entail that I (or any other noumenally free being) must be completely responsible for the laws of nature and everything that follows from them (including all the evils that transpired throughout the course of history)? It is important to understand that the laws of nature are nothing other than laws of the natures of things. That is, the laws of nature that hold in a given world are a function of the natures that are instantiated in that world. Thus, it is important to stress that what personal agents freely choose are not immediately the laws of nature but rather their own

³⁵ He also makes this point in two other passages. In the Third Antinomy, Kant notes that reason “makes its own order according to ideas, *to which it fits the empirical conditions* and according to which it even declares actions to be *necessary* that yet **have not occurred** and perhaps will not occur” (A548/B576, italics added). Similarly, in the *Critique of Practical Reason* he remarks that “the natural necessity which cannot coexist with the freedom of the subject attaches merely to the determinations of a thing which stands under conditions of time and so only to the determinations of the acting subject as appearance, and that, accordingly, the determining grounds of every action of the subject so far lie in what belongs to past time and *is no longer within his control* (in which must be counted his past deeds and the character as a phenomenon thereby determinable for him in his own eyes). But the very same subject, being on the other side conscious of himself as a thing in itself, also views his existence *insofar as it does not stand under conditions of time* and himself as determinable only through laws that he gives himself by reason; and in this existence of his nothing is, for him, antecedent to the determination of his will, but every action . . . is to be regarded in the consciousness of his intelligible existence as nothing but the consequence and never as the determining ground of his causality as a *noumenon*” (5:97–98).

³⁶ We can thus agree with Wood’s well-known description of Kant’s position in “Kant’s Compatibilism” (p. 74).

natures.³⁷ Once all the natures in the world have been determined (whether through free choice or otherwise), then the laws of nature (and the necessity that depends on them) are set.³⁸ Moreover, to say that personal agents freely choose their own natures is simply another way of saying that personal agents are responsible for their noumenal and empirical characters, which is consistent with what common sense dictates, namely that we be at least partially responsible for our characters.³⁹

In response, one might well concede that the position sketched above does render intelligible relevant quotations from Kant and may even represent a coherent philosophical position, but still have doubts about whether it can really resolve the conflict between determinism and free will. For example, in *An Essay on Free Will*, van Inwagen defends incompatibilism by means of what he dubs the Consequence Argument:

If determinism is true, then our actions are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.⁴⁰

According to the Consequence Argument, because our actions, like all other noninitial states of the universe, are simply consequences of (1) a prior state of the universe and (2) the laws of nature, and neither one of them is up to us, our actions cannot be free (or up to us).

The debate as framed by the Consequence Argument is no longer about necessity and contingency alone (or *per se*), but rather about what

³⁷ In the beginning of the Critical Elucidation of the Analytic of Pure Practical Reason of the *Critique of Practical Reason*, Kant distinguishes between theoretical and practical reason by noting that the former is given objects in intuition, while the latter is concerned with making them real. In this context, he says of practical reason that "it does not have to provide an object of intuition, but, as practical reason, *only a law* for such an object" (5:89).

³⁸ One might object that our empirical natures are caused by prior events so that our choice of our empirical natures really does entail prior events. However, there is an important distinction between causing a nature that is instantiated in the world to be efficacious in certain ways and causing a nature to be instantiated in the world.

³⁹ This explanation differs from the one given by Wood ("Kant's Compatibilism," p. 92) only terminologically. This view finds textual support in a passage in the *Critique of Practical Reason*: "A rational being can now rightly say of every unlawful action he performed that he could have omitted it even though as appearance it is sufficiently determined in the past and, so far, is inevitably necessary; for this action, with all the past which determines it, belongs to a single phenomenon of his character, which he gives to himself and in accordance with which he imputes to himself... the causality of those appearances" (5:98).

⁴⁰ Peter van Inwagen, *An Essay on Free Will* (New York: Oxford University Press, 1983), p. 56.

is and is not up to us.⁴¹ However, given that Kant's solution to the modal conflict depends on the laws of nature being up to us, their positions are still directly relevant to each other. Van Inwagen begins by assuming that "no law of nature is such that anyone can render it false" on the grounds that "the laws of nature impose limits on our abilities." In support of this point he notes the apparent absurdity of a bureaucrat ordering an engineer to build a spaceship capable of traveling faster than the speed of light, with the justification that if the laws of nature were up to us, then the engineer ought to be able to accomplish the physically impossible feat by volitional fiat. He then deals with the possibility that psychological laws might be counterexamples to this assumption by simply defining laws of nature as "propositions that apply non-vacuously to things that are not rational agents."⁴² In short, van Inwagen assumes that the concept of a law of nature implies that we cannot render it false and then, on the basis of that assumption, excludes the possibility that laws involving psychological features could really be laws of nature.

Although Kant could not have been aware of van Inwagen's particular formulation of the argument, the position ascribed to him above provides him with several responses.⁴³ First, if one supposes that each noumenal self chooses only its own character, then van Inwagen's engineer-bureaucrat example would be inapplicable as a counterexample to Kant's position. For the engineer can choose only his own character and not what nature a spaceship can or cannot have. And even in that case, Kant need not be committed to the idea that there are no restrictions at all on what character the engineer can choose for himself.⁴⁴ As a result, Kant can accept the intuitive idea that *some* laws of nature impose limits on our abilities and are not up to us, but deny the much stronger claim that this is true of *all* laws of nature.

⁴¹ The scope of necessity is smaller than the scope of what is not up to us, since there may be contingent matters of fact, beyond what is necessary, that are not up to us.

⁴² *Ibid.*, p. 64.

⁴³ In the *Critique of Practical Reason*, Kant states the problem of free will and determinism in terms that are amazingly close to van Inwagen's. After asserting the incompatibility of freedom and determinism for transcendental realists, Kant provides the following justification: "For, from the first [i.e., the necessity of causal relations] it follows that every event, and consequently every action that takes place at a point of time, is necessary under the condition of what was in the preceding time. Now, since time past is no longer within my control, every action that I perform must be necessary by determining grounds *that are not within my control*, that is, I am never free at the point of time in which I act" (5:95).

⁴⁴ In *Religion within the Bounds of Reason Alone* (6:19–53), Kant explains what dispositions human beings may and may not have as a result of original sin and the primitive choices they make.

Second, Kant's prior commitment to idealism is also significant. For if one is committed to the idea that significant features of the world (such as space, time, and causality) are subject-dependent, then the idea that the laws of nature that govern causally interacting objects in space and time are subject-dependent as well may enjoy immediate intuitive plausibility. Granted, it is one thing to assert that something is subject-dependent, quite another to say that it is up to us or depends specifically on our wills. Still, if our wills are not to be completely irrelevant to the world, then they ought to be able to have some causal effect on the world, which is just to say that, in some respect, the world ought to depend on our will, that is, be up to us.

Third, it is important to note, however, that one need not go so far as to be a full-blooded idealist to avail oneself of the pivotal features of Kant's solution and to reject van Inwagen's argument. For what is crucial about Kant's position as it pertains to van Inwagen's argument is not its idealism or even the grounding thesis *per se*, but rather the idea that the physical laws of nature are not primary (or basic) insofar as they are contingent on *something else*. In Kant's case, that is, in the case of one who accepts idealism and that the laws of nature (or at least some of them) depend on us (in the form of the grounding thesis), it is the primacy of the physical (i.e., spatio-temporal) over the nonphysical (or intelligible) that is rejected. Accordingly, just because a certain course of action is not physically possible for me does not entail that I am not free or could not have done otherwise, since physical possibility pertains merely to the world of appearances, which depends, at least in part and to some extent, on my free noumenal choice. The point to note in this context, however, is that Kant's position suggests other, nonidealistic ways in which one could reject the primacy of the physical over the nonphysical. For example, a theist, such as Leibniz, could hold that the laws of nature are not basic and are, in fact, up to us insofar as they are based on God's free decrees that are partially based, in turn, on what human beings choose to view as good for themselves.⁴⁵

⁴⁵ Susan Wolf explores such a possibility in Chapter 5 of *Freedom within Reason* (New York: Oxford University Press, 1990), when, inspired by Leibniz's position, she suggests that one can reject the idea that "the physical level of explanation is more basic than the psychological level" in favor of the opposition position, namely that "the psychological level is more basic than the physical" (p. 110), which could be understood as a version of the idea that the laws of physics depend on what is up to us psychologically or, rather, morally speaking.

One may or may not find all of the details of Kant's story philosophically plausible. However, Kant is not aiming to motivate and justify a positive account of freedom. Instead, his ambitions are ultimately quite modest. He repeatedly insists that we cannot prove theoretically that we are free. In fact, quite often he says that we cannot even have insight into the possibility of freedom. Rather, his entire purpose is to show simply that the *impossibility* of freedom cannot be demonstrated. As we have seen, he argues that the only way to avoid the impossibility of freedom lies in distinguishing between things in themselves and appearances, asserting that things in themselves ground appearances and allowing for the mere possibility that certain appearances might be up to us insofar as we are able to choose our own characters. In short, he wants to argue that some of the distinctive features of Transcendental Idealism are necessary to avoid having to declare that freedom is impossible.

Impersonal and Personal Facts

Before turning to consider specific respects in which Kant's resolution of the problem of free will and determinism is relevant to his more general account of causation, consider one further dimension to Kant's views on freedom and determinism. One can begin with a purely philosophical point by noting that, strictly speaking, no formal contradiction arises between the deterministic claim that "no (non-initial) event could have been otherwise" (given the laws of nature and the initial conditions of the universe) and the claim made in the name of freedom that "I could have done otherwise." While it is true that the former claim ascribes necessity to events and the latter entails contingency about what I do, and thus that logically contradictory kinds of modality are ascribed in each case, it is not true that this difference in modality is ascribed to one and the same entity. For the necessity of determinism is attributed to events, that is, to what happens, whereas the contingency inherent in human freedom is ascribed to what I do, that is, to my actions, where there is an essential connection to me or my personhood.

There are a variety of contrasts that one might try to draw between events and actions. Perhaps the former are essentially passive (or neutral with respect to the active-passive distinction), whereas the latter necessarily entail some sort of activity. Perhaps events are ontologically self-sufficient entities, whereas actions imply some sort of subject that is distinct from the action being performed. In the current context, however, what is at issue is the difference between events being impersonal and

actions being “personal” in the sense that they belong to someone. For it is this difference that could allow one to accept the claim that every non-initial event is necessary (given the laws of nature and previous events) but to reject the view that my actions are necessary as well since they are not identical with (or completely reducible to) the events whose necessity is entailed by the laws of nature. For example, even if this arm, attached to this organic body, had to be raised, it does not follow that I necessarily had to raise it. Accordingly, a certain event may be necessary, but the action that is a product of that event and the attribution of whatever personal element that is required for it to be *my action* may not be necessary.

This kind of compatibilist position might appear quite attractive to Kant, given his view of the role that spontaneity plays in both our thinking and acting and his emphasis on “personal” concepts in a variety of passages throughout his Critical corpus. For example, in the Transcendental Deduction Kant famously asserts that the “I think” that must be able to accompany all representations for them to be mine is produced by a spontaneous act that does not belong to sensibility (B132). If spontaneity is thus crucial to a representation being mine, it would be structurally analogous to assert that spontaneity is likewise necessary if an event in the world is to be considered an action of mine. While Kant comes to recognize that one cannot *prove* in any theoretically acceptable way that we are free on the basis of such spontaneity, he could still find such a view attractive.

Also, in a lengthy passage from the *Critique of Practical Reason*, Kant repeatedly couches the issue of free will and determinism in terms of the concept of control (5:95–97). The concept of control would seem to represent a combination of the concepts of causality and personhood (among others). That the concept of control has causal connotations is perhaps sufficiently obvious that it does not require further comment. What might be less clear immediately is the idea that control involves personhood. However, we would not ordinarily say that impersonal objects such as billiard balls control each other in their various causal interactions. Rather, we would say that some person can control the motions of billiard balls.⁴⁶ Moreover, Kant thinks that the concept of control cannot be fully explicated solely through recourse to the laws of nature that hold for the phenomenal world. In the beginning sections of the Doctrine of Right in the *Metaphysics of Morals* Kant distinguishes between physical and

⁴⁶ We also sometimes attribute control in what is perhaps a derivative sense to intelligent animals.

noumenal possession as follows: “*intelligible possession (possessio noumenon)* must be assumed to be possible if something external is to be mine or yours. Empirical possession (holding) is then only possession in *appearance (possessio phaenomenon)*” (6:249). He then argues that

since the concept of a right is simply a rational concept, it cannot be applied *directly* to objects of experience and to the concept of empirical *possession*, but must first be applied to the understanding’s pure concept of *possession* in general. So the concept to which the concept of a right is directly applied is not that of *holding (detentio)*, which is an empirical way of thinking of possession, but rather the concept of *having*, in which abstraction is made from all spatial and temporal conditions and the object is thought of only as *under my control*. (6:253)

Insofar as rights presuppose freedom and are themselves possible only if that to which we have a right is under our control, it is clear that Kant recognizes the fundamental importance of the concept of control to that of freedom.

If Kant thus holds that spontaneous acts of our noumenal selves are necessary for us to have control over those phenomenal events that we view as ours, and if Kant’s solution to the conflict expressed in the Third Antinomy invokes the distinction between noumena and phenomena, does it follow that the distinction between the impersonal and the personal represents a further dimension to Kant’s ultimate solution to the problem of free will and determinism? And if so, could one not give up the idea developed above that the laws of nature are up to us in the sense that these laws, which hold only for the phenomenal world, depend on our noumenal choices?

The first question should, I think, be answered in the affirmative. After all, not only is the textual evidence in favor of interpreting Kant as holding such a position significant, but it also represents a *strength* of Kant’s position that he is able to give at least some account of what it is that makes my actions mine (rather than simply passing over the issue in silence, as is often done). This is not to say that Kant’s position has any definitive advantage here. Should others present alternative accounts, the relative merits of the case for each one can be taken up at that time. The point right now is simply that Kant’s account already has the resources to tell a plausible story on a potentially difficult philosophical issue.

However, the second question must be answered in the negative. Even if Kant’s account of the personal dimension of our actions is given due credit (both textually and philosophically), it is incapable of resolving the modal conflict that is central to the problem of free will and determinism.

That is, even if one grants that one can avoid an explicit contradiction between the claims of determinism and freedom by pointing out that necessity pertains only to impersonal events and not to the personal dimension implicit in the idea that *I* could have done otherwise, the personal-impersonal distinction does not give a robust enough sense of contingency. For if it is necessary that this arm be raised, the fact that it is not necessary that *I* raise this arm does not correspond to the full sense of contingency that is contained in our commonsense notion of freedom. To be free, it is essential that *I* could have decided not to raise my arm, letting it sit on the armrest instead. The notion of control that is central to Kant's reflections on how "personality" can be added to events so that they become "my actions" presupposes that *I* am causally efficacious with respect to the "impersonal" aspects of my actions as well, that is, over what events they are.

In one of the "incompatibilist" passages from the *Critique* cited above, we can see Kant making precisely this point:

Reason does not give in to those grounds which are empirically given, and it does *not follow the order of things as they are presented in intuition*, but *with complete spontaneity it makes its own order* according to ideas, to which it *fits* the empirical conditions and according to which it even declares actions to be *necessary that yet have not occurred and perhaps will not occur*. (A548/B576, italics added)

If our choice consisted merely in the fact that we could decide whether or not to identify with an event that was inevitably going to occur, then Kant could not say that we do not follow the order of the phenomenal world. As mentioned above, in the *Critique of Practical Reason*, too, we find Kant complaining that simply drawing a distinction between different kinds of determining grounds is not enough for freedom. In this context, he is objecting to a distinction between whether determining grounds are "*within* the subject or *outside* him" (5:96), but the point holds for the personal-impersonal distinction as well. The problem of free will and determinism remains problematic, he says, if

these determining representations have the ground of their existence in time and indeed in the *antecedent state*, and this in turn in a preceding state, and so forth, [for] these determinations may be internal and may have psychological instead of mechanical causality. . . . ; they are always *determining grounds* of the causality of a being insofar as its existence is determinable in time and therefore under the necessitating conditions of past time, which are thus, when the subject is to act, *no longer within his control* and which may therefore bring with them psychological freedom . . . but nevertheless natural necessity; and they therefore leave no *transcendental freedom*. (5:96–97)

Psychological freedom and the freedom to identify an event as one's action do not amount, Kant says here, to what is required for that event to be under one's control. As a result, though Kant's distinction between the personal and the impersonal is an important part of his account of free will and determinism, it is not in a position to solve that problem independently of the solution to the problem that is based on the grounding thesis.

FREEDOM AND CAUSALITY

Having considered the basic framework within which Kant situates the problem of free will and determinism (the Thesis and Antithesis arguments of the Third Antinomy) as well as both his general solution to this problem (Transcendental Idealism) and the specific elements that are necessary to resolving several particular issues that arise within this framework (the grounding thesis and its various dimensions), we are now in a position to compare Kant's position on causality as it is involved in freedom and his general views on causality as described in Chapter 4. Such a comparison focuses on four distinct points. The first point that must be established is that Kant's views on these disparate topics are in fact consistent. Establishing mere consistency might seem to be a trivial point. However, not all accounts of the laws of nature are consistent with Kant's specific resolution to the problem of free will and determinism. Also, the structure of Kant's account of laws of nature has several interesting implications. As a result, determining consistency ends up being surprisingly instructive.

Second, one might challenge the interpretation of Kant's general model of causality developed above on purely textual grounds, since there are several passages discussing the problem of free will and determinism in which Kant states that causes are events, which seems to contrast with the idea that the causality of the cause is an indeterminate activity. However, having attended to Kant's views on freedom we can now see that and why these passages should not be taken to reflect Kant's considered position. As part of a larger argument for Transcendental Idealism, they reveal the importance of understanding the distinction between things in themselves and appearances in terms of what is or is not completely determined. By holding that appearances are initially indeterminate and become increasingly determinate, though never end up being completely determined, one need not (and, in fact, cannot) be committed to causes being events qua fully determinate entities. Moreover, it also illustrates

how it is that the activity of determination can play a central role in Kant's account.

Third, by considering that Kant's resolution to the problem of free will and determinism invokes a noumenal choice and by reflecting on Kant's conception of substance, one might attempt to interpret Kant's model of causality as holding that phenomenal states are determined by noumenal rather than phenomenal substances, especially if phenomenal substances turn out, for various reasons, to be problematic on Kant's account. However, the primary arguments that might be developed against the possibility of phenomenal substances are far from definitive, and since invoking noumenal substances alone as what directly grounds the determination of phenomenal states entails the temporality of noumenal substances, the price to pay for such an interpretation turns out to be too high within the context of the overall goals of Kant's project in the *Critique*.

Fourth, because Kant's general model of causality employs concepts that extend beyond the notion of a bare event, Kant's account of freedom can appeal to those concepts in resolving two difficult problems that traditionally arise for free will. The notion of a determining ground that Kant invokes in his general model of causality can be used to stop the regress that might threaten our free actions if we were to try to explain them in terms of desires alone. Similarly, rather than attempting to locate free will in either a caused or an uncaused event (or in a complex desire or set of desires), one can view the freedom of the will in terms of a substance that acts so as to determine its effect. Accordingly, undertaking a comparison of Kant's general model of causality and his account of freedom turns out to be quite fruitful.

Consistency

As we saw in the previous section, Kant's solution to the modal conflict between free will and determinism depends on the possibility that the laws of nature are grounded in things in themselves in general and our free noumenal choices in particular. Kant makes this possibility more concrete in two ways. First, Kant holds that the laws of nature depend on the *natures* of things. Second, in the specific case of human beings, Kant wants to suggest – in line with our commonsense views – that we can choose our own characters or natures. These two points, taken together, allow Kant to say, contra van Inwagen, that the laws of nature could in fact be up to us. The contingency of free will is thus compatible with

the necessity of the laws of nature because our free will could choose (at least some of) the natures on which the laws of nature, along with their necessity, are based.

However, this solution to the problem of free will and determinism raises a question that might seem to be almost trivial, namely: Is Kant's general model of causality consistent with it? According to Kant's general model of causality as described in Chapter 4, a substance acts so as to determine an event (i.e., a change of state in a distinct substance) by exercising its causal powers. Such causal powers are exercised, in turn, according to the grounds that constitute the nature of a thing. Thus, when a cause produces its effect, a substance determines that effect according to its nature.⁴⁷ If we are considering a complex case of causality, such as mutual interaction, in which substances jointly determine each other's states, this model can be specified in such a way that the natures of both substances must be taken into account in explaining how they interact. If Kant maintains that a law of nature is a law that can specify which events will occur only by taking into account the natures of whatever substances are causally efficacious, then his general model of causality entails that substances must act according to the laws of nature. Kant can thus invoke natures in a way that is consistent with the account of natures that is required for his resolution to the problem of freedom and determinism.⁴⁸

To see that establishing the consistency of Kant's general model of causality with his resolution to the problem of free will and determinism is not an inconsequential point, consider how Kant's account of the laws of nature contrasts with an understanding of laws of nature that is based exclusively on events. Empiricists such as Hume typically understand laws of nature to be nonanalytic universal generalizations that are formed by taking a set of events that have occurred in the past and abstracting certain general features from them that are then supposed to cover events that will happen in the future. While such an account might be well suited to explaining how laws can perform certain epistemological tasks (e.g., can describe how such laws might be justified on the basis of empirical evidence we actually possess), it also faces significant objections. For example, it has been objected that this account of laws forces one to classify

⁴⁷ Recall from Chapter 4 that Kant understands the notion of determination as asymmetrical precisely in virtue of a rule based in the ground.

⁴⁸ It is true that the term "nature" is being used to apply to phenomenal and noumenal substances. However, the grounding thesis maintains that the former can depend on the latter, so that there is still a close relationship between the two uses.

what seem to be purely accidental universal generalizations as laws of nature. It has also been objected that the laws of nature should be necessary in some sense (whether logically or physically), whereas nonanalytical universal generalizations must be thoroughly contingent, since they are based on (contingently occurring) events.⁴⁹

It is immediately evident, however, that such event-based accounts of the laws of nature are inconsistent with Kant's resolution of the problem of free will and determinism. As we have seen, Kant can reconcile the necessity that follows from the laws of nature with the contingency implicit in our free will because the laws of nature are conditional on the contingent natures that we freely choose. However, accounts that do not accept such natures are unable to assert that the laws of nature are conditional on them. Event-based accounts of laws of nature do not accept such natures. Therefore, they cannot reconcile contingency and necessity in the way that Kant's account does. Since there are substantive accounts of laws of nature that do not invoke natures, it is a significant, even if unsurprising, fact that Kant's model of causality is consistent with his resolution to the problem of free will and determinism.

Moreover, because Kant's account of the laws of nature relies on the natures of substances rather than the determinate events that inhere in them, it is not subject to the main objections that empiricist conceptions face. For one, Kant is under no pressure to hold that accidental generalizations based on past events might function as laws of nature, since he bases the laws of nature on natures rather than events. For another, one can see how Kant would be in a position to explain the necessity of laws. Insofar as laws of nature are based on natures and natures are, in turn, constituted by the essential grounds of substances, necessity is, as it were, built right into the nature of things.⁵⁰

Accordingly, clarifying Kant's model of causality in response to what might have appeared to be an insignificant question about its consistency with his resolution of the modal conflict that arises in the context of the problem of free will and determinism ends up being quite revealing. For in addition to showing that not all accounts of the laws of nature are consistent with Kant's distinctive resolution of this conflict, it also illustrates how Kant's account of laws of nature has a set of philosophical

⁴⁹ These objections, and several others, are developed by David Armstrong in part I (chaps. 1–5) of *What Is a Law of Nature?* (New York: Cambridge University Press, 1983).

This issue is discussed in detail in Chapter 6.

⁵⁰ We return to this topic in Chapter 6.

implications that are fundamentally different from those of an event-based account. Specifically, Kant's account of the laws of nature can easily avoid the objections that face an event-based account. None of this establishes that Kant's account is intrinsically preferable to an event-based account, but it does uncover fundamental differences that would have to be considered in any such comparison.

Textual Troubles?

In the Third Antinomy, Kant states that "every action, as appearance, insofar as it produces an occurrence, is itself an occurrence, or event, which presupposes another state in which its cause is found" (A543/B571). Likewise, in the *Metaphysics Mrongovius* transcripts, Kant says: "In all appearances of an event the causality of the cause of the event is itself an event. Now if all causes themselves have causes, then there is nothing in the world except nature. Now since there is nothing in the sensible world except events, we can go to infinity; everything that we will experience will still be either event or effect. For were it not an event, it would not be an object of experience at all" (29:860). Finally, later in the *Metaphysics Mrongovius* transcripts, Kant asserts that "in the sensible world the causality of an event is itself an event, for in the sensible world everything happens. Therefore, no totality is in the series of conditions here. The causality of an event is also itself an event. For had it been causality at all times, then the event would have been at all times. But that contradicts the concept of an event" (29:923). In short, Kant repeatedly asserts that the causality of the cause is an event, which seems to contradict Kant's general model of causality according to which the causality of the cause is an indeterminate activity by which a substance determines an event. Is Kant's general model of causality consistent with these passages?

Here one must concede that these passages are in fact inconsistent with Kant's general model of causality. However, it is imperative to recognize that all three of these quotations occur in contexts in which Kant is discussing the problem of freedom and determinism, are therefore simply part of his larger argument, and do not actually reflect his own position. More specifically, in all three cases Kant also presents a counterpoint to the idea that every cause must itself be an event. For example, later in the paragraph from which the last quotation is taken Kant speaks of noumenal freedom as follows: "Thus such a noumenon indeed acts as appearance according to the mechanism of nature, [but] its actions do not

happen thus, but rather according to its will and not by the mechanism of nature. Its causality of an event is *not itself an event*" (29:924, emphasis added). Similarly, following the other passage from the *Metaphysics* Mrongovius transcripts we find Kant acknowledging: "If an event ensues from a cause which is no event, then it is said to occur spontaneously [*sponte accidit*] from it" (29:861). Thus, in the very passages that contradict our interpretation of Kant's general model of causality, we find Kant contradicting these contradictory remarks by explicitly stating that not every cause is itself an event. As a result, these passages are best interpreted as expressing an antinomial conflict, which makes it necessary to view Kant's own position as different from both of the contradictory views being expressed in them.

What is important about this second set of passages, though, is not simply Kant's denial that every cause must be an event (which would simply counterbalance the weight of the initial quotations), but also what philosophical resources he can draw on in support of that denial that would still make sense of the first set of quotations. Part of his reason for denying a necessary link between causality and eventhood is that the causality of freedom is noumenal. Since noumena are necessarily atemporal and events essentially temporal, a noumenal cause cannot be an event. This point is of course not purely accidental, since the primary aim of ascribing freedom to the noumenal world derives from the latter's atemporality; if noumena were temporal, then they would have to be causally determined, which would, in turn, preclude the possibility of freedom that was at stake in the first place.

However, another, more significant part of Kant's reason for denying the connection between causality and eventhood stems from the conceptual space that is opened up by the way in which Transcendental Idealism distinguishes between phenomena and noumena in terms of indeterminate and determinate entities. If one accepts Transcendental Realism, then one is committed to the view that what appears to us in space and time must be fully determinate, since what is real must be fully determinate. However, if what appears in space and time must be fully determinate, then causes, as what must also appear in space in time, must be fully determinate events, too. Once one has admitted causes as fully determinate events, then (1) one is naturally led to adopt a conception of laws as generalizations based on fully determinate entities, (2) the constraints on accounting for the possibility of freedom are increased (so much so that Kant thinks that freedom becomes impossible), and (3) more generally, there would seem to be no indeterminacy in events that causal powers,

whether freely exercised or not, could remedy by means of their activity of determination.

If, by contrast, one accepts Transcendental Idealism by drawing a distinction between things in themselves and appearances and asserting that only things in themselves can be fully determinate, then appearances must be indeterminate in some respect. Once one has admitted indeterminacy in appearances, there is room to claim both that events can become determinate through the causality of a phenomenal cause and that the causality of this cause can itself be indeterminate and thus not a temporally determinate event. As a result, because Transcendental Idealism opens up room for understanding Kant's model of causality as invoking temporally indeterminate activities rather than temporally determinate events, we can see that this doctrine gives us resources that help us to understand in greater detail the general model of causality Kant develops in the Analogies of Experience.

Accordingly, by taking into account Kant's views on freedom, the details of Transcendental Idealism in particular, we are able to understand properly certain passages that contradict Kant's model of causality. Since they are, when taken in context, counterbalanced by assertions of the opposite, an antinomial conflict arises, one that can be resolved by Transcendental Idealism's distinction between appearances and things in themselves. Thus, instead of being problematic, these passages point to the resources that are necessary to form a more accurate understanding of that model (1) by showing why one need not be committed to understanding causes as fully determinate entities, (2) by revealing the centrality of the distinction between determinate and indeterminate entities, and (3) by indicating the fundamental importance of the role that determination can play in Kant's overall account in light of this distinction.

Phenomenal Substances and Noumenal Causality

Having introduced the importance of Transcendental Idealism as a means of attaining a better understanding of Kant's general model of causality, might it not also open up the possibility of viewing the substances that stand in causal interaction on that model as noumenal rather than phenomenal? For if one appeals to the noumenal causality of freedom to show that not all causality can be understood in terms of events, why not construe Kant's general model of causality in terms of noumenal causality as well? Textually, it is quite striking that Kant at least uses the same phrase ("the causality of the cause") in both cases, so one might

naturally take him to be referring to the same thing in each case. So introducing Transcendental Idealism indirectly raises the further question: Why not simply eliminate phenomenal substances from Kant's model of causality altogether?⁵¹

One could garner support for this suggestion by developing independent arguments against phenomenal substances.⁵² First, one might object to the very idea of phenomenal substances as contradictory. If "substance" is defined along traditional Cartesian lines as what is capable of independent existence (and is thus nonderivative or fundamental), while "phenomenal" entails subject-dependency (or at least that something appearing to be fundamental is in fact derivative), then a phenomenal substance would be something that is both independent of anything else (including us) and also dependent on us as subjects (or on something distinct from itself), which is a contradiction. As a consequence of this line of argument, only what is noumenal could truly be a substance, and Kant's assertions regarding phenomenal substances are to be understood accordingly, namely as asserting that they are either merely "comparative" substances or what appear (or can be thought) to be substances without actually being substances in the true sense of the term.⁵³

Second, one might pursue the same conclusion by means of a slightly different line of argument.⁵⁴ That is, if one defines substance as an absolute subject (which would also be capable of independent existence), then the only properties that such a substance can have would be intrinsic properties, since only intrinsic properties are compatible with the possibility that only that substance might exist. Since Kant repeatedly claims (e.g., at B66–67 and A277/B333) that all phenomena consist entirely in relations, it follows that phenomenal substances are impossible. For phenomenal substances would have to have intrinsic properties that at the same time consisted entirely in relations, which is a contradiction. That

⁵¹ Allen Wood ("Kant's Compatibilism," pp. 88–89) seems sympathetic to this line of interpretation when he suggests that Kant's account of causality divides into an Aristotelian, noumenal account that invokes activity as a fundamental notion and a Humean, phenomenal account that is primarily directed at empirical regularities. As shall become clearer below, the interpretation I favor holds that certain aspects of "Aristotelian" (or perhaps also Leibnizian) activity pertain to the phenomenal realm as well.

⁵² Karl Ameriks, *Kant's Theory of Mind* (New York: Oxford University Press, 1982), and Rae Langton, *Kantian Humility* (pp. 53–63), have developed arguments along these lines.

⁵³ One could support such a line of argument by appealing to Kant's *Reflexion* 5312 (18:150) and Baumgarten's *Metaphysica*, esp. §§193, 201, and 233.

⁵⁴ Langton argues along these lines (*Kantian Humility*, pp. 49–67). Ameriks rejects such an interpretation ("Kant and Short Arguments to Humility").

Kant still uses the term “phenomenal substance” is to be explained on the grounds that the permanence of certain relations can perform functions that lead us (albeit mistakenly) to reifying them into substances. As a result, only what is noumenal can truly be a substance.

Third, this general line of interpretation seems to offer a neat solution to a puzzle that arose in the context of explaining mutual interaction. As we saw above, mutual interaction between two substances entails that the grounds that constitute the natures of these substances must *jointly determine* their states as simultaneous. If, however, grounds must jointly determine the states of substances, then it might seem as if Kant’s argument is circular (insofar as it presupposes simultaneity in order to establish simultaneity) or problematic (insofar as the jointness of the grounding relationship presupposes the temporal determinacy of the grounds, which is illegitimate in the context of that argument). The current line of interpretation might seem to be able to avoid the whole puzzle as follows. Because noumenal substances are atemporal, by holding that they stand in mutual interaction in such a way that their phenomenal states are simultaneous, the grounds that constitute their essence can still be outside of time.⁵⁵ Since they are not temporal and thus do not presuppose either simultaneity or any temporal determinacy at all, one can avoid the charge that Kant’s argument is either circular or problematic in the way that was suggested.

However, as attractive as such an interpretation might appear to be in these respects, it encounters objections that ultimately rule it out as an interpretation of Kant’s position. One can start by noting that the very idea of a phenomenal substance need not be contradictory. For one can grant that Kant develops a notion of substance that applies only to noumenal substances but still deny that it is his only conception of substance or that it represents a core notion that must underlie all other senses. In fact, if Kant has any core notion of substance at all, it would be that of something that can only be a subject, never a predicate. Yet that notion is univocal between noumenal and phenomenal substances. The other notion of substance that might be considered fundamental to Kant’s philosophy is one that he adopts from Leibniz (and several of his followers, such as Wolff and Baumgarten), namely that activity is to

⁵⁵ Langton seems to adopt such a position when she concludes that “forces are not substances, but properties, relational properties, of substances unknown in themselves” (*Kantian Humility*, p. 63). She does not, however, consider in any detail how this view would be consistent with Kant’s argument in the Third Analogy or with the problematic discussed above.

be equated with substantiality. Once again, however, there is no obvious connection between activity and independent existence that would entail that this core notion of substance contradicts the idea of phenomenal substances.⁵⁶

If Kant is not committed to a core notion of substance that is inconsistent with phenomenal substances, does Kant's characterization of phenomenal substance in particular contain a hidden contradiction? Kant's most prominent characterization of phenomenal substance is in terms of permanence, and so no contradiction appears to arise on that count. Moreover, the notion of activity that Kant identifies with substantiality in the Second Analogy actually cuts against the idea that one could dispense with phenomenal substances. For if activity entails substantiality, then the causal activity of phenomenal objects (e.g., the *Prolegomena's* sun warming a body, 4:312) would seem to entail the existence of phenomenal substances. Moreover, Kant's introduction of both noumenal and phenomenal natures (in the guise of noumenal and phenomenal characters, where Kant explicitly stipulates that a thing's character is "a law of its causality, without which it would not be a cause at all," A539/B567) also tells in favor of phenomenal causality. As a result, Kant is committed to accepting rather than rejecting phenomenal substances.⁵⁷

The most important objection to an interpretation that dispenses with phenomenal substances by asserting that all causal interaction is between noumenal substances, however, is that it compromises the atemporality of noumenal substances. If phenomenal substances are actually relations that merely appear to be substantial (in virtue of their permanence), while noumena are the ultimate subjects in which those relations must inhere, then the changes that occur in the phenomenal world are actually changes in the relational properties of noumenal substances. But if noumenal substances are changing their (relational) properties, then they are clearly temporal insofar as change entails temporality.⁵⁸ Thus,

⁵⁶ One should keep in mind here Kant's explicit identification of substance and activity in the Second Analogy (A204/B249–50).

⁵⁷ While one can point to Baumgarten's use of phenomenal substances in a pejorative sense and passages in Kant's corpus that either are derivative on that usage or stem from his pre-Critical period, I am aware of no fully Critical text that unambiguously rejects phenomenal substance.

⁵⁸ One might object that change does not entail temporality if the changes are what are sometimes called "Cambridge changes," as opposed to "real changes." See Geach's discussion of Cambridge changes in *Truth, Love, and Immortality* (London: Hutchinson, 1979), p. 90ff. Moreover, if "real" changes are defined in terms of changes in intrinsic properties, Kant's claim that phenomena consist entirely in relations would entail that all

noumenal substances would be temporal, contrary to Kant's explicit requirement that noumena be atemporal.⁵⁹

One could attempt to avoid this objection by taking recourse to the conception of substance described above that identifies noumenal substances with ultimate subjects insofar as they have *intrinsic* properties. By understanding substances in this way, if it is the case that the intrinsic properties of such ultimate subjects do not change, then it would be correct to assert that noumenal substances are atemporal in some significant sense. However, two problems arise with such an interpretation. First, such an understanding of noumenal substance is incompatible both with what Kant takes to be an analytical claim, namely that noumenal substances can belong to a single world in virtue of their (causal) relations to each other, and with the grounding thesis, since both rely on the idea that specifically noumenal substances stand in relations to other entities (whether it be other noumenal substances or the phenomenal world). Second, if the ultimate subjects of intrinsic properties are also the ultimate subjects of relational properties, then one has failed to avoid the original problem. For if the relational properties are changing, then so are the ultimate subjects that have them. The only way to avoid this conclusion is to deny that relational properties require any ultimate subject, but this denial amounts to conceding that relational properties are properties of something that is not an ultimate subject. However, in that case it is difficult to see why one is not thereby admitting the possibility of phenomenal substances. Accordingly, in developing this line of interpretation, one is forced either to admit the temporality of things in themselves, which is an extremely costly admission in the context of Kant's

phenomenal changes would be Cambridge changes and thus that things in themselves (with their intrinsic properties) need not change in spite of phenomenal Cambridge changes. Three points are important in response. First, without a clear conception of what Cambridge changes are and thus a precise account of whether the subjects of such changes are temporal or not, it remains unclear whether the objection stated above carries any force. Second, in the context of Leibnizian metaphysics, it is clear that all relational changes must be reflected in the changes of monadic or intrinsic properties, so that there would be no such things as Cambridge changes on such an account. Finally, in many (and perhaps all) cases Cambridge changes occur when the intrinsic properties of one thing change without a change in the intrinsic properties of another to which it is related, with the Cambridge change in such a case being the change in the relation between the two things. However, since the objection stated above requires that all phenomenal changes are Cambridge changes, the objection violates the principle that Cambridge changes require real changes.

⁵⁹ One should also keep in mind the textual evidence from the Second Analogy described in chapter 4, according to which the activity of the cause is described as continuous.

ultimate philosophical aims, or to concede that phenomenal substances are an important element in Kant's story after all.

As a result, both textual and philosophical arguments present significant challenges to the idea that one could simply dispense with phenomenal substances as what stand in mutual interaction according to Kant's model of causality. However, it is important to note precisely what is being denied here. It is not being denied that Kant accepts noumenal substances insofar as ultimate subjects are, it seems, required for the existence of phenomenal states. Nor is it being denied that noumenal substances are causally efficacious with respect to phenomenal states; the grounding thesis is simply one expression of such a thesis. Rather, what is being denied is that noumenal causality can take over all of the tasks of phenomenal causality so that phenomenal substances could be dispensed with altogether.

Determination and the Problems of Regress and Location

In a variety of contexts we have seen the importance of the notion of determination that Kant employs throughout his philosophy. For example, we saw that Kant's general model of causality employs the notion of an activity of determination to explain the asymmetrical rule-governed relation that obtains between a ground (which contains the rule) and a determination (which follows from that rule in the relevant circumstances), and one can see that such a notion might play an analogous role in Kant's views on freedom. For it is entirely in line with common sense to say that in acting freely one *determines* one's own actions. In fact, one might even say that if one determines one's own actions in accordance with a rule that one has freely adopted, then there is a genuine sense in which one has acted autonomously. We also saw that Kant's notion of determination is to be understood in terms of Transcendental Idealism's distinction between fully determinate noumena and indeterminate appearances, since Kant understands determination as an activity that renders something indeterminate determinate. We now turn to consider how Kant's notion of determination is also relevant to two problems that traditionally arise with respect to the issue of freedom.

In discussions of free will and determinism, one problem that is often raised concerns how to stop the regress that ensues from attempting to explain a putatively free action. If one seeks to know why that action was performed, typically a desire is offered up in response. However, it might seem that for any desire that is cited, one can ask for a further explanation

of that desire. If a second desire is given in support of the first, it seems to be possible to ask for an explanation of the second desire, and so on. For example, if I take my children to school and am asked why I do so, my response is likely to involve my desire that my children become well educated. If I am then asked why I want my children to be well educated, I am likely to answer that I want them to lead a happy and interesting life (and hold the belief that being well educated is generally conducive to leading a happy and interesting life). Of course, this latest desire stands in need of explanation, too. Since any desire one gives seems to require an explanation in terms of yet further desires, an interminable regress (of desires) seems unavoidable. Accordingly, in light of what one might call the regress problem, no satisfactory explanation of a freely performed action appears to be possible.⁶⁰

Kant is aware of the regress problem in roughly this form. In fact, toward the end of his resolution of the Third Antinomy, Kant provides a response to it in the following terms:

It, reason, is present to all the actions of human beings in all conditions of time, and is one and the same, but it is not itself in time, and never enters into any new state in which it previously was not; in regard to a new state, reason is **determining** but not **determinable**. Therefore, one cannot ask: Why has reason not determined **itself** otherwise? But only: Why has it not determined **appearances** otherwise through its causality? But no answer to this is possible. For another intelligible character would have given another empirical one. (A556/B584)

Two features of Kant's notion of determination are crucial to his response to the regress problem. First, if reason, or rather a noumenal cause, is a determining ground of an action rather than something determinable, then it becomes illegitimate to ask why the ground of the action is what it is. For if it is not determinable, that is, cannot be determined by anything else, then there is nothing external to it that could explain why it is the way it is. One *can* ask why reason determines its phenomenal *actions* the way it does, but Kant wants to argue that we cannot answer this question, not because the question is illegitimate, but because *we* cannot know what would be needed to answer it. As Kant remarks in the following paragraph: "But why the intelligible character gives us exactly these appearances and this empirical character under the circumstances before us, to answer this surpasses every faculty of our reason" (A557/B585). That is, because a

⁶⁰ For an interesting description of this problem (though it goes under a different name), see Thomas Nagel, *The View from Nowhere* (New York: Oxford University Press, 1986), chap. 7, esp. p. 117.

different intelligible character would have caused a different empirical character, and we can have no knowledge of what intelligible characters there are, we cannot know why the intelligible character that we have causes these appearances rather than other ones.

Second, while part of Kant's reason for asserting that one cannot explain why grounds are the way they are is due to the fact that they are determining rather than determinable grounds, it is also important to recognize the significance of Kant's claim that determining grounds are not *temporally* determinate. In the case of freedom, the reason is clear. If the determining grounds were temporally determinate, then there would be further grounds that determined them to be the way they are at that time, which would preclude the possibility of freedom. Yet one can see that an analogous line of reasoning holds for Kant's general model of causality. If the grounds of simultaneous states were temporally determinate, then they would require further grounds that determined them to be the way that they are at that time. However, invoking further grounds runs contrary to the way that the initial grounds are supposed to function, for in that case the further grounds rather than the initial ones would be what ground the simultaneity of the states. Accordingly, Kant's answer to the regress problem reveals a completely general philosophical point, namely that determining grounds cannot themselves be temporally determinate if they are truly to function as grounds.

Kant's solution to the regress problem also corresponds to how common sense would naturally deal with examples in which the regress problem arises. While one might look for a further desire that would explain my desire for my children to lead happy lives, one could instead suggest simply that it is the kind of desire that a responsible parent would have. This suggestion could stop the regress because what has been provided is not yet another desire but rather a characteristic of a certain kind of person. While one might ask what caused me to have that characteristic – in which case the regress would not have been stopped at all, but rather simply transferred from the level of desires to that of character traits – it is also possible to understand the response as invoking the fact that I am a certain kind of person, namely a responsible parent. That is, that characteristic (among others) is constitutive of who I am, so that it is illegitimate to ask why I have it, since that reduces to the question of why I am me.⁶¹ Kant would interpret this example in terms of his model of

⁶¹ In this example, it is not essential that I am a parent (since I could have been childless), but only that if I am a parent, that I am a responsible one.

causality such that invoking the fact that I am a certain kind of person instead of appealing to desires amounts to providing an explanation in terms of determining grounds rather than determinate states. Since determining grounds involve the essence of a thing, the problem of regress ought not arise insofar as they are designed to function as ends of explanation rather than intermediary links, which corresponds precisely to the role played by the appeal to the kind of person I am in the example at hand.

Kant's notion of determination is also relevant to a second problem that often arises in the context of free will and determinism. The question is sometimes raised about where we should locate free will (provided that we have it at all). If one locates it in an event that is caused by a previous event (or set of such events), then it becomes unclear that the will is still free. After all, there is some event that causes it to be what it is and if that event is caused by a previous one, and so on, then it looks as if the ultimate source of the action (and thus of the responsibility for it as well) lies outside us and is thus not something that we control. However, if – to cut short the chain of causes that eventually leads outside of us – one locates it in an event that is not caused by a previous event at all, that is, in an event that is entirely uncaused, then it might appear that such an event is just a random occurrence and not something for which *I* should be held responsible.⁶² Since it seems problematic to locate our free will in either caused or uncaused events, we have what one might call a location problem.

One popular way of responding to the location problem has been to say that I can still be responsible for events that are caused if and only if those events are caused by desires with which I identify. For example, Frankfurt distinguishes between first- and second-order desires (the desire for X and the desire to have the desire for X), on the one hand, and second-order volitions (the desire that the desire for X be one's will), on the other. On Frankfurt's initial account, I identify with an action if it is caused by, or at least in accordance with, a second- or higher-order volition.⁶³ This account thus attempts to solve the location problem by means of a structural feature of our desires; free will should be located at a particular juncture (namely at the level of second- or higher-order volitions) in a

⁶² For discussion of this particular problem, see *Agents, Causes, and Events*, ed. T. O'Connor (New York: Oxford University Press, 1995).

⁶³ Harry Frankfurt, "Freedom of the Will and the Concept of the Person" and "Identification and Wholeheartedness," in *The Importance of What We Care About* (New York: Cambridge University Press, 1988), pp. 12ff. and 166ff.

complex hierarchy of desires.⁶⁴ However, Frankfurt came to realize that the details of his initial solution were not entirely adequate. As he notes, “the assignment of desires to different hierarchical levels does not by itself provide an explanation of what it is for someone to be *identified* with one of his desires rather than with another,” since even a complex desire is still just a desire and no explanation has been given as to why I should be identified with higher- rather than lower-order desires.

To explain identification more clearly, Frankfurt offers two further ideas. First, a “person, in making a decision by which he identifies with a desire, *constitutes himself*.”⁶⁵ Accordingly, identification now entails the self-constitution of a person. Second, Frankfurt explains the self-constitution of a person by means of acts of identification that occur when conflicts between desires (regardless of what order they are) are resolved either by rejecting certain ones or by ordering them such that some have priority over others. As Frankfurt puts it: “It is these acts of ordering and of rejection – integration and separation – that create a self out of the raw materials of inner life.” Frankfurt thus attempts to solve the location problem not as he had earlier, by according privileged status to, say, second-order volitions over first-order desires, but rather by spreading out a certain structural or functional property (namely that of a priority ordering) over the entire hierarchy of desires.⁶⁶

Despite the obvious sophistication of such a solution to the location problem, one might find it problematic on the following grounds. Frankfurt’s account ultimately posits acts of ordering and rejection that constitute the self. However, what is it that performs these acts? The most natural answer would be that these acts are performed by the self that identifies with its actions and desires. However, this answer is not available to Frankfurt, given that the self is constituted by the acts it is supposed to perform. That is, there is a kind of inconsistency in the two claims that form the basis of Frankfurt’s solution to the location problem. The first claim asserts that the self is not antecedently given, but rather must first be constituted, while the second claim states that what constitutes the self are acts of ordering and rejection. The inconsistency arises because the

⁶⁴ One might think of this solution as employing a foundationalist strategy insofar as the highest-order desire that is the cause of an action determines what one identifies with.

⁶⁵ Frankfurt, “Identification and Wholeheartedness,” p. 170.

⁶⁶ One might think of this response as adopting a coherentist strategy insofar as it does not attempt to isolate a particular base-level (even if fluctuating) order of desires as the foundation of identification, but rather tries to spread identification out over all desires (in the form of their coherence with one another and some claim of priority).

acts of ordering posited in the second claim presuppose what the first claim asserts we do not have, but rather need to constitute in the first place, namely the very self that the acts of ordering are posited to create.

Notice, however, what the ultimate source of the location problem is for Frankfurt's account. While the eventual problem it encounters is that the self must be both prior and posterior to the acts that constitute it, it faces this problem only because it makes the empiricist assumption that desires and relations among them must be the fundamental building blocks out of which the self and its concomitant notion of identification must be constituted. Since desires are simply events, Frankfurt's account implicitly accepts the terms of the dilemma that underlies the location problem – that free will must be located in an event of some sort, whether caused or uncaused – and simply tries to address the challenges that then arise in selecting the one option.

In light of this analysis, it is clear that Kant's model of causality and the notion of determination employed in it can offer a different way of thinking about these issues, a way that allows one to avoid the dilemma that gives rise to the location problem. For, as we have seen above, in addition to events, Kant's model invokes substances that determine these events. In the case of freedom, one can say more specifically that the substance that determines events is really an agent who freely determines its own actions.⁶⁷ Accordingly, instead of trying to constitute the self out of previously existing events/desires and running into the problem of how to constitute the self without implicitly relying on it, Kant can deny that the self needs constituting at all and view it as a distinct kind of entity, an agent, who is prior to its states or actions as what underlies and determines them in the first place.⁶⁸ As a result, Kant's solution to the location problem is to locate free will not in any event, but rather in an agent (or substance) who determines his or her own actions.

In the case of free will, moreover, Kant holds that the self (as a certain kind of substance) not only causes a certain event to occur (by contributing to the selection of the laws of nature from which that event then follows), but is also responsible for the personal dimension such an event

⁶⁷ Although Kant specifically argues in the Paralogisms (1) that we cannot *know* that the self is an immaterial substance that retains its personal identity over time and (2) that rationalist *arguments* attempting to establish such knowledge are fallacious, it is clear from the *Critique of Practical Reason* that Kant is ultimately quite sympathetic to such a *position* as something we might believe on practical grounds.

⁶⁸ This makes Kant's solution similar in significant respects to what an Aristotelian might say on this issue.

would have as *my action*, as something over which I think of myself as having *control*. Since I have control over my actions, there is no need to fear that they would be random (as might be the case with the uncaused events that form one horn of the dilemma of the location problem). Kant's account thus has resources with which to respond to one of the main objections that is raised against views that his own position resembles in certain respects, while still avoiding the dangers of the location problem.

As a result, by drawing on features of Kant's general model of causality, especially the notion of determination that it employs, we see how he can develop plausible responses to philosophical problems that arise in the context of free will and determinism. The regress that seems to arise in explaining a free action in terms of desires can be stopped (and the regress problem solved) because Kant's general model of causality employs the notion of something other than a desire, namely that of a determining ground, which cannot be explained by anything external to itself, since qua determining ground it is not determinable, that is, cannot be determined by anything other than itself. The location problem can likewise be solved by appealing to a notion employed in Kant's general model of causality other than that of an event, namely, the notion of a substance that determines its actions. For insofar as an agent (or self) that freely determines its actions can be a substance, it is not an event that could be determined by anything else and thus does not run the risk of being determined by something else in such a way that it must abdicate responsibility for its actions. Accordingly, Kant's notion of determination can not only be employed in both his account of freedom and his general model of causality, but also be used to solve a series of traditional problems and explain a variety of phenomena that are not easily incorporated into a single coherent philosophical account.

CONCLUSION

If Kant's account of freedom, which is absolutely central to his entire Critical project, were not consistent with the general model of natural causality described in Chapters 3 and 4, we would be faced with a difficult decision, since it is equally clear that causality is of paramount importance to him as well. Fortunately, we have found that Kant's accounts of natural causality and freedom are not only consistent, but even employ many of the same concepts and principles, even if they must be modified to be appropriate to their particular context. Central to both accounts,

for example, is the notion of determination, since both accounts explain, albeit in different ways, how appearances must initially be indeterminate and can then become determinate in some respect by means of the activity or determining ground of a substance (whether the substance be phenomenal or noumenal) in accordance with its nature.

We have also seen just how powerful the set of concepts and principles that Kant employs in these accounts can be. For one, they can explain how to resolve the modal conflict between contingency and necessity that is one central aspect of the problem of free will and determinism. For, according to Kant's account, it is possible that by exercising its causal powers, a substance might be able to choose (some aspect of) its own nature, which influences in turn which laws of nature hold and thus which events are necessary in accordance with them. For another, the notion of a substance or determining ground provides Kant with the resources to develop an agency theory that is in a position to explain where to locate our free will (namely in the agent rather than in any desire or hierarchy of desires) so that an infinite regress of explanation can be avoided. Accordingly, exploring the metaphysics of freedom allows us to see the considerable strengths of the fundamental concepts that Kant employs in his general model of natural causality.

Kant's Reply to Hume

Historical and Contemporary Considerations

INTRODUCTION

Now that we have a comprehensive and detailed description of Kant's views on causality, we can turn to consider how they are related to the views of others. Since Kant repeatedly relates his views on causality to Hume's, it is natural to begin with the historical question of what Kant's reply to Hume is in light of the account of causality attributed to Kant in the previous chapters. While it is commonly assumed (especially in the context of the Second Analogy of Experience) that Kant's reply to Hume is (at least supposed to be) a direct refutation of Hume's position, I argue for three contrary theses in the first half of this chapter. First, when the reception of Hume in Germany is taken into account, we see that Kant would have been justified in assuming that the majority of his readers (especially those interested in what "pure reason" can establish) would not have thought that a refutation of Hume's views on causality was at all necessary. Second, Kant's explicit references to Hume in the first and second editions of the *Critique* suggest that Hume's views on causality were important to him not primarily in their own right, but rather as an illuminating illustration of Hume's more general skeptical approach, which, due to its inherent instability, should be replaced with his own Critical methodology. Finally, when Kant's and Hume's models of causality are compared, one can see that they are so radically different that they do not share enough assumptions for a refutation to be possible. Instead, one should view Kant as attempting to develop an alternative account of causality, one that competes against rather than refutes Hume's views.

However, Kant's views on causality can also be shown to be relevant to contemporary philosophical discussions. In the second half of this chapter, I show that Kant's views can be used to clarify certain specific issues pertaining to the nature of causal powers (as distinct from event causation), necessitarian conceptions of laws of nature (as contrasted with regularity-based conceptions), and agent causation (as opposed to belief-desire models of action). I conclude by suggesting that Kant's contribution to contemporary discussions does not end with the clarification of points pertaining to specific issues, but can also be viewed in terms of the development of a comprehensive metaphysical picture of the world that, by being based in the notion of activity (or its correlate notion of determination), represents a systematic alternative to empiricist (e.g., Humean) positions, just as Kant himself had done with respect to Hume over two centuries ago.

THE HISTORICAL QUESTION OF KANT'S REPLY TO HUME

While a great deal of attention was devoted to the reception of Leibniz's views on causality in Part I, Hume's views, by comparison, were given short shrift. It is true that Kant's immediate reaction to Hume (in 1762–1764) was discussed at some length. However, the implications of this reaction were not considered in detail, and Kant's relation to Hume in the *Critique* was mentioned in Part II only in passing. It is now time to take a somewhat broader look at Hume's importance to make good on this deficit and to put us in a position to provide an answer to the historical question of what Kant's reply to Hume is on the issue of causality.

Although a comprehensive treatment of the reception of Hume extends beyond the scope of this study, a selective portrayal of how Hume's views on causality were received in Germany in the 1750s, 1760s, and 1770s can still be quite useful. For considering how figures such as Johann Georg Sulzer and Johann Nicolas Tetens understood and reacted to Hume provides a fuller context for appreciating the pre-Critical Kant's relation to Hume presented in Chapter 2.¹ Moreover, our description of these figures' reactions to Hume can also help to put us in a position to address the historical question of how to understand Kant's reply to

¹ For the reception of Hume in Germany, see Günter Gawlick and Lothar Kreimendahl, *Hume in der deutschen Aufklärung: Umrisse einer Rezeptionsgeschichte* (Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1987), and Manfred Kuehn, *Scottish Common Sense in Germany, 1768–1800: A Contribution to the History of Critical Philosophy*, with a Preface by Lewis White Beck (Kingston and Montreal: McGill-Queen's University Press, 1987).

Hume in the *Critique*, since it gives us a more accurate sense of what kind of response to Hume Kant and his intended readers would have thought necessary.

German Reactions to Hume: Sulzer and Tetens

Consider how Hume's *Enquiry Concerning Human Understanding* first appeared in German translation in 1755.² It was included as the second part of Hume's *Vermischte Schriften*, for which Johann Georg Sulzer, the leading Wolffian at the Academy of Sciences in Berlin, served as editor.³ In some respects this edition is not especially noteworthy. The translation itself is neither outstanding nor filled with egregious flaws that would significantly distort Hume's intentions or preclude the development of otherwise viable interpretations. The anonymous translator has still not been identified; in the preface, Sulzer states simply that it came to him "from good hands" and that he "examined it very carefully against the original" so that he can assure his readers that it accurately represents Hume's views.⁴

However, two other features of this edition are remarkable for current purposes. First, in his preface Sulzer explains that one of his primary reasons for publishing the translation of the *Enquiry* is not the philosophical content Hume advances – the positions articulated and the arguments adduced on their behalf – but rather Hume's philosophical style.⁵ To clarify this point Sulzer describes a fundamental contrast between two ways of attaining and presenting philosophical truths. The one way, which was prominent in Germany at the time, is based on reflection, is thorough, and is therefore long and arduous, but has the virtue of displaying analytic certainty every step of the way. The other, which was not well represented in Germany at the time and of which Hume's *Enquiry* is a prime example, relies on common sense or instinct, enables quick insight, and is reliable

² Hume's *A Treatise of Human Nature* was not translated into German (in its entirety) until after the publication of the *Critique*.

³ This volume appears as the first of seven in *Reception of the Scottish Enlightenment in Germany: Six Significant Translations, 1755–1782*, ed. H. Klemme (1755; rpt., Bristol: Thoemmes Press, 2000).

⁴ Sulzer's Preface is unpaginated. All quotations lacking references to page numbers are from the Preface.

⁵ Sulzer's other reason stems from the fact that German philosophers are in danger of suffering the effects of a drawn-out "philosophical peace"; their philosophical weapons have become dull or rusty for lack of challengers. Sulzer hopes that publishing Hume's *Enquiry*, which is clearly the work of a "skeptical mind," will "wake these philosophers up a bit from their lazy [*müßigen*] quietude and give them a new activity."

enough in most cases, but must sacrifice absolute certainty for its easy accessibility. Because Hume follows the latter method, he is able to “lead his readers into the most hidden and obscure depths of philosophy through a path that is easy, pleasant and, as it were, strewn with roses.” The state of philosophy in Germany is quite different; it “is not in fact lacking in great philosophers [such as Leibniz and Wolff], but German attire does not appear befitting to philosophy as its inner beauty requires.” As a result, Sulzer recommends that Germans attempt to appropriate Hume’s accessible style in order to present to a wider audience the “deepest secrets of philosophy,” secrets that he believes philosophers such as Leibniz and Wolff have already discovered.

Second, Sulzer attaches substantive critical commentaries to each section of the *Enquiry*.⁶ In the preface he claims that his comments do not amount to a “proper refutation,” but rather simply present thoughts that readers “might find useful in examining Hume’s views.” Still, his critical remarks are intended at least to vindicate his contention that it is at most Hume’s style and not his distinctive doctrines that ought to be adopted. In line with this expectation, Sulzer’s critique of Hume’s *Enquiry* starts in §2 with a rejection of Hume’s distinction between impressions and ideas and ends in §12 with the remark that “we have seen, through the entire course of this work, so many and such clear instances [*Proben*] of the author regularly contradicting his own doctrines that it would seem that he is concerned less with the truth than with finding it entertaining to say something that contradicts the customary opinions of philosophers.”⁷ Sulzer’s general attitude is thus clear: Hume’s *Enquiry* does not propose fertile philosophical doctrines that would be worth developing in greater detail and extending in novel directions; rather, one should admire how artfully Hume presents his views to the general public.

Sulzer’s general views on the content of Hume’s philosophy are also reflected in his treatment of the issue of causality. In his remarks on §4 of the *Enquiry*, for example, he claims that Hume fails to distinguish (as Leibniz and Wolff do) between absolute and conditional necessity and that, *pace* Hume, necessity of the latter sort obtains for matters of fact. In support of this claim, he argues that if we know that a cause is given, then we also know that its effect must necessarily follow.⁸ Further, he claims

⁶ Sulzer’s comments are nearly as long as Hume’s *Enquiry* itself.

⁷ Hume, *Vermischte Schriften*, p. 374.

⁸ *Ibid.*, pp. 92–93. He also adds that we know that the magnitude of the effect must correspond to the magnitude of the cause. “It would be superfluous should I want to demonstrate the certainty of these principles in detail” (p. 94).

that we make so many such inferences concerning matters of fact that we are typically unaware of the entire chain of reasoning we go through. Thus, the fact that Hume takes to establish the impotence of reason – namely that when we are presented with an entirely new object, we are unable to know prior to experience, and therefore simply on the basis of reason, what effects it can bring about – can be explained in a way that is consistent with the view that reason *can* have insight into causally necessary connections (even if it does not happen to have such insight in any explicit way in most ordinary cases).⁹ While this kind of alternative explanation obviates the need for Hume's positive explanation of causality in terms of custom or habit, Sulzer develops an independent criticism of such an explanation in his remarks on §5. According to Sulzer, custom, as an irrational and hence unintelligible principle within the mind, is tantamount to an occult quality and should thus be rejected.¹⁰ He supports this criticism by arguing, in line with his critical remarks on §4 and his widespread agreement with Leibniz's position, that every change in the mind must be a consequence of concepts or representations that it has had previously.

Sulzer continues his attack on Hume's views on causality with his commentary on §7 of the *Enquiry*. In §7 Hume implicitly assumes, Sulzer argues, that one cannot be aware of a power without also distinctly representing the specific *way* in which it produces its effect. Sulzer's criticism of this assumption is based on a more general argument for the existence of powers that stems from Leibniz and Wolff.¹¹ The idea is that since there must be something in the world at one moment in time that causes it to change to its state at the next moment of time, powers must exist in the world (since there must be a reason for the change). This general line of thought is then related to particular cases as follows. While admitting that we do not have an impression of the primitive forces that constitute bodies, Sulzer thinks that we do have an impression of both the derivative forces of bodies and, even more clearly, a force or power within the mind.¹² In the case of our minds, for example, he notes that we cannot still our thoughts for long. Try as we might, thoughts will soon emerge, and the power of the mind is what is responsible for the

⁹ *Ibid.*, p. 96.

¹⁰ *Ibid.*, p. 131.

¹¹ See Wolff, *Rational Thoughts*, §32.

¹² Hume, *Vermischte Schriften*, p. 187.

emergence of these thoughts. Sulzer's point is even more plausible for the case of (macroscopic) bodies. Even if one does not know all the details of a clock's mechanism, one can still know that it has the power to tell time. As a result, we can know powers without being aware of exactly how they operate, *contra* Hume's assumption in §7.

In sum, while Sulzer repeatedly praises Hume's acuity of intellect and presentation style, he does not find the position Hume endorses at all tempting. Neither Hume's general empiricist position nor his skeptical attitude toward causality is attractive to him. In terms of content, he finds Leibniz's and Wolff's rationalist positions much more plausible and repeatedly uses their views to develop his criticisms of Hume. Despite his critical attitude toward the content of Hume's philosophy, Sulzer does still find value in Hume's philosophical style in the *Enquiry*. Hume's "easy and pleasant" method, which makes even the deepest of truths accessible to all, is something that German authors should try to emulate in their attempts at reaching a broader audience.

Sulzer's interest in and reaction to Hume was no isolated incident. As part of their heightened interest in British philosophy in general, German thinkers were continuously engaged with Hume throughout the late 1750s, 1760s, and 1770s. However, as was the case with Sulzer, their engagement with Hume did not lead them, by and large, to accept his position.¹³ Moses Mendelssohn, who paid quite a bit of attention to Hume (e.g., to Hume's views on probability), did not reject Leibniz's and Wolff's basic metaphysical principles so as to accept Hume's skeptical position; rather, he attempted to synthesize the empirical detail found in empiricist writers such as Hume with the broadly Leibnizian philosophical framework he had already accepted.¹⁴ Gotthold Ephraim Lessing was interested enough in Hutcheson to translate the latter's *System of Moral Philosophy* in 1756, only a year after it was originally published, despite the fact that his own position ended up being notoriously close to Spinoza's.¹⁵

¹³ Prior to the 1780s, Hamann is perhaps the most noteworthy exception to this general statement, and even in his case there are radical differences between Hume's actual position and Hamann's appropriation of it for his own (religious) purposes.

¹⁴ On the issue of causality, Mendelssohn argued – surprisingly, from a contemporary perspective – that Hume's account of causality was actually already present in Leibniz and Wolff's philosophy. See Manfred Kuehn, "Mendelssohn's Critique of Hume," *Hume Studies* 21 (1995): 197–220.

¹⁵ See Frederick Beiser, *Kant and the Fate of Reason* (Cambridge: Harvard University Press, 1989), for a discussion of Lessing's relation to Spinoza and of what consequences this would have in Germany in the 1780s.

Moreover, interest in British authors is clearly documented by the fact that their major works were being translated into German shortly after their publication. Adam Smith's *The Theory of Moral Sentiment* was translated by Christian Rautenberg in 1770, James Beattie's *An Essay on the Nature and Immutability of Truth* was published in German translation in 1772, while Christian Garve translated Smith's *Wealth of Nations*, Edmund Burke's *Observations on the Sublime and Beautiful*, and Adam Ferguson's *Institutes of Moral Philosophy*. Still, one cannot identify a significant group of German thinkers in the late 1750s, 1760s, and 1770s who could be considered *followers* of these authors, even if they were very interested in their works.¹⁶

One important exception to this general reaction, however, is Johann Nicolas Tetens. Specifically, in his *Philosophical Essays on Human Nature and Its Development* (*Philosophische Versuche über die menschliche Natur und ihre Entwicklung*, 1777) Tetens adopts the "method of observation," which he ascribes to Locke, and repeatedly discusses Hume's position without immediately rejecting it on the basis of prior commitments (as Sulzer had, in effect, done).¹⁷ It is also relevant to note that Tetens's philosophical abilities did not go unnoticed. Kant repeatedly praised him and, in marked contrast to several of his other reviewers, thought Tetens capable of understanding and perhaps even carrying his own Critical project out to completion.¹⁸ Given Tetens's openness to empiricism and the direct connection between Tetens and Kant, Tetens's reaction to Hume is worth investigating more closely.

The most striking feature of Tetens's views is that, despite his sympathies with empiricism in general and Locke's observational method in particular, he does not follow Hume on the issue of causality. While admitting that there is much that is correct in Hume's account of causality, in

¹⁶ For an excellent discussion of these issues, see Kuehn's *Scottish Common Sense in Germany, 1768–1800*, esp. pp. 36–51. Though Kuehn classifies Garve (along with Lambert, Tetens, and the pre-Critical Kant) as one of the "critical empiricists," he immediately points to severe limitations to the accuracy of this label. This group "is less coherent than any of the other three groups previously mentioned [Berlin Enlightenment figures, common sense philosophers, and sensationists]. In fact, it is not really a group at all" (p. 46). Though Beiser refers to Garve as "essentially an empiricist" and a Lockean (*Kant and the Fate of Reason*, p. 178), he likewise admits that this description must be significantly qualified.

¹⁷ Tetens is often referred to as the "German Locke" due to his adherence to the method of observation. While the term has perhaps stuck, few endorse its accuracy, since a careful reading of his *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* reveals that Tetens is indebted to Leibniz and Wolff at least as much as he is to Locke.

¹⁸ See, for example, Kant's remarks at 10:232 and 10:341, but also 10:270.

essay four of his *Philosophical Essays*, Tetens claims that “Hume overlooked one of its [i.e., the concept of causality’s] essential components, which at the same time served as the primary occasion for his making the same mistake with respect to the entirety of human cognition, and, because he was not aware of its inner strength, he believed that he could make it falter through his skeptical sophistry.”¹⁹ After recounting Hume’s analysis of the concept of causality in terms of “constant succession,” Tetens explains which “essential components” Hume is missing:

From the representation of a **constant succession** [*einer beständigen Folge*] of the one upon the other he derives our entire concept of the causation of the one **through** the other? But we represent it to ourselves as if **the effect depended on the cause**, were **produced** by it, and made actual **through** it. Does not this last way of representing it contain further ideas beyond constant succession? We view the effect as something that is **intelligible** [*begreiflich*] on the basis of its cause! (p. 316)

That is, Tetens rejects Hume’s empiricist conception of causality by arguing that an analysis of our customary concept of causality reveals that it must also contain (1) the idea that the effect *depends* on its cause (or that the cause *produces* its effect) as well as (2) the idea that this dependency (or production) relationship ought to be *intelligible*.

Since Hume fails to identify these two elements of our concept of causality, it falls to Tetens to offer an account of them. Tetens provides such an account (or at least certain elements of it) in the broader context of his discussion of primitive relations. In the section that immediately precedes his discussion of Hume (quoted above), Tetens explains how we represent the relations that hold between absolutes that are given to us from without (e.g., sense data) as follows:

When we view two things as **identical**, when we think of them as standing in **causal** relations, when we represent one thing in another as features in a subject, or both of them at the same time next to each other or as following each other, there is a certain **act of thinking**, and the relation in us that we are thinking is something **subjective** that we attribute to the objects as something **objective** and that arises from the thinking. These acts of thinking are the **first original thoughts of relations**. (p. 303)

Tetens’s idea is that the content of the concept of causality arises from (1) an activity of thinking by means of which our ideas are related and (2) from attributing the subjective relation that holds between our ideas

¹⁹ Johann Nicolas Tetens, *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* (Leipzig, 1777), pp. 312–313.

to the relation that obtains between the objects so represented. While the idea that we transfer a subjective relation to external objects may derive from Hume, what faculty does Tetens think is responsible for representing this relation? The imagination (as a Humean might argue) or the understanding (as a Leibnizian might hold)?

In the section that then follows his discussion of Hume's account of causality, Tetens identifies three kinds of primitive relations: those of comparison (identity and diversity), those of combination (temporal relations such as simultaneity and succession), and causal relations. Tetens adopts the first two kinds of relations from Leibniz, explicitly referring to the *New Essays*.²⁰ And in this context Tetens makes it clear that causal relations must be represented by the understanding, not the imagination.²¹ But it is significant that Tetens introduces causal relations as a third, distinct kind of primitive relation. Thus, rather than relying exclusively on Hume, Locke, or Leibniz, Tetens is attempting to work out an independent position, where his account of causality forms simply a part of his more general theory of relations that involves the faculty of the understanding.

Moreover, Tetens seems to be aware of the implications that follow from conceiving of causality as a distinct primitive relation brought about by an activity of the understanding. For example, he realizes that if causal relations are distinct from relations of identity, then they cannot be based on the principle of identity or contradiction.²² While discussing Hume's account, Tetens explicitly draws this consequence by noting that "the understanding combines [causes and effects] according to a habitual law of thinking that it follows, although it does not follow it with the irresistible force had by those [laws] that the understanding assumes in thinking the necessary truths of reason, e.g., the principle of contradiction."²³ That is, Tetens sees that the kind of necessity that causal relations have is different from logical necessity and that it cannot therefore be explained in the same way that logical necessity is.

Tetens's explicit denial that causal relations could be based solely on the principle of contradiction and his mention of "habit" in this context might seem to move his view closer to Hume's, according to which it is the

²⁰ *Ibid.*, p. 331.

²¹ This interpretation is reinforced by his initial account of relations, since he asks how these relations are to be *thought*, not *imagined*, and then definitively established by various passages that directly assert that it is the understanding that thinks causal relations.

²² Tetens could be familiar with this point from Crusius.

²³ Tetens, *Philosophische Versuche*, p. 320.

imagination that is responsible for the representation of causal relations. However, in addition to his ubiquitous use of the faculty of understanding in accounting for the concepts of relation in general, Tetens explicitly distances himself from such a position: "Let us also observe the connection of thoughts when we say: 'We understand a consequence from its principles.' Is it not clear that **deriving, drawing a conclusion, and inferring one truth from another** is a connection of ideas that is essentially distinct from associations in our fantasy?"²⁴ But if our knowledge of causal relations is based neither on the principle of contradiction nor on the imagination's habits, what is its foundation? Tetens simply remarks that "such general thoughts are true thoughts, prior to all experience."²⁵ Since he never turns this idea into a fully developed argument, it is plausible to assume that he is merely reiterating the idea, familiar from Crusius and the pre-Critical Kant, that certain concepts or principles are derived not from experience per se, but rather from the understanding as an innate material principle.

Tetens summarizes his reaction to Hume's account of causality as follows:

First, it is certainly not the **mere succession of impressions** upon each other from which the conception of a **causal** connection is derived. Rather, there are certain special kinds of associations of ideas from which it is abstracted, and, in fact, those in which something more is noticed than that one idea occurs and that the other then follows it. Undoubtedly we initially get this concept from the feeling of our own striving and its effects. . . .

Second, we transfer this concept, which we have gotten from our feeling of ourselves, to external objects. . . .

Third, the understanding can derive the concepts of a **ground** (*ratio*) and of **what is grounded** in it and of the **intelligibility** of the latter on the basis of the former only from the activity of its understanding, of inferring and deriving.²⁶

What we thus see is Tetens attempting to chart a middle course between Hume and Leibniz. On the one hand, while he can agree with Hume that a constant conjunction can serve as an indicator or sign of causality, he clearly rejects what he takes Hume's view to be, namely that it constitutes the entire content of our concept of causality.²⁷ Leibniz and his followers

²⁴ Ibid., p. 322.

²⁵ Ibid., pp. 320–321.

²⁶ Ibid., pp. 323–325.

²⁷ Ibid., p. 315. Whether Tetens's view is fair to Hume is a separate question. For discussion of how Hume should be understood, see, for example, Barry Stroud, *Hume* (London: Routledge & Kegan Paul, 1978); John Wright, *The Septical Realism of David Hume*

are right to insist that the concept of causality contains the notions of dependence and intelligibility and must be represented by an activity of the understanding, an activity that is distinct from what the imagination can accomplish. On the other hand, Tetens rejects the Wolffian idea that causality might ultimately be based solely on the principle of contradiction, which means that Tetens has not specified how the intelligibility of causal relations is to be understood. That is, Tetens does not have ready at hand an explanation that would satisfy those who deny that there is any such intelligible connection or those who think that there could be several different alternative ways of understanding such a connection.

This brief and highly selective description of the reception of Hume's *Enquiry* in Germany from the 1750s through the 1770s reveals a number of points. First, Hume's views became known through an edition that immediately put his position in a less than favorable light. For Sulzer makes it clear why Hume's distinctive doctrines, both in general and on the issue of causality, ought to be rejected. Hume's literary style is attractive, but the content of his philosophy is not. Moreover, this attitude toward Hume was common to many leading philosophers in Germany at the time, such as Lessing and Mendelssohn. Finally, even Tetens, who is much more sympathetic to empiricism than most others at the time, rejects Hume's position on the issue of causality. For one, his analysis of our concept of causality reveals two significant nonempirical components that he thinks Hume overlooks. For another, he argues that causal relations must be (1) represented by the understanding (rather than the imagination, as Hume would have it) and (2) understood as a primitive relation that is distinct from both temporal relations and the relations of identity and diversity, whose necessity is based on the principle of contradiction. Although Tetens's explanation may leave unspecified what his justification of the source of causal relations is supposed to be, that in no way diminishes the differences he sees between Hume's position and his own.

By thus considering how Hume was received in the 1750s, 1760s, and 1770s in Germany, we find that most of those figures who were interested in Hume's thought were not tempted by the philosophical content he advanced. Sulzer, who was responsible for bringing Hume to the attention of

(Minneapolis: University of Minnesota Press, 1983); Richard Fogelin, *Hume's Skepticism in the Treatise of Human Nature* (Boston: Routledge & Kegan Paul, 1985); Galen Strawson, *The Secret Connexion: Causation, Realism, and David Hume* (New York: Oxford University Press, 1989); Ken Winkler, "The New Hume," *Philosophical Review* 100 (1991): 541-579; and Don Garrett, "The Representation of Causation and Hume's Two Definitions of 'Cause,'" *Nous* 27 (1993): 167-190.

a wider German audience, objects to several of Hume's central philosophical doctrines (including his views on causality). Tetens, who is otherwise extremely sympathetic to empiricism in general, explicitly rejects Hume's account of causality, since he thinks it obvious that Hume's account omits central components of our concept of causality. As a result, neither Kant nor his readers would have thought that Hume's position stood in need of refutation in the first place. While considering in detail only Sulzer and Tetens's reactions to Hume is of limited scope, they can still be viewed as representative of what German philosophers thought was significant about Hume's views on causality at the time, and thus provide us with a context in which Kant's views can be understood.

Kant and Hume in the *Critique*

Before we can turn to the question of how Kant intends to reply to Hume on the issue of causality in the *Critique*, however, we must establish one final piece of the puzzle. At this point, several of the requisite pieces are already in place. First, we understand how Kant initially becomes aware of Hume in his pre-Critical period. For sometime after 1755 Kant sees Hume as posing a serious challenge not to rationalism per se, but rather to the idea that causal relations between independently existing substances are logically necessary, an idea that was one of the cornerstones of the position that he had developed very early in his pre-Critical period in opposition to Leibnizian views. By 1764 (at the latest) Kant revises his views by distinguishing between logical and real grounds and by arguing that causal relations are to be understood in terms of the latter rather than the former. While these revisions allow Kant to avoid the immediate problem that Hume's statements entail for his position, he also recognizes that it simply relocates the problem, since real grounds, unlike logical grounds, cannot be explained by means of the principle of contradiction and thus stand in need of an explanation in terms of some new principle. In his search between 1764 and 1781 for an account that could describe and explain this new principle and also address other dimensions of his increasingly comprehensive metaphysics, Kant explores an extremely far-reaching set of issues – the distinction between sensibility, the understanding, reason, the nature of judgment, cosmological antinomies, psychological paralogsms, potentially fallacious theistic proofs in natural theology, and Transcendental Idealism.

Second, we now see how other German philosophers understood and reacted to Hume in the 1750s, 1760s, and 1770s. Sulzer and Tetens (along with most others at the time) did not accept Hume's argument against

understanding causality in terms of objectively necessary connections and thus did not feel compelled to follow him in attempting to reduce causal relations to constant conjunctions of events where, by force of habit, we come to expect an instance of one kind of event whenever we encounter an instance of another. Instead, they held fast to our commonsense conception of causality, according to which a cause produces its effect in an intelligible way, and attempted to provide an account of how this was to occur (with Sulzer relying on Leibniz and Wolff, and Tetens attempting to develop an independent position). Accordingly, Kant's immediate (pre-Critical) reaction to Hume is not out of line with the reaction of his contemporaries. In sum, few are tempted to accept the account of causality Hume adopts, but they do recognize that one must provide an account of the nature of causal relations in light of the challenges that Hume's skeptical remarks pose.

Third, and most significantly, we now have a detailed story about Kant's arguments and positions regarding causality in the *Critique*. Kant's Second and Third Analogies argue that causality and mutual interaction are necessary for experience, that is, for knowledge of a single world, since they make possible knowledge of the successive and coexistent states of the world. Moreover, the model of causality they presuppose does not consist of events, but is rather based on causal powers. One substance can cause an effect (or change of state) in another substance just in case the one acts on the other, or exercises its causal powers, in accordance with its nature and its external circumstances. In the case of mutual interaction, the exercise of causal powers must be joint, since simultaneity is a reciprocal temporal relation.

Kant's Explicit Remarks about Hume in the Critique

But what is still missing is an appreciation of Kant's explicit attitude toward Hume in the *Critique* and of how it fits in with these other pieces of the puzzle so as to form a single, coherent response to Hume. While it is true that Kant seems to have Hume in mind in several passages in the *Critique* (the Transcendental Deduction, the Second Analogy, and throughout the Transcendental Dialectic in the Paralogisms, Antinomies, and the Ideal of Pure Reason), it is quite striking that he makes explicit mention of Hume in only three passages in the first edition and that his additional references to Hume in the second edition are designed primarily to complement what is said in the first edition. Kant mentions Hume once on the final page of the *Critique*, where he lays out which *scientific* methods reason could pursue – Wolff's dogmatic method, Hume's skeptical

method, and his own critical method – in order to recommend his own. The two earlier passages, which set up this final reference, both occur in the Doctrine of Method's Discipline of Pure Reason in Its Polemical Use. Kant's treatment of Hume in this section of the text proves instructive for understanding his implicit attitude toward Hume in the rest of the *Critique* and thus toward Hume's views on causality as well.

Kant's primary goal in the Discipline of Pure Reason is to describe how pure reason should limit its propensity to stray beyond its proper boundaries in its various possible uses. Thus, in the first section of the Discipline of Pure Reason, "The Discipline of Pure Reason in Dogmatic Use," Kant argues that, despite certain similarities, philosophy cannot imitate the method that mathematics uses so successfully, and thus that pure reason in its dogmatic or "merely speculative use" (A736/B764), that is, the use of reason when it is not restricted to what can be given in intuition or possible experience, is not legitimate. In the second section of the Discipline of Pure Reason, "The Discipline of Pure Reason in Its Polemical Use," Kant then attempts to establish that there cannot be any polemical use of pure reason, either. Whereas the dogmatic use of reason would consist in reason asserting the truth of a certain proposition, the polemical use of reason would consist in "the defense of [pure reason's] propositions against dogmatic denials of them" (A739/B767). That is, there are two different kinds of dogmatic uses of reason, with one (the dogmatic use proper) consisting in the assertion of a certain claim (and in the presentation of arguments that would justify it) and the other (the polemical use) consisting in the defense of such a claim against attacks that might be raised against it. Kant argues that the polemical use of reason is improper because it falsely presupposes that the two combating parties can "conduct a dispute about a matter the reality of which neither of them can exhibit in an actual or even in a merely possible experience" (A750/B778). Kant illustrates this point poignantly as follows:

There is accordingly no real polemic in the field of pure reason. Both parties fence in the air and wrestle with their shadows, for they go beyond nature, where there is nothing that their dogmatic grasp can seize and hold. Fight as they may, the shadows that they cleave apart grow back together in an instant, like the heroes of Valhalla, to amuse themselves anew in bloodless battles. (A756/B784)

Despite the impossibility of making progress in such fashion, Kant recognizes a significant positive value to the polemical use of pure reason.

Thus let your opponent speak only reason, and fight him solely with weapons of reason. For the rest, do not worry about the good cause (of practical reason), for

that never comes into play in a merely speculative dispute. In this case the dispute reveals nothing but a certain antinomy of reason, which . . . must necessarily be heard and examined. The conflict cultivates reason by the consideration of its object on both sides, and corrects its judgment by thus limiting it. What is in dispute here is not the **matter** but the **tone**. For enough remains left to you to speak the language, justified by the sharpest reason, of a firm **belief**, even though you must surrender that of **knowledge**. (A744/B772)

That is, even if pure speculative reason cannot attain *knowledge* of the objects that most interest it and thus must rest content with *belief* supported by *practical* reason, the polemical use of pure reason can still cultivate reason, leading it to see that it can correct itself by limiting the domain over which it judges. It is in this context that Kant explicitly praises Hume for his critique of natural religion, since “he rightly holds that its object lies entirely beyond the boundaries of natural science, in the field of pure ideas” (A746/B774). This is the second passage in which Kant explicitly refers to Hume in the first edition of the *Critique*.

Yet Kant’s agreement with and praise of Hume can extend only so far. For, in Kant’s view, Hume’s use of reason is ultimately not *critical*, but rather merely *skeptical*.²⁸ Immediately after concluding that there is no real polemical use of reason, Kant argues that “there is also no permissible *skeptical* use of pure reason” (A756/B784, emphasis added). Kant explains that

for reason to leave just these [skeptical] doubts standing, and to set out to recommend the conviction and confession of its ignorance, not merely as a cure for dogmatic self-conceit but also as the way in which to end the conflict of reason with itself, is an entirely vain attempt, by no means suitable for arranging a peaceful retirement for reason; rather it is at best only a means for awaking it from its sweet dogmatic dreams in order to undertake a more careful examination of its condition. (A757/B785)

In short, if the ultimate aim of Hume’s skepticism is neither to combat the dogmatic use of reason nor to cultivate reason for some other purpose, but rather simply to document the ignorance of reason, it cannot be viewed as a solution that satisfies reason’s own interests.²⁹

Kant discusses this issue at greater length in a separate subsection entitled “On the Impossibility of a Skeptical Satisfaction of Pure Reason That Is Divided against Itself.” Here, in the third passage that explicitly

²⁸ In addressing Hume’s skepticism, Kant is focusing on Hume’s skepticism about reason, rather than his skepticism about either the senses or induction.

²⁹ As Kant puts it a few pages later, “skepticism is a resting-place for human reason . . . but it is not a dwelling-place for permanent residence” (A761/B789).

mentions Hume, while repeatedly praising the value of Hume's skepticism for combating dogmatism, Kant raises two main objections to viewing skepticism as a stable solution of reason's interests. First, Kant argues that reason, whose "momentum is not in the least disturbed, but only hindered [by skeptical attacks], does not feel that the room for its expansion is cut off, and although it is annoyed here and there it can never be entirely dissuaded from its efforts" (A768/B796). That is, even if Hume has raised some significant skeptical arguments against specific claims that reason might make, these arguments do not by themselves entail that reason cannot establish any claims at all and thus do not keep reason from making other assertions. Second, because Hume, "while rightly denying to understanding what it really cannot accomplish, goes further and disputes all its capacity to expand itself *a priori* without having assessed this entire capacity, the same thing happens to him that always brings down skepticism, namely that he is himself doubted, for his objections rest only on *facta*, which are contingent, but not on principles that could effect a necessary renunciation of the right to dogmatic assertions" (A767/B795). In other words, censuring whatever dogmatic claims happen to be made in a given context may be appropriate, but such a reactive procedure is too contingent to provide a clear indication of what reason can and cannot establish in principle. One consequence of this, Kant notes, is that Hume's skeptical attacks are susceptible to self-referential doubts, since nothing keeps reason from calling them into question in turn.³⁰

If the dogmatic use of reason is rightly called into doubt by the skeptical use of reason, but skepticism does not represent a stable position either (since nothing can keep it from being used against itself), how does Kant think that this impasse can be resolved? He explicitly addresses this question as follows:

The first step in matters of pure reason, which characterizes its childhood, is **dogmatic**. The . . . second step is **skeptical**, and gives evidence of the caution of the power of judgment sharpened by experience. Now, however, a third step is still necessary, which pertains only to the mature and adult power of judgments, . . . which subjects to evaluation not the *facta* of reason, but reason itself, as concerns its entire capacity and suitability for pure *a priori* cognitions; this is not the censorship, but the **critique** of pure reason, whereby not merely **limits** but rather the determinate **boundaries** of it . . . are not merely suspected but are proved from principles. (A761/B789)

³⁰ In fact, Hume himself notes this consequence in the section of *A Treatise of Human Nature* entitled "Skeptical Doubts Concerning Reason."

Hume's skepticism rightly shows that pure reason, in the face of its own dogmatic assertions to the contrary, cannot attain knowledge of objects that cannot be given to our senses (i.e., there are limits to what reason can know). However, because Hume investigates the "*facta*" of reason (i.e., what reason *does*) rather than the faculty of reason itself, he cannot delineate in a principled way what reason can and cannot know, and since he cannot demarcate the "determinate boundaries" that circumscribe its knowledge, it is open to the objections just noted.³¹ Accordingly, Kant holds that reason can attain a stable philosophical position solely by means of a critique of the faculty of pure reason, an investigation whose primary intent is to address neither particular dogmatic claims nor skeptical attacks on them, but rather what the very nature of our cognitive powers are and what kind of a priori knowledge they are capable of attaining as a result.³²

In the course of carrying out his investigation of the nature and boundaries of our cognitive powers in the Doctrine of Elements, Kant discovers that our cognitive faculties must be reconceived in fundamentally new ways. He finds that we must have distinct cognitive faculties – sensibility, understanding, and reason – each of which brings about a different *kind* of representation – intuition, concept, and idea. In light of his analysis of knowledge, Kant then argues that both intuitions and concepts are required for knowledge and thus that substantive knowledge cannot be gained by means of the use of pure reason's ideas. He can also provide an account of a priori knowledge on this basis. Kant's distinctive idea here is that even if we cannot attain knowledge of what lies entirely beyond

³¹ Kant twice provides an analogy between the shape of the earth and the nature of Transcendental Idealism that is intended to illustrate the distinction between our knowledge having limits and its having determinate boundaries. In one passage Kant notes: "If I represent the surface of the earth (in accordance with sensible appearance) as a plate, I cannot know how far it extends. But experience teaches me this: that wherever I go, I always see a space around me in which I could proceed farther; thus I cognize the limits of my actual knowledge of the earth at any time, but not the boundaries of all possible description of the earth. But if I have gotten as far as knowing that the earth is a sphere and its surface the surface of a sphere, then from a small part of the latter . . . I can cognize its diameter and, by means of this, the complete boundary, i.e., surface of the earth, determinately and in accordance with *a priori* principles; and although I am ignorant in regard to the objects that this surface might contain, I am not ignorant in regard to the magnitude and limits of the domain that contains them" (A759/B787). See also A762/B790.

³² Kant describes the contrast between the dogmatic and skeptical uses of reason and the critical use of reason further as follows: "[T]he critique is not involved in these disputes, which pertain immediately to objects, but is rather set the task of determining and judging what is lawful in reason in general" (A751/B779).

experience, we can still have knowledge that does not depend on particular experiences. His justification for this claim lies in the fact that we can apply the understanding's concepts not only to empirical intuitions, but also to a priori intuitions due to their relation to possible experience: "although of course we can never **immediately** go beyond the content of the concept which is given to us, nevertheless we can still cognize the law of the connection with other things *a priori*, although in relation to a third thing, namely **possible** experience, but still *a priori*" (A766/B794). It is thus not only *actual* experience (as Hume thought), but also *possible* experience that provides warrant for knowledge. Specifically, outside of mathematics, it is possible experience that makes synthetic a priori knowledge possible, since possible experience is what justifies the a priori connection between concepts that are not related by means of identity.³³ However, despite Kant's disagreement with Hume about the importance of possible experience, he agrees that we cannot have knowledge of the objects of traditional metaphysics, since they lie beyond *possible* experience, and the conclusion follows that traditional metaphysics is not possible (at least not in the sense in which it is ordinarily thought to be possible, namely as a body of knowledge, as opposed to belief).

Thus, in the Doctrine of Method Kant is attempting not to develop a general refutation of Hume's position, but rather to show how it can be improved on in several significant ways. Kant grants that the skeptical use of reason is valuable in combating its dogmatic use, but since skepticism is neither principled in its criticisms nor stable on its own, he views the task at hand to be to find a stable position that can demarcate what reason can know in principle. Kant argues that this task can be accomplished only if we undertake a critique of pure reason, that is, an investigation of the faculty of reason itself. After Kant charts our cognitive faculties in a radically new way so that both actual and possible experience can contribute to our knowledge, he is in a position to explain how synthetic a priori knowledge is possible and to decide on the fate of traditional metaphysics (where it turns out that he can agree with the content of many, if not all, of Hume's skeptical conclusions).

These links that Kant draws in the Doctrine of Method between the inadequacy of skepticism, the necessity of a critique of reason, and the possibility of synthetic a priori knowledge as it relates to metaphysics

³³ See also A154/B193ff. for a discussion of how possible experience makes synthetic a priori knowledge possible, and especially A155/B194, where Kant explains that time, imagination, and apperception are the sources of a priori representations.

also provide the context for the passages referring to Hume that Kant added in the second edition of the *Critique*.³⁴ For example, at B19 in the Introduction, Kant reiterates the connection between the possibility of metaphysics and the status of synthetic a priori knowledge. From the perspective that Kant describes in the Discipline of Pure Reason it is clear that Hume, who came closest among his predecessors to seeing metaphysics as standing or falling with synthetic a priori knowledge, was correct to focus on the principle of causality as a synthetic principle (A722/B750). However, due to the absence of a proper analysis of our cognitive faculties, Hume could not see (1) that the principle of causality can in fact be established a priori (due to the way in which it makes experience possible) and (2) that it is just one of several instances of synthetic a priori claims that we can make (since, as Kant points out at A767/B795, there are other principles, such as that of persistence).

The second mistake kept Hume from seeing that his denial of synthetic a priori knowledge entails a denial of the proper status of pure mathematics (and also of pure natural science). That is, if Hume had recognized that principles other than that of causality might be synthetic a priori, then he would have been in a position to view pure mathematics and natural science as synthetic a priori as well.³⁵ The first mistake led Hume to infer that metaphysics as the science of what can be known a priori can be nothing more than a delusion of reason. After Hume drew this skeptical conclusion about reason and hence about the possibility of metaphysics he was forced to maintain that the concept of causality has “in fact merely been borrowed from experience and from habit has taken on the appearance of necessity” (B20).

In the two other passages from the second edition that mention Hume by name Kant simply repeats these complaints in different forms.³⁶ At B5

³⁴ It also makes sense of Kant's famous remarks in the Preface of the *Prolegomena* at 4:257–261.

³⁵ Kant sometimes accords this mistake primacy: “The skeptical aberrations of this otherwise extremely acute man, however, arose primarily from a failing that he had in common with all dogmatists, namely, that he did not systematically survey all the kinds of *a priori* synthesis of the understanding” (A767/B795).

³⁶ It thus turns out that Kant discusses Hume's treatment of the principle of causality in greatest detail in the Doctrine of Method. There, Kant also remarks that Hume “falsely inferred from the contingency of our determination **in accordance with the law** the contingency of **the law** itself, and he confused going beyond the concept of a thing to possible experience (which takes place *a priori* and constitutes the objective reality of the concept) with the synthesis of the objects of actual experience, which is of course always empirical; thereby, however, he made a principle of affinity, which has its seat in

Kant remarks that “the very concept of a cause so obviously contains the concept of a necessity of connection with an effect and a strict universality of rule that it would be entirely lost if one sought, as Hume did, to derive it from a frequent association . . . and a habit . . . of connecting representations arising from that association.” At B127–129 he again complains about the way in which Hume (like Locke) asserts an entirely empirical origin for what are pure concepts of the understanding, since such an origin “cannot be reconciled with the reality of scientific cognition a priori that we possess” (B128). Thus, the passages that Kant adds in the second edition of the *Critique* do not represent a fundamental change of position or an especially significant addition on the basis of further reflection that he engaged in while writing the *Prolegomena* in 1783 or revising the *Critique* prior to 1787. Rather, Kant’s basic attitude toward Hume is already determined in the first edition of the *Critique* in his explicit remarks in the Discipline of Pure Reason.

Kant's Reply to Hume

Now that we understand Kant’s attitude toward Hume in the *Critique* in general, we can finally answer the historical question: What is Kant’s reply to Hume on the issue of causality? The first point to note is that one common answer to this question must be incorrect. Previous commentators have assumed that Kant’s reply to Hume is perfectly straightforward – at least in principle, even if certain crucial details have been a matter of dispute. For in the Second Analogy it seemed that Kant is simply affirming what Hume is denying, namely that a cause is necessarily connected with its effect (perhaps even according to causal laws, if a strong interpretation of the Second Analogy is defended). Given the general consensus concerning the extent of the agreement about the assumptions that Kant and Hume share, the extensive scholarly debate on this issue has centered on whether Kant’s argument in the Second Analogy can be understood in such a way that it is in fact cogent, that is, on whether an argument could be found in the text that neither appeals to premises that Hume would (or at least need) not accept nor requires the acceptance of subtle fallacious inferences in order to obtain a conclusion that contradicts Hume’s position. It is true that describing this debate at such a high level of generality hides important complications. For example, as we saw in Chapter 3,

the understanding and asserts a necessary connection, into a rule of association, which is found merely in the imitative imagination and which can present only contingent combinations, not objective ones at all” (A766/B794).

much has been written on whether Kant is attempting to establish the “every event–some cause” principle or whether he thinks he can establish the “same cause–same effect” principle.³⁷ However, both parties to this debate have agreed that Kant is attempting to *refute* Hume’s position on terms that he accepts; their disagreement simply focuses on which of Hume’s skeptical principles Kant is attempting to refute.

However, our investigation of (1) Kant’s pre-Critical philosophy, (2) the reception of Hume in Germany in the 1750s, 1760s, and 1770s, (3) Kant’s explicit remarks about Hume in the *Critique* (especially in the Doctrine of Method), and (4) Kant’s model of causality in the Second and Third Analogies establishes that Kant is not even attempting to refute Hume’s position. Consider first Kant’s pre-Critical reaction to Hume. To avoid the problem that Hume’s remarks about causality pose for the version of causal interaction that he had developed early in his pre-Critical period, Kant introduces the notion of a real ground, which can support not logical necessity, but rather something we might call natural necessity. By contrast, Hume, who holds that we can have knowledge only of relations of ideas and matters of fact, thinks that all necessity derives from relations of ideas and must be purely logical. But notice that Kant does not present real grounds and the necessity that depends on them in the course of a refutation of Hume, but rather as part of a response to a problem that Hume’s objections entail for his *own* position (which he is developing primarily in response to the German debate about pre-established harmony and physical influx). While the Critical Kant could have abandoned his initial pre-Critical reaction to Hume, the evidence that we have presented above suggests otherwise. For Kant quickly recognizes that real grounds require both a faculty other than pure reason and a principle other than that of identity or contradiction, and it would appear that the way he distinguishes between the logical and real uses of the intellect in the Inaugural Dissertation and his subsequent explanation of our knowledge in terms of the real use of the understanding in the *Critique* is supposed to provide the explanation real grounds require.³⁸

³⁷ See, for example, Lewis White Beck, *Essays on Kant and Hume* (New Haven: Yale University Press, 1978); Henry Allison, *Kant’s Transcendental Idealism; An Interpretation and Defense* (New Haven: Yale University Press, 1982); Paul Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987); Michael Friedman, “Causal Laws and the Foundations of Natural Science,” in *The Cambridge Companion to Kant*, ed. P. Guyer (New York: Cambridge University Press, 1992), pp. 161–199; and James Van Cleve, *Problems from Kant*, (New York: Oxford University Press, 1999).

³⁸ More specifically, the problem with real grounds was that they could not involve what logical grounds were based on, namely the principle of contradiction, since real grounds

Thus, Kant's immediate reaction to Hume in his pre-Critical period looks nothing like a refutation based on commonly accepted assumptions.

Second, what is striking about the reception of Hume in Germany from the 1750s through the 1770s is the fact that almost no one accepted what was novel about Hume's account of causality. However, what is especially relevant for our current purposes is the fact that no one felt that Hume's position stood in need of an explicit refutation (even if Sulzer did suggest reasons that might be developed into one). For the most part, it was thought to be completely obvious that Hume's analysis of causality omitted several crucial components, such as the notions of objective necessity and intelligibility. The primary area of disagreement concerned whether one could explain these non-Humean components in terms already available within a Leibnizian framework, as Sulzer proposed, or whether one would have to provide an explanation that went beyond what Leibniz had said, as Tetens and Crusius thought. On this point, Kant seems to agree with Tetens and Crusius, since his account of the understanding in the Critical period not only provides room for real grounds, as we just saw, but also involves categories, that is, discursive (and hence intelligible) representations of objects that contain necessary connections. Accordingly, the reception of Hume in Germany does not suggest that a refutation of Hume's position was viewed as a priority, and Kant's reflections on causality fit into such a pattern.

Third, our analysis of Kant's explicit remarks about Hume in the Doctrine of Method indicates that his primary interest does not lie in refuting Hume's position on causality. For one, Kant does not so much present detailed *arguments* against skepticism as provide an *orientation* that views skepticism as an unstable position. As we saw above, Kant suggests (1) that skepticism goes too far when it infers the complete ignorance of reason from reason's incapacity to defend whatever dogmatic claims happen to be made and (2) that skepticism is susceptible to self-referential problems (since nothing about skepticism per se can prevent it from being called into question itself). Yet the first point need not bother a more cautious

were defined in opposition to logical grounds. If real grounds were to be invoked only for causal relations, then it could seem ad hoc to posit a separate faculty or principle for it alone. However, by reflecting on the general problem of metaphysics (and the status of synthetic a priori propositions), Kant, unlike Hume, recognizes that the problem of real grounds is not specific to causality, but rather can be generalized. But since real grounds are commonplace in such contexts, it is perfectly legitimate to posit a separate faculty, with a distinct set of abilities and functions, to account for a broad range of cases. As I understand this, it is not designed to refute Hume's position.

skeptic (i.e., a skeptic who is careful about stating the exact scope of his conclusions), and with respect to the second, Kant himself never actually develops a detailed self-referential attack on skepticism. Instead, Kant's reaction to skepticism is to undertake a critique of pure reason, that is, an investigation of the nature and boundaries of our cognitive faculties. In this way, Kant hopes not only to account for synthetic a priori knowledge (e.g., mathematics and pure natural science) and to determine the possibility of metaphysics, but also to explain what skepticism gets right (e.g., when it denies that reason can have knowledge of objects that cannot be given to our senses). In this respect, Kant is not even disagreeing with Hume's position, much less refuting it.³⁹

For another, although Kant does mention the issue of causality when discussing Hume's position in the Doctrine of Method, it is certainly not his primary focus. Rather, just as is the case in several other passages (including his famous discussion of Hume in the *Prolegomena*), he refers to causality as one prominent instance of a more general issue that has important implications for his critique of pure reason. As a result, Kant's main interest in undertaking a critique of pure reason concerns establishing the importance of the understanding as a distinct faculty that is responsible for a priori representations of necessary connections, and therefore he addresses the issue of causality only indirectly by means of it. Thus, Kant's explicit remarks about Hume in the Doctrine of Method neither supply nor lead one to expect a refutation of Hume's account of causality.

Finally, the strongest argument for thinking that Kant is not attempting to refute Hume's position derives from the fact that Kant's model of causality in the Second and Third Analogies is radically different from Hume's. As we saw in Chapter 4, what Kant takes an effect to be, namely a continuous change from one determinate state to another, is not what Hume understands an effect to be, namely a determinate state of an object at a particular moment in time. Thus they do not agree on what the explanandum is. Nor is the cause for Kant – a phenomenal substance – identical to the cause for Hume – a determinate state at a particular moment in time (that is regularly followed by another determinate state). Accordingly, they do not agree on what the explanans is. As a result, Hume's denial and Kant's assertion of a necessary connection between cause and effect do not directly contradict each other, because they are

³⁹ Though it is true that Kant rejects the implication Hume draws, namely that there can be no synthetic a priori truths.

talking about the possibility or impossibility of a necessary connection between completely different kinds of entities. Moreover, the relationship between the cause and the effect is not the same for Hume and Kant, since – setting aside the contested modal status of causal relations – Hume asserts mere constant conjunction, whereas Kant invokes an indeterminate activity of a substance in accordance with its nature that determines the change of state in the effect. Since Hume and Kant do not agree on what the explanandum, the explanans, and the relationship between the cause and effect are supposed to be, the model of causality that Kant employs is simply not the same as Hume's, and since it is open to Hume to reject the very starting point of Kant's argument (on the grounds that he does not share his most fundamental assumptions), Kant's argument cannot refute Hume's position.

One natural strategy at this point would be to attempt to reinterpret Kant's model of causality on the basis of resources inherent in Hume's model in such a way that Kant's argument could be made intelligible on Humean terms. On this strategy, while Hume expresses skeptical doubts about a necessary connection *between* cause and effect, as he understands these terms, Kant could reinterpret Hume's view as raising doubts about such a connection *within* the effect, as Kant understands the term, since it is only within the effect that there is the kind of temporal asymmetry that was required in Hume's account. If Kant could show that the first state of the effect is necessarily connected with the second state of the effect, then he might be able to reply to Hume directly and on Hume's own terms. For in that case he will have established a necessary connection where Hume thought that none could be found.

However, there are two difficulties with this strategy. First, such a reinterpretation does not immediately put one in a position to reconstruct a cogent argument against Hume on Kant's behalf. For example, Kant can assert that there is a necessary connection within the effect, that is, between the initial state of one substance and its later state, since the later state *must* be different from the earlier state if one is to be able to infer that it is an event that stands in need of a cause. However, the kind of necessity involved in the change from an initial state to a later state is not causal, but rather conceptual, since the change in question would not be an event at all (and thus would not require a cause) unless it contained two different states. But it is far from clear how to move from this kind of conceptual necessity to the kind of causal necessity at issue in the Analogies.⁴⁰

⁴⁰ See Guyer's discussion of this kind of point in *Kant and the Claims of Knowledge*, p. 249.

Second, and more seriously, the reinterpretation strategy presupposes that one can translate the various elements of Kant's model of causality so that they are expressible in Humean terms. However, the stark contrast between Hume's and Kant's fundamental ontologies precludes any such reinterpretation. Hume's events are states of affairs at instantaneous moments in time, whereas Kant's events are continuous changes of state over time. Hume accepts only events (or matters of fact), while Kant accepts noumenal and phenomenal substances, causal powers, essential natures, as well as indeterminate relations, such as inherence and "the causality of the cause." Hume attempts to construct an account of the world solely on the basis of such discrete events, whereas Kant is concerned with explaining how certain concepts and principles are necessary conditions of knowledge of a single, spatio-temporal world. Hume's events are distinct from each other, while Kant attempts to establish grounding or dependency relationships between substances and their determinate states. Hume's events are neither active nor passive, Kant's determinations are passive and his causes active. Hume and Kant thus share no neutral philosophical vocabulary that would allow Kant to formulate a refutation of Hume on Hume's own terms. As a result, one must conclude that the translation strategy cannot work and that Kant's reply to Hume cannot take the form of a refutation.

If Kant is not attempting to refute Hume's position on causality, then how should his reply be understood positively? The various considerations just discussed reveal, I believe, that what Kant is attempting to accomplish is to develop a comprehensive philosophical account that represents a fundamentally new alternative to Hume's position (an alternative that he thinks has significant advantages when compared with Hume's at an appropriate level of generality). That is, Kant's strategy is not to use a set of explanatory terms and concepts he shares with Hume to show how Hume failed to see which implications they had (which would amount to a refutation of Hume's position), but rather to provide a different set of concepts and doctrines that are supposed to obviate the very framework that Hume's approach presupposes.

The same body of evidence that was used above to show that Kant's reply to Hume does not consist in a refutation supports the view that his basic intention is to provide an alternative to Hume's position. As we saw above, other German philosophers in the 1750s, 1760s, and 1770s often understood their own views as alternatives to Hume's position, whether they relied on Leibniz's and Wolff's established doctrines (Sulzer and Mendelssohn) or attempted to develop independent views (Crusius and Tetens). Kant's introduction of real grounds in the early 1760s can also

be seen as an attempt to make available a concept that is entirely alien to Hume's philosophy. Since Hume admits only relations of ideas (which can be understood as logical grounds) and matters of fact (which, Hume argues, cannot be understood in terms of logical grounds, and the beliefs concerning which can derive only from immediate experience or habit), there is no room for real grounds within Hume's view.

What we have seen in the *Critique* establishes even more clearly that Kant is developing an alternative to Hume. Kant does not adopt the basic elements of Hume's model of causality in order to show that there must be a necessary connection between cause and effect. Rather, he assumes that the category of causality contains the idea of a necessary connection and argues that the category of causality is necessary for a relation (successive determinations within an object) that makes experience of a single world possible, and in the course of his argument he employs a model of causality whose fundamental constituents are completely different from Hume's. Seen from this perspective, the differences between Hume's and Kant's models are inevitable given the different ontologies they accept and the different projects they undertake. In particular, Kant's use of the term "determination," which might appear to be simply another name for property or state of affair, is actually much richer and ends up driving his argument (e.g., in the guise of the problem of time-determination).

Kant's account of the various uses of reason in the *Discipline of Pure Reason* not only supports the idea that he is providing an alternative to Hume's position, but also introduces a perspective for understanding how the other elements of his view fit together. As we saw above, Kant rejects both Wolff's dogmatism and Hume's skepticism and proposes that a stable philosophical position can be attained only by undertaking a critique of the faculty of pure reason itself. A critique of pure reason, however, leads him to work out a novel account both of our cognitive faculties and of the ontology that is consistent with such faculties. Thus, Kant's methodological reflections in the *Discipline of Pure Reason* provide a general framework within which real grounds, the real use of the understanding (in the form of the categories of causality and mutual interaction), a new model of causality, and the ontology on which this model draws can be understood. In short, it offers a comprehensive alternative to the skeptical system advocated by Hume.

However, Kant's actual intentions in the *Discipline of Pure Reason* may not be limited to presenting an alternative to Hume's position. For Kant clearly asserts that his position has significant advantages over Hume's, and he cites both specific and general points to back up this assertion. In

addition to describing the proper status of our mathematical knowledge (as synthetic a priori), on his account our concept of causality really does contain necessity (just as common usage would have it) and a valid derivation of that necessity (i.e., one that does not attempt to derive necessity from empirical cases) is possible. At a more general level, Kant holds that the instability of skepticism is unattractive compared with what his critique of pure reason can deliver, since the latter can determine the boundaries of our knowledge in a principled and hence stable way.

Yet if it is accurate to interpret Kant's position as a radical alternative to Hume's position, one might think not only that a refutation of Hume's position is impossible – since Kant's argument inevitably draws on resources foreign to Hume's basic position – but also that neither one is ultimately in a position to claim advantages over the other. That is, if Hume and Kant are truly supposed to be competing against each other, one would have to be able to provide criteria that are neutral between them. However, if the differences between their basic views are as great as it now appears, then, besides internal consistency, which is a criterion that both views must satisfy, it is far from clear what other criteria are available. One might think, for example, that Hume's view ought to be preferred because invoking only matters of fact and relations of ideas accords better with Ockham's razor (since Kant views the world as consisting of noumenal and phenomenal substances, their causal powers and activities, and the determinate states for which they are responsible). Yet one can agree that entities should not be multiplied needlessly without agreeing about which entities are in fact needed. Less obviously, but more importantly, this kind of argument depends on equating Hume's matters of fact with Kant's determinations so that substances, causal powers, and activities can be viewed as additional, extraneous entities. The argument developed above suggests, however, that such an equation cannot be made and that therefore Ockham's razor turns out to be too dull to cut against Kant's position on this point. If no other criteria that are truly neutral between Hume's and Kant's positions can be identified, they may still be competitors, but only in a rather unusual sense, since it may be unclear how to judge who wins the competition.

We now finally have an answer to the historical question of how Kant is replying to Hume on the issue of causality. Kant is not attempting to refute Hume by showing how Hume makes an assumption that implicitly commits him to maintaining a necessary connection between cause and effect. For any such refutation would presuppose that Kant and Hume share the same basic philosophical framework. However, as we have seen,

Hume and Kant do not agree about what model of causality is presupposed in considering whether there is a necessary connection between a cause and an effect, since they proceed from very different assumptions about what events are, about what kinds of entities causes and effects are, and about what kind of relationship could hold between them. Instead of attempting to discover what sort of position results from the skeptical use of pure reason, we have seen that Kant believes that he must undertake a critique of the faculty of pure reason itself, which leads him to develop an entirely novel philosophical system that competes against Hume's skeptical position. As a result, what is important in understanding Kant's views on causality properly lies not in how they might directly relate to Hume's views, but rather in how they can be understood on their own terms and within the context of Kant's larger philosophical project.

CONTEMPORARY CONSIDERATIONS

With an answer in hand to the historical question of how one should understand the nature of Kant's reply to Hume, we can now address, briefly, the systematic question of the importance of Kant's views on causality in a contemporary setting. However, historical and systematic issues are, in this case, closely related, since the answer to the historical question developed above can provide guidance for understanding how Kant's views can be relevant today. For if Kant does not even attempt to refute Hume's empiricist account of causality, then there is no reason to expect that he would have any special critical insights into the weaknesses of empiricist accounts, and one can dispense with potentially fruitless attempts to locate some inconsistency that allegedly lies deeply hidden in Humean commitments.

Instead, one can immediately use Kant's views on causality either to articulate the fundamental contrast between accounts that are roughly Humean in character and those that can be viewed as similar in essential respects to his own or to clarify certain issues that have arisen concerning the latter kind of account. In the following, we focus mainly on current debate about three issues: (1) the metaphysics of causality, (2) the status or nature of the laws of nature, and (3) the conflict between freedom and determinism. In the case of the metaphysics of causality, we see below how Kant's views on causality can bring out the contrast between empiricist (event-based) and nonempiricist (causal powers-based) accounts. With respect to the issues of the status of the laws of nature and of free will and determinism, Kant's position will be of help primarily in allowing us to see

how one might respond to objections that have been raised against various leading accounts of these issues. However, by understanding how Kant's views on causality are helpful with respect to several specific issues under discussion today, one can also see how Kant's views on causality might also put one in a position to develop a more comprehensive systematic metaphysics.

The Metaphysics of Causality

Consider first metaphysical accounts of the nature of causality. Over the last several decades, a tremendous amount of effort has been devoted to understanding the metaphysics of causality. The question that has attracted significantly more attention than any other has been about the nature of causal relations: Can causal relations be understood in terms of other relations or are they irreducible and primitive? Contemporary analytic philosophers, in particular, have often attempted either to provide an analysis of causal relations in other terms – such as regularity relations, relations of counterfactual dependence, statistical correlations, or relations of nomological subsumption – or to refute these attempts, typically by way of counterexamples (whether they be the existence of accidental regularities, common cause scenarios, or cases of preemption).⁴¹ However, their preoccupation with this question should not obscure the fact that they have also addressed, albeit to a somewhat lesser degree, questions concerning the nature of causal relata. Specifically, how many causal relata are there (two, three, or four or more)? Must causal relata be immanent in the world or could they transcend it? How should the causal relata be individuated (e.g., in terms of coarse-grained or fine-grained events)?⁴²

⁴¹ The literature is vast, but a few central contributions are: John Mackie, *The Cement of the Universe* (New York: Oxford University Press, 1973); David Lewis, "Causation," in *Philosophical Papers 2* (New York: Oxford University Press, 1986), pp. 159–213, and "Causation as Influence," *Journal of Philosophy* 97 (2002): 182–197; Jaegwon Kim, "Causation, Nomic Subsumption, and the Concept of an Event," *Journal of Philosophy* 70 (1973): 217–236; Patrick Suppes, *A Probabilistic Theory of Causality* (Amsterdam: North Holland, 1970); Donald Davidson, "Mental Events," in *Essays on Actions and Events* (Oxford: Clarendon, 1980), pp. 207–227; Ellery Eells, *Probabilistic Causality* (New York: Cambridge University Press, 1991); and Judea Pearl, *Causality: Models, Reasoning, and Inference* (New York: Cambridge University Press, 2000).

⁴² See, for example, Donald Davidson, "Actions, Reasons, and Causes," in *Essays on Actions and Events*, pp. 3–19, and "Causal Relations," also in *ibid.*, pp. 149–162; David Armstrong, *A World of States of Affairs* (New York: Cambridge University Press, 1997); Jonathan Bennett, *Events and Their Names* (Indianapolis: Hackett, 1988); L. A. Paul, "Aspect

As we have seen above, Kant and Hume develop contrasting answers to many of these questions. For example, on the issue of the nature of causal relations, Hume argues that causal relations are reducible to constant conjunctions or brute regularities (plus a subjective expectation), since no logically necessary connection obtains between any two causally related events, while Kant thinks that causal relations are represented by a category, that is, a primitive, nonempirical concept of the understanding, and are therefore irreducible to any other, more primitive objective relation.⁴³ However, by investigating Kant's model of causality in detail and considering how best to understand his position as a reply to Hume, it became apparent that the difference in their answers about the nature of causal relations is simply part of a more fundamental contrast in their basic philosophical frameworks (which includes both their ontologies and the projects in which they are invoked). For Hume's empiricist project starts with events, where events are understood as distinct states of affairs at instantaneous moments, and then attempts to construct the causally connected world out of them. Thus, for Hume, causal relations between events will seem to stand in need of explanation, since his starting point implies that only events (along with spatio-temporal and purely logical relations between them) can be accepted as basic ontological constituents. Kant, by contrast, does not presuppose events as his starting point. He holds that there are permanent substances along with changes of state in these substances and argues that such changes of state are possible only if substances exercise their causal powers in accordance with their natures. By thus accepting substances endowed with causal powers, Kant can grant that they stand in real relations to each other, relations that do not require (or even permit) further analysis, since it is built into the very idea of substances exercising their causal powers that they can determine the state of another substance in this way.

However, Kant's and Hume's views on causality do not merely illustrate the general point that the ontology one accepts influences how one can understand both what causal relata are and which kind of causal relations they can stand in, but also allow one to raise the question of precisely which feature of their ontologies requires these contrasts between their accounts of causal relations. That is, Hume's events differ in several

Causation," *Journal of Philosophy* 97 (2000): 223–234; Christopher Hitchcock, "The Role of Contrast in Causal and Explanatory Claims," *Synthese* 107 (1996): 395–419; as well as several contributions mentioned in the previous note.

⁴³ For present purposes, we can ignore the differences between the categories of causality and mutual interaction.

respects from Kant's substances exercising their causal powers according to their natures. Events exist only at an instant, whereas causal powers endure over time in independently existing things. Events are contingent occurrences, whereas substances act in accordance with the necessity of their natures. Events are neither active nor passive, whereas the exercise of a causal power is an activity that contrasts with the passivity of the effect (i.e., of the change of state that another substance suffers). Which of these respects is most fundamental to the differences that emerge between Hume's and Kant's views about how to understand causal relations (specifically, to causal relations being contingent or necessary and reducible or irreducible)?

Our interpretation of Kant's model of causality allows us to see that several of these contrasts are not in fact fundamental to the differences between their views on causality. While it is true that Hume defines events as instantaneous, it is not clear that events could not be allowed to endure (especially if one granted the possibility of complex events that had simpler events as their constituents), and if events can endure, then the fact that causal powers endure does not represent a fundamental difference between Kant's and Hume's accounts.⁴⁴ Also, as Hume himself notes, if substances are defined in terms of independent existence, one could naturally view each and every event as a substance.⁴⁵ Accordingly, simply understanding substances as independently existing entities does not constitute an essential difference between their respective accounts of causality, either. Similarly, whether or not to ascribe modal features to entities would not seem to be the decisive feature standing in the way of a reductive story. For there is no obvious contradiction in asserting that two events are related by some kind of necessity relation (even if not by logical necessity), and events themselves can have essential properties (since, e.g., an event might be thought to be necessarily contingent).⁴⁶

⁴⁴ For a clear illustration of an account of events according to which they endure, see L. A. Paul's "Logical Parts," *Nous* 36 (2002): 578–596.

⁴⁵ See, for example, book I, part IV, section V, of Hume's *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised by P. H. Nidditch (Oxford: Clarendon, 1978), p. 233.

⁴⁶ See David Lewis's "Events," in *Philosophical Papers* 2, pp. 247–254. However, David Armstrong, for example, holds that events can stand in necessitation relations in virtue of a necessitation relation between the universals that they instantiate. See his *What Is a Law of Nature?* (New York: Cambridge University Press, 1983), and the discussion of his conception of laws of nature below. But note that Armstrong admits that there is a further question here that he does not address in this work, when he remarks: "I regard Actualism as the most difficult and uncertain of my three assumptions. It is bound up with the difficult question whether the laws of nature involve logical necessities in things. . . . For

Accordingly, nothing prohibits the attribution of modal features to events. Thus, several of the contrasting features of Kant's and Hume's models of causality are not essential to the fundamental differences between their views.⁴⁷

Rather, what our reflection on Kant's and Hume's accounts of causality suggests is that the crucial difference between their views (i.e., the difference that precludes the possibility of reducing causal powers to events) lies in the idea that a substance can be *active*, that is, that an effect can be produced only through the *exercise* of a causal power.⁴⁸ That is, Hume's events are not active, since we have, he thinks, no impression of any event being productive of anything else (or of any event being constrained by anything else). Nor, for that matter, are they passive, since passivity is possible only if something active is presupposed with respect to which something else is passive. Instead, events are best characterized as inert, namely as lacking entirely the ability to act or be acted on, since they are simply what happens to occur in a certain spatio-temporal region. This position contrasts starkly with Kant's view that the activity of a substance is the temporally indeterminate exercise of its causal power in accordance with its nature and circumstances so as to bring about a determinate effect.

The idea that the notion of activity distinguishes Kant's model of causality most fundamentally from Hume's can be confirmed by noting that understanding activity in this way allows one to see how several of the other features that Kant attributes to causal powers follow naturally. Imagine that you are convinced that no reductive explanation of the causal relation will succeed. At this point, you could still simply posit a real causal relation between events. You could even attribute some kind of necessity to such a relation (though such an attribution should be viewed as

dispositions and powers, if they are conceived of as the non-Actualist conceives them, involve logical or quasi-logical connections in the world between the dispositions and powers, on the one hand, and their actualizations, on the other" (p. 9).

⁴⁷ It is true that if one were to make these concessions (e.g., that events have such modal features), then the resultant position would not necessarily be very close to Hume's own position. However, such positions are still relevant insofar as they help us to identify which feature of the position that Hume does hold represents the most basic point of contrast to Kant's.

⁴⁸ This is a controversial claim, even among those who adhere to some version of a theory of causal powers. While Harré and Madden can be read as sympathetic to such a claim (even if they do not explicitly focus on it), Nancy Cartwright, *Nature's Capacities and Their Measurement* (Oxford: Clarendon, 1989), who prefers to talk of capacities, does not identify capacities as having activities, much less as making activities essential to capacities.

optional). However, without invoking the notion of activity, no explanation of how such a relation comes into existence is readily available. It is posited as a brute fact, with no foundation in the events that it relates. Yet one can notice that a significant difference arises as soon as one accepts causal powers and views activity as the core difference between events and causal powers, since this makes possible an explanation of how a real causal relation comes into existence. For that is precisely what the activity of the causal power is in a position to establish, namely a real relation between (the states of) substances that are ontologically distinct.

Moreover, if one holds that causality does involve some kind of natural necessity, viewing activity as the distinctive element of causal powers also provides insight into how natural necessity can emerge. For the mere existence of a nature consisting in necessary properties does not obviously or immediately entail the necessity of the relations in which whatever has that nature stands. However, if the substance having that nature must act in accordance with it, then one can understand why the causal relation might be necessary in a corresponding sense. For the substance's nature provides the rule that determines what kind of activity it is and what it is able to bring about.

Let us return now to the contemporary debate to see how such a focus on Kant's model of causality and on the notion of activity it employs can illuminate this debate. As we saw above, most contemporary discussions of causality have assumed, along with Hume, that causes and effects are simply events so as to focus on the nature of the causal relation that holds between them. However, some attention has also been devoted to the idea that causes should be understood not in terms of events, but rather in terms of causal powers.⁴⁹ The view that causality should be explained in this way goes back at least to Aristotle and was dominant for centuries thereafter. However, it came under attack in the early modern period on the grounds that, given its connection to the scholastic doctrine of real qualities and substantial forms, it posited an occult quality that could have no explanatory power. A variant of this view was then reintroduced into the contemporary debate by Rom Harré and E. H. Madden, who argued that what should be viewed as ontologically primitive is not events, but rather what they call a "powerful particular." Their

⁴⁹ See, for example, Rom Harré and E. H. Madden, *Causal Powers: A Theory of Natural Necessity* (Oxford: Basil Blackwell, 1975); Fred Freddoso, "The Necessity of Nature," *Midwest Studies in Philosophy* 11 (1986): 215–242; and Cartwright, *Nature's Capacities and Their Measurement*.

explanation of what a powerful particular is involves causal powers, natures, and natural necessity in a way that is intended to be diametrically opposed to Humean accounts and sympathetic to broadly Aristotelian accounts. Nancy Cartwright has likewise argued for the indispensability of what she calls "capacities" or "enduring tendencies," which she explicitly characterizes as being "anti-Humean."⁵⁰

However, if causal powers, so understood, are to amount to a genuine alternative to events, one must have a clear account of exactly what causal powers are. In particular, one must know which specific feature makes causal powers different from events. Otherwise, Humeans can simply reduce causal powers to (sets of) events and deny that causal powers represent any genuine alternative to their own views.⁵¹ Harré and Madden are aware of this issue and explicitly claim that "the conceptual analysis of causal notions must clearly be drawn from a wider ontology than events."⁵² But what is it about a powerful particular that is supposed to generate a genuine and irreducible difference with events?

Harré and Madden provide an analysis of events according to which they are "atomistic" and "absolutely independent" of each other. However, without providing an explicit and detailed characterization of what atomism amounts to, it is not clear that powerful particulars are not atomistic in some sense, since each one presumably cannot be divided into more ultimate constituents.⁵³ Similarly, each powerful particular is at least logically independent of every other in the sense that the existence of one does not logically entail the existence of another (even if there may be some nonlogical, natural necessitation relations). Thus, Harré and Madden's analysis of events does not clearly identify any contrast that could represent the fundamental difference between causal powers and events.

⁵⁰ Cartwright, *Nature's Capacities and Their Measurement*, p. 4. Cartwright prefers not to use the term power, because she wants to focus not on individuals, but rather on the abstract relation between capacities and properties. See p. 9 and chap. 4, and also *The Dappled World* (New York: Cambridge University Press, 1999), esp. chap. 3.

⁵¹ See, for example, Sydney Shoemaker, "Causality and Properties," in *Identity, Cause and Mind*, ed. S. Shoemaker (New York: Cambridge University Press, 1984). In other words, one could introduce a non-Humean conception of events, such that the exercise of a causal power (or, for that matter, even a noumenal state of affairs) could be considered an event. However, for current purposes, such a conception would simply obscure important differences between Humean and non-Humean accounts.

⁵² Harré and Madden, *Causal Powers*, p. 5.

⁵³ Harré and Madden seem to derive atomism about events from Hume's atomistic psychology, but it is clear that such a connection need not form an essential feature of a Humean account of causality.

Nor do the most prominent features of Harré and Madden's positive analysis of "powerful particulars" fully succeed at distinguishing causal powers from events. Though there is a point to contrasting Humean regularities with singular occurrences, it is clear that what is central about a powerful particular is not its singularity or particularity, but rather the idea that it has a power. Harré and Madden explain what it means to ascribe a power to a thing as follows: "'X has the power to A' means 'X (will)/(can) do A, in the appropriate conditions, *in virtue of its intrinsic character.*'"⁵⁴ Now their emphasis in the final phrase of this explanation seems to suggest that the distinguishing feature of a power lies in the intrinsic character of the thing that has the power. However, as we saw above, it is at least logically possible that Humean events could be characterized essentially by the intrinsic qualities that they exemplify in a certain spatio-temporal region. Yet if this is possible, then one cannot draw the ultimate contrast between causal powers and event-based views in terms of intrinsic features alone.

Fortunately, our understanding of Kant's model of causality and how it represents an alternative to Hume's account can help to suggest what the underlying difference between causal powers and events-based accounts might be. For, as we saw above, what is central to Kant's model of causality and what makes it essentially different from Hume's account is the role that activity plays in it. Specifically, what distinguishes the exercise of a causal power from an event is that an event simply occurs somewhere at some time, whereas the exercise of a causal power is a temporally indeterminate activity that is constrained by a rule provided (1) by the nature of the thing whose causal power is being exercised, (2) by the relations in which the thing stands to other things, and (3) by their natures. As we saw in Chapters 4 and 5, it is particularly important to Kant that such an activity is not temporally determinate. However, for current purposes one can abstract from the specifically temporal features of activity and still hold that an activity is essentially different from an event. If a cause determines its effect, then there is, so Kant thinks, an activity that is essentially constrained by (though, more positively, it also receives guidance from) a rule by means of which it brings about a passive determination of an object. Causal powers can thus be understood as essentially distinct from event-based models of causality by means of this notion of activity. To put the case the other way around, without some such notion of activity, it becomes increasingly difficult to see how one could hold onto any

⁵⁴ Harré and Madden, *Causal Powers*, p. 86.

essential difference between causal powers and events-based models of causality, and one would be forced to declare that what appeared to be a significant debate is in fact a merely verbal dispute based on confusion or the improper use of terms.

Moreover, viewed in this way, there is reason to believe that Harré and Madden (and also Cartwright) could see Kant's position as suggesting either a clarification of or a friendly amendment to their view.⁵⁵ For in their account of powerful particulars Harré and Madden invoke the idea of "generative mechanisms, whose structure and components constitute the essential nature of the permanent things and materials in the world."⁵⁶ Further, these generative mechanisms play a crucial role in explaining what is distinctive about their causal powers account: "There is an ontological tie that binds sequential events together, but it is not event-like. It is the persisting generative mechanism consisting of powerful particulars and natural agents which produces the sequence of events and states and endures throughout."⁵⁷ If one focuses on the idea that causal powers are powers *to do* something and takes seriously the idea of a *generative* mechanism, one can see that Harré and Madden themselves take implicit recourse to the notion of activity, even if they do not explicitly cite it as a separate feature of their view and do not emphasize how it could play a central role in their account. In this way, taking recourse to Kant's views on causality can contribute to the contemporary debate about causation by helping us to understand more precisely what the foundation of causal powers-based models of causality might be such that they represent a distinct alternative to Humean event-based models.

The contrast between Kant's and Hume's views on causality can also contribute to the contemporary debate by shedding light on what is essential about event-based models. One common line of thought about events that has clear roots in Hume proceeds as follows. First, accept events as ontological primitives, and then construct causality, laws of nature, chance, and other nomic concepts on the basis of such primitive

⁵⁵ Nancy Cartwright seems, as much as possible, to steer clear of the metaphysics of causality, for example, by being noncommittal on whether causality requires necessity of any sort, and she could consistently do the same with respect to activities. At the same time, just as her agnosticism about modality leaves room for necessity, she could certainly be open to activities and could even embrace them *if* she thought that doing so was the only way for her account to be distinct from the Humean views she finds problematic.

⁵⁶ Harré and Madden, *Causal Powers*, p. 130.

⁵⁷ *Ibid.*, p. 131.

events. David Lewis, for example, is one prominent and very explicit proponent of such a project. In the introduction to the second volume of his *Philosophical Papers*, Lewis describes this kind of project in terms of a doctrine he calls “Humean supervenience”:

It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. . . . We have geometry: a system of external relations of spatio-temporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else [e.g., causality, laws of nature, counterfactuals, persistence through time, etc.] supervenes on that.⁵⁸

Such a doctrine might appear to be extremely attractive. After all, if Hume is read as emphasizing that we can have no nontrivial knowledge without empirical evidence, and all the empirical evidence we happen to have is of particular matters of fact or events, then the world that we know consists in nothing other than such events (and spatio-temporal relations between them). Further, the view might also seem to be especially straightforward and clear, since it ultimately invokes only events (and certain very clear relations between them), and events, it might be thought, are uncontroversial, because the contested issue might seem to be about whether one can (or must) go further and accept something *in addition* to events.

However, the contrast between Kant’s and Hume’s accounts of causality described above provides a helpful perspective on such a project by allowing us to focus more clearly on what events are and to see that they may not ultimately have all the advantages that they might at first appear to have. For reflection on Kant’s model of causality (again, abstracting from the issue of temporal indeterminacy) can lead one to ask a Humean such as Lewis why causal relations are not themselves events. If an event “is a localised matter of contingent fact” that occurs in a certain place and time, then why is the causing of an effect at a certain place and time not an event just as the redness of a certain surface (e.g., of an apple) is?⁵⁹ This question runs counter to the very setup of Lewis’s project, since he starts off by assuming that causality is to be understood as a causal *relation*

⁵⁸ David Lewis, “Introduction,” in *Philosophical Papers* 2, pp. ix – x.

⁵⁹ David Lewis, “Events,” in *Philosophical Papers* 2, p. 243. One might object to Lewis’s stipulation that events must be contingent. However, we can abstract from this issue for the present discussion.

between events (rather than itself a primitive event) and then attempts to provide an analysis of causal relations in terms of other relations that are supposed to be clearer. But what could justify this initial assumption?

Lewis's most plausible reply to this line of inquiry is to draw attention to certain details of his theory of events, since it is this theory that must contain the reason for excluding the possibility that the causing of an effect at a certain place and time could be an event.⁶⁰ One detail worth considering is his claim that events must be "perfectly natural intrinsic properties," since it might be thought that the causing of an effect at a certain place and time is not perfectly natural and thus not an event. But drawing attention to this detail simply raises the question of what a perfectly natural intrinsic property is, and unless Lewis provides a clear criterion for distinguishing between perfectly natural intrinsic properties and unnatural intrinsic properties (or perhaps imperfectly natural intrinsic properties) and motivates that criterion, the advantages that Lewis's account might have seemed to enjoy could easily disappear.⁶¹ For without providing and justifying a criterion for determining which intrinsic qualities are natural, what qualifies as an event is not clear, and denying that a cause bringing about an effect at a certain place and time is an event is ad hoc.

At this stage in the argument, Lewis might point to a second detail of his theory of events. As we saw above, he seems to suggest that events are local qualities that require nothing more than a point to be instantiated, and he might think that the locality of events excludes the possibility that a cause bringing about an effect at a certain place and time could be an event. For there is one moment in space-time at which the quality associated with the cause exists and then a later (or at least different) moment in space-time at which the quality that is the effect exists. That is, the cause and the effect exist at different spatio-temporal points. But

⁶⁰ For an alternative account of events, see Jaegwon Kim, "Causation, Nomic Subsumption, and the Concept of an Event," *Journal of Philosophy* 70 (1973): 217–236, esp. p. 222, as well as his "Causes and Counterfactuals," *Journal of Philosophy* 70 (1973): 570–572, though the debate between Lewis and Kim about events focuses on a different issue. However, it is significant, I think, that after describing several difficulties faced by the two accounts of Humean causation that he finds most attractive, Kim notes that it "seems likely that clues to a correct account of these cases will be found not at the level of analysis in this paper, but at a deeper metaphysical level, involving such concepts as substance, power, and accident, or at a pragmatic level" ("Causation, Nomic Subsumption, and the Concept of an Event," p. 235).

⁶¹ Lewis attempts to develop such an account in "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 71 (1983): 343–377.

if an event can occupy only a single point, and the cause and effect exist at different points, then it follows that the cause and the effect cannot be (included in) the same event.

However, just as the previous response led us to inquire into what intrinsic qualities are natural, this response raises the question of why the locality condition should be accepted for the individuation of events. On this point, Lewis explicitly acknowledges that the locality condition is merely contingent, since “there might be emergent natural properties of more-than-point-sized things.”⁶² Whether or not there are such properties is an issue that science, perhaps physics, at a very advanced stage ought to determine. But such a qualification has significant costs, since it amounts to admitting that the Humean account does not obviously have empirical evidence in its favor and is also not any clearer philosophically than alternative accounts. That is, the advantages that Humeans sometimes claim on behalf of their account have disappeared. Lewis seems to recognize as much when he admits: “Really, what I uphold is not so much the truth of Humean supervenience as the *tenability* of it. If physics itself were to teach me that it is false, I wouldn’t grieve.”⁶³

In sum, reflection on Kant’s model of causality and the way in which it contrasts with Hume’s can contribute to the contemporary debate about the metaphysics of causality in two ways. First, one can draw on Kant’s notion of activity in order to explain what makes a causal power different from and irreducible to an event, an explanation that is indispensable for debate about the nature of causal relations to make sense in the first place. Second, reflection on Kant’s model of causality allows one to see that Humean theories of events may not have the advantages sometimes claimed for them. At the same time, since we have also learned that Kant’s account of causality is not meant to be a refutation of Hume’s, it is plausible to think that neither of these two points should be taken to amount to a criticism of the tenability of Humean accounts of causation (even if the second point is a criticism of claims that empiricists sometimes make about the relative advantages of their views).

Laws of Nature

Another issue that has received considerable attention in recent literature concerns the status of the laws of nature. One popular position, which

⁶² Lewis, “Introduction,” *Philosophical Papers* 2, p. x.

⁶³ *Ibid.*, xi.

finds inspiration in Hume and has been articulated more recently in a powerful way by David Lewis, holds that laws of nature are at base universal statements of regularities (such as "All Fs are Gs"), though this view is often strengthened so as to incorporate simplicity, informativeness, and other pragmatic notions into the choice of which regularities are to count as laws of nature.⁶⁴ Another position, articulated in a contemporary setting by David Armstrong, Michael Tooley, and Fred Dretske, holds that laws of nature express not regularities between particular instances, but rather necessitation relations between universals.⁶⁵ As Armstrong puts it: "Something's being F necessitates that same something's being G, in virtue of the universals F and G," though, to be more precise, what he means is that something's being F necessitates its being G in virtue of the necessitation relation between the universals F and G.⁶⁶ Thus, on the Humean view it is a law of nature that all ravens are black because all ravens (or perhaps all ravens that have been observed to date) are (or have been) black, whereas on Armstrong's necessitarian view this statement is a law because the universal "raven" necessitates the universal "blackness."

While these contrasting accounts of the status of the laws of nature share (or can share) many of the same features, one point on which they disagree concerns whether the laws of nature can be said to "govern" their instances, as our pretheoretical intuitions about laws of natures would suggest. According to Lewis's Humean view, laws cannot govern their instances since the laws, as generalizations of instances observed in the world, must *follow* from these instances and therefore cannot in turn constrain them. However, because Armstrong defines laws of nature not in terms of particular instances, but rather as expressing a relationship between universals, he can maintain that the laws of nature are prior to their instances. Thus, unlike Lewis, he can say that the laws govern

⁶⁴ See David Lewis, "Humean Supervenience Debugged," *Mind* 103 (1994): 473–489. Lewis defines a law of nature as follows: "Take all deductive systems whose theorems are true. Some are simpler and better systematized than others. Some are stronger, more informative than others. These virtues compete: An uninformative system can be very simple, an unsystematized compendium of miscellaneous information can be very informative. The best system is the one that strikes as good a balance as truth will allow between simplicity and strength. How good a balance that is will depend on how kind nature is. A regularity is a law iff it is a theorem of the best system" (p. 478).

⁶⁵ David Armstrong, *What Is a Law of Nature?*; Michael Tooley, "The Nature of Laws," *Canadian Journal of Philosophy* 7 (1977): 667–698; and Fred Dretske, "Laws of Nature," *Philosophy of Science* 44 (1977): 248–268.

⁶⁶ Armstrong, *What Is a Law of Nature?*, p. 96.

those instances that fall under the universals whose necessary relation is expressed by the law. That is, according to Lewis, the law that ravens are black does not independently constrain the color of any particular raven, since the law would not hold in the first place if all relevant ravens were not black. By contrast, Armstrong can say that the law that ravens are black necessitates that this particular raven be black, since it would violate the necessitation relation between the universals “raven” and “black” if an instance of the universal raven were not also an instance of the universal black.

This difference concerning whether laws are defined in terms of particular instances that constitute regularities or in terms of necessity relations between universals also provides a point of contrast for their respective accounts of how laws of nature are supposed to be explanatory. According to the Humean account, a law of nature explains its instances because it entails them deductively given that, on this view, a law of nature is defined in terms of its instances. There is no gap, so to speak, between the laws and the instances of them. Armstrong, by contrast, rejects this account of explanation on the grounds that it seems vacuous or circular to him to use a law to explain its instances if the law is *constituted* by these instances in the first place; nothing has been explained if a certain set of instances is being invoked to explain that same set of instances.

Whatever the merits of Armstrong’s objection to the Humean view of explanation, his account of the laws of nature opens up the possibility of a different account of explanation. By holding that the laws of nature express relations of necessity between universals, Armstrong can say that laws of nature explain their instances because these instances must take on the same necessitation relation that their corresponding universals have to each other. Even if (or, from Armstrong’s perspective, precisely because) laws of nature are not ultimately identical to the totality of their instances, they can (still) explain their instances. For just as the relations between the universals that make $2 + 2 = 4$ can explain why adding two apples to two oranges yields four pieces of fruit, so the relations between being a raven and being black can explain why this particular raven is black. In light of the fact that Armstrong posits necessitation relations between universals, his accounts of explanation can be quite different from Humean accounts, which are restricted to regularities and pragmatic criteria.

To see how Kant’s views on causality can be relevant to the debate on the status of the laws of nature between Humean regularity theorists and Armstrongian necessitarians, consider two objections that have been

raised against Armstrong's view.⁶⁷ First, Armstrong has been pressed on the exact nature of the necessity that holds between the universals referred to in the statement of a law of nature. Despite obvious parallels, the necessity in this case does not appear to be logical, since it would not violate the principle of contradiction if there were a raven that happened not to be black. In *What Is a Law of Nature?*, Armstrong states simply that one universal "necessitates" or "brings along" another without clarifying how these terms should be understood more precisely. In "The Identification Problem and the Inference Problem," however, where Armstrong responds to two of van Fraassen's criticisms, he specifies that the necessitation relation is (at least in paradigm cases) causal: "The required relation is the causation relation, the very same relation that is actually experienced in the experience of singular causal relations, now hypothesized to relate types not tokens."⁶⁸

However, one might find this response implausible. For even if one rejects a Platonist theory of universals as Armstrong does, it would still seem to be the case that a universal, understood as a *type* of entity (rather than a particular entity of a certain type), cannot cause anything, much less another universal or type of thing. What could it mean for one universal to cause another universal if not simply that instances of the one are the cause of instances of the other (a view that Armstrong has to reject, since affirming it would nullify the difference between Lewis's and his accounts)? Unless Armstrong can provide a more satisfying account of the necessity relationship that holds between universals, one can understand why the initial charge of obscurity can seem to retain its force.

Second, objections have also been raised to Armstrong's account of how it is that the laws of nature relate to what happens in the world. In one formulation, the objection is posed in terms of how the laws of nature are supposed to govern what happens in the world. If the laws of nature are supposed to govern what happens in the world, how can they do so unless the laws of nature are causally efficacious? We typically say, for example, that a ruler governs his country because he makes certain things happen in certain ways, as when he enforces the laws enacted in the proper way by the legislature. If the laws of nature or the universals they are based on are not entities that can act causally (so as to enforce

⁶⁷ These criticisms have been raised in different forms by, among others, Bas van Fraassen, *Laws and Symmetry* (New York: Oxford University Press, 1989), and Barry Loewer "Humean Supervenience," *Philosophical Topics* 24 (1996), 101–127, esp. pp. 118–119.

⁶⁸ *Philosophy and Phenomenological Research* 53 (1993): 421–422.

themselves), then they cannot literally govern as a ruler does and the suspicion arises that use of the term “govern” is purely metaphorical. Since, as we saw above, the idea that laws of nature govern what happens in the world parallels the idea that laws of nature are supposed to explain what happens, a slightly different version of this objection can be formulated for the nature of explanation. How is it that a necessitation relation between universals can explain what happens in the world? If laws of nature simply express a relation between universals, how can such a relation explain why specific things in the world are the way they are? From Hume’s perspective, one can express the force of this point by saying that if a law of nature were simply a relation between universals, then a law of nature would be nothing more than a relation among ideas, which is irrelevant to the contingent matters of fact that we find in the world.

Armstrong would presumably respond in different ways to the different formulations of this objection. To the governing formulation, Armstrong can concede that “governing” does not mean “causal,” but still maintain that the term is not entirely metaphorical, since it is expressing the idea that what happens in the world depends on the relation between the relevant universals, though the precise nature of this dependency may still not be completely transparent. This point becomes clearer when we consider his explicit response to the second formulation, which Armstrong states as follows: “If a certain type of state of affairs has certain causal effects, how can it not be that the tokens of this type cause tokens of that type of effect? The inference is analytic or conceptual.”⁶⁹ Armstrong’s basic idea is that whatever relations hold between types of things, or universals, must also hold between the things themselves, that is, their tokens or instances. As a result, the law that all ravens are black entails that this particular bird must be black, because (1) it is a raven and (2) anything that instantiates the universal raven must also instantiate the universal black (due to the necessitation relation between the two universals and the principle that whatever relations hold between universals must also hold between their instances).

However, again, one might still find this response not to be entirely satisfactory, since it may not be completely obvious that whatever relations hold between universals must also hold between their instances. For one must keep in mind that if Armstrong’s position is to represent a distinct alternative to Humean accounts, then a universal cannot be simply the set of instances that exemplify a certain feature. But if a universal is not identical

⁶⁹ *Ibid.*, pp. 421–422.

to its particular instances, then it becomes less clear that Armstrong can back up his assertion that the principle that whatever relations hold between universals must also hold between their instances is in fact analytic. It is true that Armstrong could back off the claim to analyticity and maintain that the entailment relation still holds, but it is also true that Armstrong's characterization of the relation between a universal and its instances has been slight enough to make the grounds for maintaining an entailment relation between the two less than completely transparent.

Let us now return to Kant. Two features of Kant's views about various kinds of laws are directly relevant to the debate between Humean and necessitarian views. First, we noted that Kant accords a privileged status to certain laws by maintaining that they are transcendental and hence a priori laws. For example, as we saw above, the Analogies of Experience are transcendental (rather than empirical) laws that make possible experience of temporal relations, which are, in turn, essential to the possibility of experiencing anything that would count as a single world. Though our primary focus has been on the Analogies of Experience, it is clear that the Laws of Mechanics, which are discussed in the Mechanics of the *Metaphysical Foundations of Natural Science*, have a similar status. Since these laws first make experience possible, they cannot in turn depend on whatever instances we happen to experience in the world. Second, we saw that Kant's views on freedom relied on the idea that the laws of nature do not depend on whatever events happen to occur in the world; rather, the laws of nature depend directly on the natures of things. Both of these claims support the idea that Kant would be in fundamental agreement with Armstrong that the Humean account of the status of the laws of nature ought to be rejected.

However, what is perhaps more significant is the fact that Kant's model of causality can provide a somewhat different perspective on Armstrong's view, a perspective that sheds important light on the objections that have been raised against Armstrong's view. Recall that according to Kant's model of causality, a substance determines the state of another substance by exercising its causal powers according to its nature and the external circumstances in which it exists. Thus, to use a familiar example, one body attracts another body by exercising its attractive power according to its mass and the distance between the two bodies. (Of course, the other body does the same so that action is equal to reaction and the masses of both bodies are relevant to the net force exercised.) While one could attempt to assimilate this view to Armstrong's by saying that there is a law of nature that relates the masses of the respective bodies (with the masses being

treated as universals), it is important to see that the further resources invoked in Kant's model of causality can yield a somewhat different point of view on the objections raised against Armstrong's position. For as we saw above, what is essential to Kant's account is the idea that one substance *act* on another, and because an activity is not essentially an event that occurs at some point in space-time, but rather something that happens according to a rule, the substance's nature supplies the rule that the activity requires (along with the relevant external circumstances).

One can thus draw on the resources inherent in Kant's model of causality to develop different and perhaps more plausible responses to the objections that Armstrong's view encountered. As for explaining the kind of necessity that is involved in the laws of nature, one can immediately note that, unlike some, Kant would not be hostile to the very idea of there being a necessary relation between universals, since he is sympathetic to the more general idea that there are necessary relations in nature. For it is a distinctive feature of the categories that they are able to represent necessary connections. If there were no necessary connections in nature, the categories would be of no use. However, Kant would not say that the necessity expressed by the laws of nature is causal. That is, it is not the mass of a body that causes a certain kind of attraction in the other. Rather, it is the body acting in accordance with its mass that causes the other body to be attracted. The relation between the universal mass and the universal attraction may be necessary, but it is not causal, even if it depends on causality in the body that has the mass. Instead, this necessity is based on the natures of the relevant bodies. (In contemporary contexts, this kind of necessity is, understandably, often called natural necessity.)

The Kantian response to the second objection raised against Armstrong's position proceeds in two different ways, depending on which version of the objection is under consideration. As regards the first version (concerning governing), the relevant Kantian claim is that the laws of nature do not govern what happens by directly causing these events (as a ruler might); rather, the laws of nature depend on the natures that things have, and it is these natures that govern the activities of substances in determining what happens. For the natures of things (and the laws of nature that are based on them) contain rules that govern how the activities they are engaged in determine what occurs in the world. This notion of governing cannot be causal because it is involved as a basic element in a model that is supposed to constitute causality in the first place. It is rather a *sui generis* sense of governing, since it is tied up with the very notion of activity, without therefore being mysterious.

The Kantian response to the second formulation of the objection, concerning how laws of nature explain what happens in the world, depends on first taking note of a peculiar feature of Armstrong's account. Because Armstrong holds that the relation between universals is causal (such that one type of thing causes another type of thing), he has, in effect, multiplied the number of instances of causation in the world. For in any given case of causality, there is (presumably) a particular thing that causes its effect and then, in addition, there is a kind of thing (to which that particular thing belongs) that causes its kind of effect.⁷⁰ The second version of this objection then inquires into why the causation relation between universals explains or entails a causation relation between the instances of the relevant universals. However, because Kant denies that the necessitation relation between the natures of things is causal, he is not forced to accept a plurality of causal necessitation relations and then try to derive one from the other. That is, because there are not two instances of causal necessitation between two different orders, one need not ask why the causal necessity of the one governs, explains, or entails the causal necessity of the other. Rather, Kant can simply say that what explains what happens in the world is the fact that substances exercise their causal powers in accordance with their natures, that is, in accordance with the laws or rules provided by their natures, and in response to their external relations to each other. In short, the laws of nature explain what happens because substances act in accordance with their natures and the laws based on them.

In sum, Kant's views of causality can contribute to the contemporary debate concerning laws of nature by providing one with a slightly different (and perhaps fuller) account of causation that then puts one in a position to provide more plausible responses to some of the central objections that have been raised against anti-Humean, necessitarian accounts. Specifically, because Kant grounds the laws of nature in the natures of substances and because substances must act in accordance with those natures, one can clarify both the kind of necessity that the laws of nature have (namely, natural rather than causal or logical necessity) and the sense in which a Kantian can say that the laws of nature govern what

⁷⁰ Armstrong might attempt to avoid this objection by denying that a particular thing causes its effect, which could make his view closer to a Humean position, according to which the basic regularities that occur in the world are not explained in terms of anything else, but are rather brute facts. If, by contrast, Kant's position can explain these basic regularities in terms of more primitive features, then his view would still seem to have a significant advantage (even if not the one identified above).

happens in the world without having to assert that the laws of nature are themselves directly causally efficacious.

Free Will and Determinism

Though the topic of free will and determinism was discussed in Chapter 5, the discussion there focused primarily on articulating Kant's own views on freedom and their relation to his general model of natural causality, and only a few issues from contemporary discussions were selected to clarify Kant's position. Our intention now is to see whether Kant's views can be put to use in a contemporary setting by suggesting a position that is able to respond in more satisfying ways to certain objections that have been raised in recent literature.

While there are many different ways of describing the issue of free will and determinism, one common way starts by distinguishing between causal determinism and indeterminism. Causal determinism is the thesis that every event is causally determined by previous events in accordance with the laws of nature. Indeterminism, by contrast, is simply the denial of causal determinism: Some events are *not* causally determined by previous events in accordance with the laws of nature. Causal determinists can then either deny the existence of freedom (on the grounds that causal determinism leaves no room for freedom) or be compatibilists (by holding that determinism and freedom can somehow be reconciled). Causal indeterminists can likewise either deny the existence of freedom (though not because determinism leaves no room for freedom, but perhaps rather because indeterminism does not leave a gap in exactly the right place) or accept libertarian freedom (by arguing that the causal gap permitted by the rejection of causal determinism opens up precisely the room that is needed for free actions).

While compatibilism has been perhaps the dominant position among analytic philosophers over the last several decades, significant attention has also been devoted to causal indeterminism, libertarianism in particular. In fact, the indeterminist position has been developed in three distinct ways. First, there are simple indeterminists, such as Carl Ginet, who think that we are free insofar as an *uncaused* mental event occurs that has an "actish *phenomenal* quality."⁷¹ Second, there are causal indeterminists, such as Robert Nozick and Robert Kane, who differ from simple indeterminists in asserting that our actions are caused by prior events,

⁷¹ Carl Ginet, *On Action* (New York: Cambridge University Press, 1990), p. 13.

but who claim that they are not caused by them *deterministically*.⁷² Third, there are proponents of agent causation, such as Roderick Chisholm, Richard Taylor, and, more recently, Timothy O'Connor, who hold that *agents* cause their actions and, in some cases, do so freely.⁷³

At the same time, each of these versions of indeterminism has come under attack.⁷⁴ The main challenge for simple indeterminism is that if a mental event of a certain kind (e.g., the forming of a volition or intention) is uncaused, then it is hard to see why that event should be thought to be under the control of an agent (since the agent could not, by hypothesis, have caused it). The mere fact that a certain mental event has an "actish phenomenal quality" is not sufficient to establish the presence of control, which is typically what is thought to be required for freedom and moral responsibility. Causal indeterminism, by contrast, seems to improve on simple indeterminism insofar as it allows that there is a cause of one's actions (even if only an indeterministic cause). However, the indeterministic element of causal indeterminism can appear to threaten agents' control over their actions, since agents do not determine whether or not they will cause their actions and when they will do so (if they do).

While the difficulties with simple indeterminism and causal indeterminism typically focus on features of control, the main objection leveled against causal agency theories is to the very notion of an agent. For example, C. D. Broad has objected to agent causation as follows:

[I]n so far as an event *is* determined, an essential factor in its total cause must be other *events*. How can an event possibly be determined to happen at a certain date if its total cause contained no factor to which the notion of date has any application? And how can the notion of date have any application to anything that is not an event?⁷⁵

⁷² Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981), and Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996).

⁷³ Taylor develops such a position in *Action and Purpose* (Englewood Cliffs, N.J.: Prentice Hall, 1966), while Chisholm develops the view in a series of works, starting in 1964 with "Human Freedom and the Self," in *On Metaphysics*, ed. R. Chisholm (Minneapolis: University of Minnesota Press, 1989), and "Freedom and Action," in *Freedom and Determinism*, ed. K. Lehrer (New York: Random House, 1966), pp. 11–44. In *Person and Object* (LaSalle, Ill.: Open Court, 1976), Chisholm seems to depart significantly from his earlier views. Timothy O'Connor articulates his position in *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000).

⁷⁴ See chap. 2 of O'Connor's *Persons and Causes: The Metaphysics of Free Will* for more detailed critical discussions of these versions of indeterminism.

⁷⁵ C. D. Broad, "Determinism, Indeterminism, and Libertarianism," in *Ethics and the History of Philosophy* (London: Routledge & Kegan Paul, 1952), p. 215.

Broad's objection is that since an agent is an enduring and nondatable thing rather than an event (or the state of a thing at a particular time), proponents of agent causation can give no explanation of why the effect that the agent brings about, namely e , occurred at t_0 rather than at any other time (given that there can be no difference in the agent at these times that could account for e 's occurring at that rather than any other time).

Timothy O'Connor has recently responded to Broad's objection by arguing that one should not "hold that there is an event that is e 's occurring at t_0 rather than at t_{-1} , in addition to e itself. There is," he says, "no reason to hold that corresponding to every contrastive fact about a contingent occurrence there is a distinct 'contrastive event,' where an event is understood to be a concrete entity."⁷⁶ In short, O'Connor rejects the idea that there is, in addition to e , a contrastive event that stands in need of explanation so that it is no defect if the agent causal theorist does not explain it.

However, two issues need to be kept distinct in O'Connor's reply. The first concerns the number of events that are relevant in the case in question. Does the existence of the effect require that there be only one event, e , or rather that there must be two events, e , plus the contrastive event of e 's occurring at t_0 rather than at t_{-1} ? The second concerns whether the occurrence of e at a certain moment in time is a feature of the effect that requires explanation. One can perhaps be agnostic about the first issue, since the question of how finely or coarsely events are to be described and thus whether there must be a distinct event for every contrastive fact is not obviously any more problematic for agent causation than it is for event causation.⁷⁷ However, the second issue is crucial for establishing the coherence of agent causation. For the real point of Broad's objection lies in the idea that the temporal index of an event requires explanation just as much as its objective content (i.e., in its being a red rather than a blue patch of color in a certain region of space), and proponents of agent causation cannot account for the temporal index of an event, Broad alleges, because they appeal not to a datable event at a particular time, but rather to an entity whose existence endures over an extended period of time.

⁷⁶ O'Connor, *Persons and Causes*, p. 76.

⁷⁷ David Lewis and Jaegwon Kim debate this issue in their articles cited above in notes 59–60.

Anticipating this reply, O'Connor points out that causal indeterminists are likewise unable to explain why indeterministically caused events occur when they do, and he infers from this that there is nothing incoherent in the very idea of rejecting the requirement that the temporal index of an event be explained. However, if one of the central objections to causal indeterminism is that it gives an agent no control over whether and when a cause brings about its effect, then a parallel concern would seem to arise for agent causation. Moreover, and regardless of whether O'Connor's solution faces the same kind of objection as does that of the causal indeterminist, if he rejects the idea that the temporal indices of events require explanation, it does create an asymmetry when compared with accounts that invoke only events and causal determinism. For proponents of deterministic event causation clearly *can* explain the temporal index of events. As a result, one can appreciate why Broad would think that accounts that invoke agent causation are *not* able to give as full an explanation of what happens as are determinist accounts that invoke only events.

It is at this point that Kant's account of causality can begin to be of help in the current debate. As we saw above (especially in Chapters 2 and 4), Kant thinks it crucial to distinguish between the cause and the external circumstances that must obtain for this cause to be efficacious in a particular way (i.e., to bring about an effect). As Kant emphasized in the *Nova dilucidatio*, since the determinations of substances are posited along with their unchanging grounds, only a change in external relations between such substances would allow them to bring about changes of state in each other. This position suggests the following response to Broad's objection. Although a cause, properly speaking, is an enduring substance with grounds that do not change, a substance can bring about different effects at different times due to the (changing) external relations in which it stands toward other substances, since the external relations function as the conditions that specify which effects a substance will bring about at any given time. Therefore, one can agree with Broad that *some* factor *relevant* to the cause must be a datable entity but hold that the datable factor is not itself a cause per se, but rather simply one of the conditions under which the cause is efficacious.

Now Broad and advocates of event causation will want to deny that one can distinguish in this way between an agent and the circumstances in which that agent is efficacious, since, on their view, all references to an agent can be translated into talk about events such that the total cause is simply the set of events that are necessary and jointly sufficient for the

effect. Broad is quite explicit about this point: "Of course I am well aware that we constantly use phrases, describing causal transactions, in which a continuant is named as the cause and no event in that continuant is mentioned. Thus we say: 'The stone broke the window.' . . . But it is quite evident that all such phrases are elliptical. The first, e.g., expresses what would be more fully expressed by the sentence: 'The coming in contact of the moving stone with the window at a certain moment caused a process of disintegration to begin in the window at that moment.'" ⁷⁸ That is, according to Broad one can reduce talk of agents to descriptions of datable events, and once one has done that, there is no reason not to view both agents and the circumstances under which they act as homogeneous and entirely datable parts of the complete cause of an effect.

It is clear from the understanding of Kant's reply to Hume described above, however, that a Kantian ought to reject the reductionist move that serves as the basis for denying his distinction between a cause and the circumstances necessary for its exercise. For from a Kantian perspective, talk about agents cannot be reduced to talk about events, since agents are active substances (endowed with a special set of causal powers, namely an understanding and a will) that contrast with passive states of substances or changes therein (i.e., events). And the fact that agents are essentially different from events provides a basis for distinguishing between a cause and the circumstances in which it is efficacious. As a result, by understanding agents as irreducible to events, it becomes clear that the circumstances in which an agent acts cannot be viewed as causes (or even parts of the total cause), since they are not active and hence cannot themselves *bring about* anything. Accordingly, one can avoid the incoherence Broad describes for accounts of agent causation without giving up the idea that the effect must be a datable event.

If the very idea of agent causation is thus not incoherent, as Broad's objection maintains, how exactly should agent causation be understood? Two main kinds of accounts of agent causation have been developed in detail in the contemporary literature. While all proponents of agent causation assert that an agent, *S*, causes an action or event, *e*, they part ways about whether *S*-causing-*e* is an event that is itself caused or uncaused. If one pursues the first option, as Chisholm and Taylor do, then one faces two problems. First, one runs the risk of an infinite regress, since whatever causes *S*-causing-*e* will itself have a cause, which will, in turn, require a further cause, and so on. Now one might suggest in response

⁷⁸ Broad, "Determinism, Indeterminism, and Libertarianism," pp. 215–216.

not that there is some other event that causes S-causing-*e*, but rather that S itself does so.⁷⁹ However, if S can bring about an effect only by means of the event S-causing-*e*, then the infinite regress has not been avoided, since S can cause S-causing-*e* only by means of S-causing-[S-causing-*e*], which is possible in turn only if there is a further event S-causing-[S-causing-[S-causing-*e*]], and so on. This kind of infinite regress is vicious and therefore must be rejected. Even if it were not, one could still doubt that citing S as the cause of the complex event S-causing-*e* is intelligible or actually explains why S-causing-*e* occurs. As O'Connor points out: "It is true that A's causing B would not have occurred if A hadn't, but this is merely a consequence of the fact that A is a constituent of the more complex event" of A-causing-B.⁸⁰ In other words, S can be used in an explanation of why *e* occurs, but not of why S-causing-*e* occurs, since to do so one would have to explain why S itself exists, an explanatory task in which S cannot be involved.

One could attempt to avoid these difficulties by stipulating that there be a *previous* event that causes S-causing-*e* (rather than either S itself or S-causing-[S-causing-*e*]). Now if such an event was itself caused by a further previous event (etc.), then one would, again, be forced to accept an infinite regress, though in this case it would extend back through time (rather than being internal to S-causing-*e*). Such an infinite regress would not necessarily be vicious, but it does suggest a different problem for this kind of account, namely that if S-causes-*e* is caused by a previous event, then S has no control over its actions (since S is caused to do so). This problem is particularly clear if S-causing-*e* is caused by a series of ever-distant prior events, since eventually the events from which S-causing-*e* follows will be located in the distant past, far from anything that S does. Therefore, accounts of agent causation according to which S-causing-*e* is an event that requires a cause encounter significant difficulties.

However, the second kind of account of agent causation, which O'Connor develops, seems equally unattractive. It is true that this kind of account does not immediately face an infinite regress (in the way in which Chisholm's does), since it accepts the fact that certain events are not caused, which means that one need not invoke further events outside one's control in explaining the cause of one's actions. However, one can object that claiming that the event of S-causing-*e* is uncaused is ad hoc.

⁷⁹ Chisholm seems to suggest such a position in "Replies," in *Philosophia* 8 (1978): 620–636, esp. p. 626.

⁸⁰ O'Connor, *Persons and Causes*, p. 58.

After all, if S-causing- e is a complex, but still datable event, then why does it not have a cause just as “regular” events do? In particular, if the event of S-causing- e occurs at t_0 , one would naturally think that something must have caused it to occur at precisely that time. Why should the event of S-causing- e be different in this respect from any other event?⁸¹

O’Connor develops the following argument for the claim that an agent-causal event (e.g., S-causing- e) cannot be caused (and is therefore unlike other events). He first presents an analysis of when it is permissible to say that one event has caused another event to cause a third, and then argues that the legitimate ways of speaking of event-causal causings (e.g., e_1 causing e_2 -causing- e_3) do not hold for agent-causal events being caused and that therefore agent-causal events are uncaused. To see how the argument works *in concreto*, consider O’Connor’s analysis of the example of when my finger presses the doorbell button, the doorbell rings, and your cat jumps in fright. In such a case one can say that my finger’s pressing the button causes the ringing of the bell’s causing the cat to jump, but when we speak this way we are not saying anything different from what is said when we say that one event causes a third indirectly, namely by causing a second event. But this kind of example, O’Connor argues, has no analogue in the context of agent-causal events, “because the cause within the causally complex event, *agent S’s causing e*, is not itself an event, but an enduring substance.”⁸² In other words, if one *can* say that e_1 causes e_2 -causing- e_3 because e_1 causes e_2 , which in turn causes e_3 , one *cannot* similarly say that e_1 causes S-causing- e_2 because e_1 causes S, which in turn causes e_2 , since e_1 cannot cause S, an enduring substance.

O’Connor is surely right to point to the disanalogies between event-causal events being caused and agent-causal events being caused. However, these disanalogies may not actually establish the inferences that he draws from them, since one could still question whether it is appropriate that O’Connor’s argument relies on an analysis of *event-causal* events in order to determine how *agent-causal* events can be understood. Specifically, why should one think that event-causal events should set the standard, so to speak, for what counts as intelligible for agent-causal events? In fact, one would naturally expect the opposite. Insofar as agent causation is conceived of as an alternative to event causation, differences in how

⁸¹ This point would seem to depend on the same point that was raised regarding O’Connor’s reply to Broad’s objection, since the main question in both cases is whether the temporal indices of events (whether complex or not) must be caused just as other features of events are.

⁸² *Ibid.*, p. 53.

agents and events bring about their effects should not surprise us. For O'Connor's argument to be cogent, he would have to provide reason for us to think that agent-causal and event-causal events ought to occur in exactly the same way (despite one's initial expectations to the contrary).

In fact, further reflection on the differences between agent-causal events and event-causal events suggests that O'Connor's account faces a more serious difficulty. One of the primary motivations for being a libertarian is to block the infinite regress of causes that entails that one's actions are caused by something other than the agent, which therefore opens up the possibility that the agent could decide whether or not to do a certain action and thus to act freely. So far, so good. But simply positing an uncaused event is not necessarily of any help (as the main objection to simple indeterminism made clear). One must make sure that it is one's action (or, in O'Connor's parlance, one's volition/intention) that is not caused by previous events. In other words, to capture what libertarians want, it is essential that what is uncaused by previous events is not the event of S-causing-*e*, but rather *e* itself.⁸³ In fact, this point is clear from O'Connor's own analysis because what the agent causes and thus can have control over is not S-causing-*e*, but rather *e*.⁸⁴ (If the agent caused S-causing-*e*, then it would do so in virtue of S-causing-[S-causing-*e*], and Chisholm's infinite regress would loom large. Further, it is the defining feature of O'Connor's view that S-causing-*e* is uncaused, which contradicts the supposition that S would cause S-causing-*e*.) But notice that if S-causing-*e* were uncaused (as O'Connor holds), but *e* were (also) caused by some prior event, *e*₁ (a possibility that is not excluded by O'Connor's account), then we would (at least in some instances) say that the action was determined and not free (which goes against what O'Connor wants to say). This counterexample shows that O'Connor's claim that S-causing-*e*

⁸³ It is crucial to distinguish between the agent, which is an enduring entity, the effect, which is an action (or intention/volition to perform that action), and the activity of the agent, by means of which the agent causes the action. (Above, we saw Kant refer to this final element by the phrase "the causality of the cause.") Although the agent's activity is in between the agent and its effect, it is distinct from the relation that O'Connor describes as S-causing-*e*, since that relation includes all three elements, whereas the agent's activity is simply *one* of the elements of the complex relation.

⁸⁴ O'Connor's argument for understanding S-causing-*e* as an event that is itself uncaused seems to rely on this point insofar as it assumes that S-causing-*e* cannot be caused because something would have to cause S or S's activity in order for S-causing-*e* to be caused. In other words, the problem is not that *e* cannot be caused (since in fact, it is caused on O'Connor's account, namely by the agent), but rather that the agent's causal activity cannot be caused.

must be an uncaused event is not in a position to secure the advantages libertarians hope to attain. Rather, what is supposed to accomplish significant work for this kind of libertarian is the notion of an agent (insofar as an agent is the sole cause of its actions).⁸⁵

At this point, one might, once again, be tempted to claim that agent causation is incoherent (or at least insufficient to its own purposes) and thus cannot be the central notion in a viable account of how we might act freely and hence responsibly. For if difficulties arise regardless of whether one assumes that *S*-causes-*e* is a caused or an uncaused event, it might seem that all the options available to proponents of agent causation have been considered and rejected as too problematic. However, such a conclusion is premature. To see why, recall Kant's general model of causality. As we saw in Chapter 4, Kant thinks that a substance brings about an effect by means of a temporally indeterminate, asymmetrical activity in accordance with a rule provided by the natures of the relevant substances and their external relations to each other. What is important about Kant's model of causality for current purposes is that he allows for the possibility that the activity by which a substance causes its effect is not itself an event, but rather what makes events (understood as changes of determinate state) possible.

If one understands agent causation analogously to Kant's model of causality (allowing for differences that stem from the special demands of the possibility of freedom), then one can simply deny that the agent's activity by which an event is brought about is itself an event, which allows one to avoid the difficulties that Chisholm's and O'Connor's accounts of agent causation encountered. One can bypass at the start the problems that Chisholm's account faced, namely the problems of an infinite regress of causes and the loss of control that results from allowing one's agency to be caused, since denying that an agent's causal activity is an event leaves one with no obvious reason to think that it would stand in need of a cause in the first place.⁸⁶ Further, this denial also allows one to avoid both of the difficulties of O'Connor's account. First, because the agent's causal

⁸⁵ One can also see the objection raised in this paragraph by considering the view of a determinist who thinks of our actions solely in terms of events. All that such a determinist needs to maintain is that every simple event is caused by a previous event (or set of events). That is, a determinist could admit that there are complex events, such as event *e*₁-causing-*e*₂, that are not caused, without such an admission posing a genuine threat to determinism.

⁸⁶ This is not to say that one could not argue that it might stand in need of explanation for some other reason.

activity is not a datable event, there is no temporal index that would be unexplained. It is true that the effect of this causal activity, namely *e*, can be ascribed a temporal index and thus stands in need of explanation, but that is precisely what *S* (along with the relevant circumstances) is invoked to explain, so that no difficulty arises in explaining that feature of how an agent causes its effect (or action). Second, this view locates indeterminism at the right place, namely in between the agent and his actions rather than in a complex event that includes this relation. For instead of asserting that the complex event of the agent causing its action is uncaused by previous events, one can say simply that one's action (*e*) is uncaused by previous events, though it is not, for that reason, entirely uncaused; it is caused by the agent and by the agent alone, which can secure the advantages that the libertarian claims for his account.

The idea that the activity by which an agent causes its action (e.g., exercises its will) is not an event is not, however, an entirely new one, though in contemporary contexts it is often attributed to Reid rather than to Kant.⁸⁷ Accordingly, it has received some critical attention. For example, O'Connor, after noting that Reid develops strict criteria of what is to count as an event, objects to this view as follows: "[T]here surely is a wider sense of the notion [of an event] . . . on which it [i.e., the obtaining of an agent-causal relation] is properly so called. After all, don't agent-causal relations obtain at certain times and not at others? If so, aren't such originating activities things that *happen*? For this reason, the agency theorist should demur from Reid on whether an agent's originating activity is an event."⁸⁸ That is, since the activities that agents engage in appear to obtain at some times and not at others, they should, O'Connor thinks, count as events.

Three points can be made in response to O'Connor's claim that the activities that agents engage in must be events (and thus that one must accept the broad construal of eventhood that he endorses). First, there is no compelling argument that would establish that the activities that agents engage in must be events. For example, one might argue that any such activity must happen at a specific time, since the effect it brings about, namely the action (or else the intention or volition to do the action), occurs at a specific time. However, the mere fact that the effect is temporally

⁸⁷ One important exception is Chisholm, who recognizes that the idea of agent causation goes back at least to Aristotle and was developed further by medieval thinkers such as Suarez. See Chisholm's "Human Freedom and the Self."

⁸⁸ O'Connor, *Persons and Causes*, p. 49.

determinate does not obviously entail that the cause is temporally determinate, since the temporal determinacy of the effect still leaves it open whether the cause is prior to or simultaneous with it, which suggests that it also leaves open the possibility that there could be no determinate fact of the matter, as Kant would hold.

Second, if the activity by means of which an enduring agent causes an effect (i.e., its action or its intention/volition to perform an action) were an event that occurred at a specific time, then one faces a serious dilemma. Since the activity occurs at a determinate time, it cannot be identified with the agent, because the agent is an enduring substance. But if the activity is thus an event distinct from the agent, then either it is caused by the agent or it is not. If it is not caused by the agent, then the agent will obviously not have control over it (granting the assumption that the concept of control contains the concept of causality as one of its necessary conditions). However, if it is caused by the agent, then one faces the very same problem that one encountered at the start. For either the agent causes the activity by means of some further event or it does not. If the former is the case, then (as we saw in the case of Taylor's view above) a vicious infinite regress looms large. If the latter is the case, then one has thereby accepted the idea that there is a primitive relationship between an agent and its causal effects, and it is most economical to do that not one stage in, but rather immediately, which is exactly what Kant's position does. The general point here is simply that as soon as one accepts an enduring entity and a relation between it and a temporally determinate event, one is saddled with questions about that relation that cannot be best answered in terms of temporally determinate events.

Third, even if one were to concede that the activities of agents occur at specific times, such a concession would still not directly entail that the activities in which agents engage are events. For temporality could simply be a necessary and not a sufficient condition for eventhood. In fact, Kant's view that substances are necessarily active and events (i.e., determinations) essentially passive can serve as an illustration of precisely such a position. For Kant's position is that the activity-passivity distinction is more fundamental than is the question of whether or not something is temporally determinate insofar as what makes something a temporally determinate event is the fact that it is passive, that is, is produced by an activity distinct from itself.⁸⁹ But the point at issue here relies not on

⁸⁹ Kant's distinction between activity and passivity is also relevant to establishing that the activity by which an agent brings about a cause cannot itself be caused. For if one

Kant's detailed views about activity and passivity, but rather on the fact that there can be more than one criterion for eventhood and on the possibility that different kinds of arguments and positions will lead to different criteria of what should count as an event. As we have seen in Chapter 4, Kant accepts the view that if there is some entity that makes an event possible, then that entity need not have all the traits that the event has and thus may not itself be a determinate event, and his argument is based on the idea that reality consists not in one flat layer of events, but rather in a two- or multilayered complex of events, substances, and relations among them that cannot be assimilated to either.

Kant's views on causality can thus be of considerable use in a contemporary setting by allowing one to develop an account of agent causation that can respond to objections commonly raised against it. Against Broad's objection to the very idea of agent causation one can follow Kant in distinguishing between the activity of the agent and the circumstances under which the agent undertakes that activity so that one can admit that there are datable factors relevant to a cause bringing about an effect without the cause itself being datable. Moreover, one can draw on Kant's model of causality to develop an account of agent causation that does not require the exercise of an agent's will to be viewed as a temporally determinate event. This allows one to avoid the infinite regress that ensues if the exercise of an agent's will is itself caused, while also allowing one to deny that it is an uncaused event not under the control of the agent.

Systematic Metaphysics

At this point, we have seen how Kant's views on causality can be directly relevant to contemporary discussions of the metaphysics of causality, the laws of nature, and agent causation. However, I would suggest that they have an additional value on the contemporary scene. For Kant's various reflections on causality not only provide insights into different issues currently being debated, but they do so in the service of a comprehensive and systematic metaphysical system.

understands activity such that nothing can *cause* anything else to be active (but rather merely satisfies conditions for it to be active), then it will be a conceptual truth that a substance's agency cannot be caused. At the same time, one can avoid being committed to the claim that an agent cannot be causally influenced by other things, since the causal powers that an agent has can be affected by other things. That is, even if a substance's nature is unchanging, the causal powers that are possible according to its nature can be restricted further by the changing state of the substance.

Kant's own statements of what he thinks he has accomplished with his Critical philosophy are both well known and, in my view, something of an exaggeration. In the Preface to the first edition of the *Critique*, for example, Kant claims that he can establish his philosophy with completeness, comprehensiveness, certainty, and clarity (Axiv–xviii). And although Kant's second edition Preface is in some ways more modest, his ambitious claims cannot be attributed to a momentary lapse of judgment while composing the first edition Preface, since he reiterates and articulates in greater detail several of these claims at the very end of the *Critique* in the Doctrine of Method. In the Architectonic of Pure Reason, he defines systematic philosophy in terms of "the unity of the manifold cognitions under one idea . . . [that is] the rational concept of the form of a whole, insofar as through this the domain of the manifold as well as the position of the parts with respect to each other is determined *a priori*" (A832/B860). While Kant thus claims that his philosophy is systematic by being ordered *a priori* according to a single idea, one must, I think, admit that he does not identify clearly enough what that idea is nor how the various parts of his system are ordered by it.

But even if one questions whether Kant can really make good on such far-reaching claims, it is clear that he does not attempt to solve just one particular problem at a time (in abstraction from all others), as is commonplace in contemporary analytic philosophy. In fact, in the previous sections of this chapter that focus on contemporary discussions, we have seen how Kant's general model of causality can be fruitfully used to address at least three distinct issues, issues that are, for the most part, discussed by completely different groups of philosophers.⁹⁰ What is particularly striking, however, is not only the comprehensive scope of the contributions that can be made by drawing on Kant's views on causality, but also the fact that a single notion of activity or determination that lies at the core of his views is central to these contributions. In considering the metaphysics of causality, Kant's notion of activity turns out to be the concept that must be used to distinguish event causation from causal powers. In considering how to understand laws of nature as distinct from Humean regularities, it is necessary to appeal to Kant's notion of activity in order to explain with what kind of necessity universals might be related to each other and how such a necessitation relation can be relevant to what actually happens in the world (i.e., to explaining how the laws of

⁹⁰ Lewis is a clear exception to this claim insofar as the scope of his view is very comprehensive.

nature can “govern” the events that fall under them if they are just expressions of relations between universals). Finally, concerning the issue of free will and determinism, certain central features of Kant’s notion of activity were important in helping us to see why the very idea of agent causation can be saved from objections that have been raised against it and how to develop a more detailed account of agent causation that can avoid problems that have been raised against other versions of it in the literature. Even if Kant’s notion of activity does not amount to a single idea that would structure his entire philosophy, it is clear that it is a crucial notion that one can draw on in addressing a wide range of philosophical problems from a unified perspective.

It is true that nothing that has been said so far amounts to a defense, much less a justification, of a systematic metaphysics based on the notion of activity that is central to Kant’s model of causality. Nor has any reason at all been given for the necessity of a systematic metaphysics, the establishment of which has fallen out of favor as of late. However, without calling for the kind of system-building characteristic of German Idealism, one can still find value in the articulation of a *single* philosophical framework that presents an *alternative* to the empiricist (e.g., Humean) outlook that has dominated contemporary analytic philosophy since the era of the logical positivists.

CONCLUSION

In this chapter, we have discussed how Kant’s account of causality relates both to Hume’s views and to those of certain contemporary Humeans and anti-Humeans. Consideration of how Hume’s philosophy was received in Germany in the eighteenth century and of Kant’s own explicit statements about Hume in the *Critique* has helped us to put Kant’s arguments about causality in a larger historical perspective. Given that the German audience for which the *Critique* was written did not, in general, think that Hume’s position stood in need of refutation and given that Kant’s principal arguments – in the Second and Third Analogies and with respect to the model of causality they employ – make assumptions that Hume would have rejected, it became clear that Kant did not intend to refute Hume’s position. Instead of trying to refute Hume on Hume’s own terms, we saw that Kant develops a different set of concepts – by borrowing but also modifying significant elements from the Leibnizian tradition in which he was educated – in order to present an alternative account of causality that competes against Hume’s. Thus, it would be a mistake to try to

reduce Kant's position to Hume's such that Kant's position is very much like Hume's, but with the addition of necessary connections where Hume has none. Rather, each account must be understood on its own terms.

Understanding Kant's account of causality in this way has also helped us to see one way in which it can be relevant today. Rather than attempting to find resources in Kant that would lead to a refutation of contemporary Humean positions, we found that one can use Kant's position to illustrate more clearly what the primitive concepts are on which contemporary Humeans and anti-Humeans might differ. Both Lewis's Humean and Harré and Madden's anti-Humean accounts of causality (which invoked a certain theory of events and a particular view of causal powers, respectively) were illuminated by Kant's perspective in this way. But we also found that one could use the insights that are worked out in Kant's position to develop more plausible non-Humean accounts. Thus, Kant's position was of help in responding to several important objections to Armstrong's necessitarian account of the laws of nature and to Chisholm's and O'Connor's accounts of agent causation. Finally, in the course of seeing how Kant's position on causality was relevant to several different contemporary issues, it became clear that appeal was repeatedly made to one and the same set of basic concepts that are employed in his model of causality (the notion of a substance exercising its causal powers in accordance with its nature). This fact suggests that Kant's position could be used to articulate a comprehensive metaphysical system that presents a fundamentally different alternative to Humean approaches. In these ways, Kant's reflections on causality can be seen to be of interest and importance both historically and philosophically.

Conclusion

The goal of this book was to present Kant's views on causality in their proper historical context. One can abstract the following general picture from the details of the historical and philosophical arguments that have been presented in the preceding chapters. Kant was keenly interested in articulating a comprehensive account of causality throughout his career that could (1) account for our knowledge of temporality, (2) explain how freedom can be reconciled with determinism, and (3) be used to formulate a response to Hume's skeptical arguments. Because Kant's account of causality also has implications for his views on other topics, such as his philosophy of science, it constitutes a central element not only of the *Critique of Pure Reason*, but also of his philosophical thought more generally, insofar as it develops the resources with which he hopes to describe how different kinds of beings can act in different kinds of contexts.

In his first pre-Critical works, Kant held that causality in the form of mutual interaction is required for substances (1) to be connected so as to form a single (spatio-temporal) world and (2) to bring about changes in their states. For he argued (1) that causally isolated substances would, as Leibniz asserted, be "worlds apart" rather than members of a single world, and (2) that such substances would also, as Leibniz denied, be just as static and unchanging as the essential grounds that must be posited for these substances to exist in the first place. Moreover, Kant recognized that while adding causal relations to otherwise isolated substances does allow them to form a single world, this, *by itself*, does not allow one to explain how these substances could change. For if the unchanging grounds within a substance cannot bring about a change, then the unchanging

grounds of other substances would seem to be equally impotent to produce change. As a result, Kant argued that substances' states can change only if the relations between them are changing. That is, Kant maintained that causal interaction between substances can be responsible for change only if the external circumstances or conditions in which they exist are relevant to how they act, that is, to which of their grounds determine the states of other substances.

The pre-Critical Kant also thought that since neither causal nor noncausal (e.g., spatio-temporal) relations between substances are required for the mere existence of these substances, such relations require a special ground beyond (the grounds of) the substances that are to be related, a ground that would be able to coordinate their activities and states. At this point, Kant argued that only God, who creates substances in the first place, is in a position to coordinate the relations between them and can thus serve as the requisite special ground. This line of thought distinguishes Kant's pre-Critical account of causality from the main anti-Leibnizian accounts developed by Crusius and Knutzen, both of whom did not specifically maintain that God must serve as a special ground for the causal interaction between substances. Later in the pre-Critical period, Kant, aided by the German translation of Hume's *Enquiry Concerning Human Understanding*, also came to acknowledge that causality cannot be a purely logical relation. For substances, understood as independently existing entities, cannot logically entail each other. Instead, Kant introduced the notion of a *real* ground and understood causal relations in terms of it, though he did not immediately grasp which principle real grounds are based on (in the sense in which logical grounds are supported by the principle of identity).

In his Critical period, Kant continued to assert a close connection between causality and temporality but changed his views in a number of important respects. First, in line with his Critical denial that we can have knowledge of things in themselves, Kant backed away from claiming to be able to know (in theoretical contexts) that God coordinates causal interaction between substances (though in practical contexts he argued that God must exist so as to coordinate virtue and happiness in conformity with the highest good). Second, while Kant continued to claim that causality is necessary for temporality, he embedded this claim in a broader, epistemological context, since temporality (specifically, succession and coexistence) is an essential element of the kind of experience whose possibility is the object of investigation in the *Critique of Pure Reason*. That is, Kant argued that causality and mutual interaction between

substances are necessary if we are to have knowledge of objective succession and co-existence, which is, in turn, itself necessary if we are to have knowledge of a single (spatio-) temporal world, or, as Kant sometimes puts it, *one experience*. Accordingly, Kant's arguments in the Second and Third Analogies unite ontological and epistemological elements by claiming that certain ontological commitments – causality and mutual interaction – are presupposed by the epistemic claims concerning our knowledge of succession and coexistence that are central structural features of our experience.

Despite these (and other) differences between Kant's pre-Critical and Critical views, it is quite striking that he still maintained several fundamental features of the model of causality he had developed early on. For even in the Critical period Kant held that substances act as causes and that they do so according to unchanging essential grounds (A205/B250). As a result, rather than thinking of causes and effects as events, as Hume often did, Kant held that one substance determines the state of another by actively exercising its causal powers in accordance with its essential grounds (or nature) and the (changing) circumstances in which they exist. Only by means of substances acting in this way is our experience of temporality possible.

Kant's views on causality are also relevant for his position on the issue of freedom and determinism. The first point to note is that several features of Kant's model of causality are employed in his account of free actions as well. Specifically, just as Kant understands phenomenal causality in terms of a substance acting according to its essential grounds so as to determine the state of another substance at a particular time, so too free actions are to be conceived of in terms of a substance, or agent, acting according to the intelligible nature, or character, that it chooses for itself so as to determine its actions in the world. Whether one talks of substances or agents, causal powers or faculties, essential natures or characters, determining or acting, the same kind of basic ontological structure is instantiated in both cases, and that structure is represented by the same set of concepts, namely the categories of substance and causality and the predicables of activity (*Handlung*) and power or force (*Kraft*).

Second, despite these structural parallels, Kant stressed a crucial dissimilarity between phenomenal causality and free actions. In the context of freedom, he divorced causality from temporality by arguing that free actions cannot be events in the phenomenal world. For, as he argued in the Third Antinomy, the phenomenal world's temporality entails the truth of determinism, which contradicts the absolute spontaneity required for

an action to be free. As a result, Kant inferred that if free actions are to be possible at all, then they must occur in the noumenal world to avoid conflict with determinism, which reigns over the phenomenal world.

At the same time, one could object that while positing freedom at the noumenal level does in fact avoid the threat of determinism, it does so at the cost of making freedom seem entirely superfluous insofar as it can appear to be irrelevant to what happens in our ordinary lives in the phenomenal world. Now, *if* Kant's model of causality consisted simply of temporally determinate events (and if laws of nature were simply regularities among such events), then the presence or absence of freedom at the noumenal level would in fact be irrelevant to determining what happens in the phenomenal world, because every event in the phenomenal world would be fully determined by prior temporally determinate events and the laws of nature (which ultimately reduce to a larger class of such events). However, Kant's position can be saved from this objection because his model of phenomenal causality does *not* consist only of temporally determinate events, but rather asserts that substances exercise their causal powers according to their natures such that temporally determinate states occur, and this, in turn, makes it possible for freedom at the noumenal level to have an influence on, and thus be relevant to, what occurs at the phenomenal level. For (1) the *activities* of phenomenal substances that determine events in the phenomenal world are not temporally determinate events and are also not caused by previous events, and (2) the *natures* according to which phenomenal substances act are not caused by previous events, but rather are grounded in free acts at the noumenal level. Since causality at the phenomenal level can be influenced in these ways by noumenal beings, locating freedom in the noumenal world does not necessarily make it irrelevant to what happens in the phenomenal world.

Finally, these remarks about noumenal freedom and phenomenal causality allow one to see how Kant can solve the apparent *modal* conflict between the *contingency* of free action (implicit in the idea that an agent could have done otherwise) and the *necessity* of events that is entailed by determinism (implicit in the idea that given the laws of nature and a complete set of initial conditions, all other events follow necessarily). Specifically, he can maintain that the necessity of determinism is not itself a brute fact; rather, it can depend, *at least in part*, on our free actions. For while it is agreed on all sides that determinism relies on the laws of nature, Kant understood the laws of nature as laws that are based on the natures of things, and since the natures of things can include our

freely chosen natures (which each of us typically calls our character), our free choices *can* determine, at least in part, what the laws of nature are and thus what occurs of necessity in the phenomenal world as well. As a result, Kant's views on causality provide him with conceptual resources that make it possible to resolve the modal conflict between freedom and determinism.

While many important details have been left out of this general picture of Kant's views on causality, it can still present us with the main outlines of his reply to Hume. Since several fundamental elements of Kant's model of causality remained the same throughout his career, one can see that the replies to Hume's skeptical doubts that Kant could appeal to in the pre-Critical period were still available to him in the *Critique of Pure Reason*. First, Kant can agree that there is no *logically* necessary connection between cause and effect, but still hold that a cause is a *real* ground of its effect, where a real ground is responsible for necessity of some kind (e.g., natural or physical necessity) between it and its effect, a necessity that, in the Critical period, must be represented by the categories. Second, while Kant can again agree with Hume's skeptical argument that our *empirical evidence*, which is based solely on past experience, is insufficient to establish that the future must resemble the past, he can still argue on *nonempirical* grounds that since substances must act in accordance with their unchanging essential grounds, they must continue to act in the future as they have in the past. That is, even if we do *not* know the exact content of the unchanging real grounds that have been the source of a wide range of empirical effects in different circumstances in the past, we *can* know that whatever grounds existed in the past must exist unchanged in the future and therefore that the future must resemble the past.

Although Kant can thus respond to Hume's famous skeptical doubts by appealing to central features of his model of causality (its use of real and unchanging grounds in particular), one might object that invoking these features illegitimately presupposes resources that go beyond what Hume would grant, and thus that Kant has no truly satisfactory reply to Hume (e.g., one that Hume himself would have to accept). However, this objection misses its mark insofar as it misunderstands the nature of Kant's most fundamental reply to Hume. There is ample historical evidence to believe that the audience for whom Kant was writing the *Critique of Pure Reason* would not have viewed Hume's position as standing in need of refutation in the first place, so that presupposing resources that Hume (would have) rejected is not an obvious and egregious mistake. In fact, instead of trying to *refute* Hume's radically skeptical or empiricist

position, Kant viewed the challenge facing him as that of articulating a viable *alternative* to Hume's position – one that calls on more traditional philosophical resources, such as substances, causal powers, and natures, while at the same time reconceiving our epistemic faculties on the basis of an analysis of pure reason itself in order to be able to combat the skeptical challenges that Hume and his “nomadic” followers may mount against this alternative, which could otherwise appear dogmatic.

Kant's ultimate response to Hume thus highlights a contrast between the two thinkers that is more fundamental than is often supposed. Hume and his followers started by assuming that the most basic level of reality consists of logically independent, temporally determinate entities, namely events, and then tried to construct the larger world (including the macroscopic objects of common sense, causal relations, laws of nature, etc.) out of such entities. By contrast, Kant held that temporally determinate entities are not the most fundamental level of reality, but rather themselves require further explanation, since, *prima facie*, the principle of sufficient reason holds for them insofar as one can ask not only why a certain event happened rather than some other one, but also why it happened at this point in time rather than at any other. Given this very different starting point, Kant reasoned that the only way to avoid a metaphysically unacceptable infinite regress is to posit something that is not itself temporally determinate, causal activity, as the ground of temporally determinate events. The notion of a temporally indeterminate causal activity that thus determines the state of a substance at a certain moment in time is then embedded in a broader metaphysical account, according to which the most basic level of phenomenal reality is not spatio-temporal events, but rather substances exercising their causal powers according to their natures and relations to each other so that temporally determinate states are possible in the first place.

While the metaphysical account of causality that Kant developed as an alternative to Hume's empiricist view thus constitutes a central part of his project in the *Critique of Pure Reason* of explaining what is required for the possibility of experience, it also has consequences for his Critical undertaking more generally. One prominent example of such a consequence can be found in the *Metaphysical Foundations of Natural Science*, where Kant attempted to show how the principles he had argued for in the *Critique* can be applied to the concept of matter so as to provide Newtonian physics with the kind of metaphysical foundation that he thought it requires. Now if his account of causality in the *Critique* is a central ingredient of his more general metaphysical position, then it is to be expected

that it will also bear on how the metaphysical principles of Newtonian physics are to be understood. Specifically, if Kant's account of causality invokes substances, causal powers, and natures, then it is likely that the metaphysical principles that Kant developed for matter will involve these same notions, even if they must be instantiated in more concrete ways. That is, it is plausible to suspect that in the *Metaphysical Foundations* Kant's conception of matter as what has mass and exercises its attractive and repulsive forces is best understood in terms provided by the metaphysical account of causality described above: Matter is simply spatial substance, the nature of this kind of substance is to have the mass it does, and the way in which one bit of matter exercises its attractive and/or repulsive forces in a certain situation so as to cause another bit of matter to change its motion is simply an instance of a particular kind of substance acting according to its nature and circumstances so as to determine the state of another substance at a particular time.

At the same time, highlighting the way in which Kant's metaphysical account of causality can be relevant to his account of matter in physics is not to suggest that his project in the *Metaphysical Foundations* amounts to nothing more than describing an abstract theory in more concrete ways. For Kant also hoped to accomplish several quite controversial goals in his transition project from the *Critique's* transcendental principles to empirical science, such as showing (1) that cognition, in order to qualify as science proper, must be systematically ordered according to rational principles and known a priori with apodictic certainty, that is, with "consciousness of their necessity" (4:468), (2) that such cognition is scientific only to the extent that it contains mathematics, since it is only through the construction of concepts in pure intuition (which is mathematical) that the possibility of an object can be established (4:470), and (3) that matter, as the object of physics, can be experienced only if certain principles (e.g., the laws of mechanics) that involve the categories obtain (4:472–475).

To understand Kant's transition project properly would require a detailed study in its own right, since one must not only document how Kant's metaphysical account of causality is applied to the concrete case of matter, but also determine how such a metaphysics of nature can be incorporated into an argument that establishes the additional goals just sketched. In fact, it is complicated further by the fact that one would also need to take into account how both Newton's *Principia* and Leibniz's writings on physics (e.g., the "Specimen Dynamicum") were received in eighteenth-century Germany and how Kant's pre-Critical publications

provide a helpful context for understanding his project in the *Metaphysical Foundations*. The central point for present purposes is simply that it is important to be clear from the start about the nature of the rich metaphysical framework that he assumed in his more scientific work and that he found useful, as we have seen, in a number of other contexts as well. In the end, therefore, we can see that Kant's account of the metaphysics of causality is not only fundamental to what he wanted to accomplish in the *Critique of Pure Reason*, but also relevant throughout his Critical project, because it reveals at a very basic level – one that guides his philosophical thought on many topics – how it is that any being must act.

Bibliography

PRIMARY LITERATURE

- Acht neue merckwürdige Schriften, die in der Wolffischen Philosophie von neuem erregte Streitigkeiten betreffend.* N.p., n.p.: 1737.
- Baumeister, Friedrich Christian. *Institutiones metaphysicae.* Wittenberg, 1743.
- Baumgarten, Alexander. *Metaphysica.* Reprinted in Immanuel Kant's *Gesammelte Schriften*, vols. 15,1 (pp. 5–54) and 17 (pp. 5–226). 1739. Rpt. Berlin: de Gruyter, 1902–.
- Bertram, Johann Friedrich. *Beleuchtung der Neu-ge-tünchten Meynung von der Harmonia Praestabilita durch Veranlassung der jüngst-edirten Reinbeckischen Erörterung.* Bremen, 1737.
- Bilfinger, Georg Bernhard. *De Harmonia animi et corporis humani maxime praestabilita ex mente illustris Leibnitii, commentatio hypothetica.* Reprinted in Abt. 3, Bd. 21 of Christian Wolff's *Gesammelte Werke.* 1723. Rpt. Hildesheim: G. Olms Verlag, 1984.
- . *Dilucidationes philosophicae de Deo, anima humana, mundo, et generalibus rerum affectionibus.* Reprinted in Abt. 3, Bd. 18, of Christian Wolff's *Gesammelte Werke.* 1725. Rpt. Hildesheim: G. Olms Verlag, 1982.
- Crusius, Christian August. *Entwurf der nothwendigen Vernunft-Wahrheiten.* Reprinted in Christian August Crusius, *Die philosophischen Hauptwerke*, ed. G. Tonelli, vol. 2. 1745. Rpt. Hildesheim: Georg Olms Verlag, 1964.
- . *Weg zur Gewißheit und Zuverlässigkeit der menschlichen Erkenntnis.* Reprinted in Crusius, *Die philosophischen Hauptwerke*, ed. G. Tonelli, vol. 3. 1747. Rpt. Hildesheim: Georg Olms Verlag, 1964.
- Darjes, Joachim. *Elementa Metaphysices*, vol. 2. Jena, 1743–1744.
- Euler, Leonard. *Gedancken von den Elementen der Körper.* Berlin, 1746.
- . *Lettres à une princesse d'Allemagne sur divers sujets de physique et de philosophie.* St. Petersburg, 1768–1772.
- . *Reflexions sur l'espace et le temps.* Berlin, 1748.
- Gottsched, Johann Christoph. *Erste Gründe der gesammten Weltweisheit.* Reprinted in *Ausgewählte Werke*, vol. 5, Part I. 1733–1734. Rpt. Berlin: de Gruyter, 1983.

- . *Vindiciarum systematis influxus physici*. Leipzig, 1727–1729.
- Hartmann, Georg Volckmar. *Anleitung zur Historie der Leibnitzisch-Wolffischen Philosophie*. Reprinted in Abt. 3, Bd. 4, of Christian Wolff's *Gesammelte Werke*. 1737. Rpt. Hildesheim: Georg Olms Verlag, 1973.
- Hollmann, Samuel Christian. *De Harmonia inter animam et corpus praestabilita*. Wittenberg, 1724.
- . *Dissertatio epistolica ad virum doctissimum Dn. Georg. Bernhardum Bülfingerum*. N.p., n.p.: 1726.
- . *Epistolae amoebae de harmonia praestabilita*. Frankfurt and Leipzig, 1728.
- . *Observationes elencticae in Controversia Wolffiana, disputatori cuidam Halensi, ad vindicandas suas de harmonia inter animam et corpus praestabilita habitas dissertationes*. Frankfurt, 1724.
- Hume, David. *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, revised P. H. Niddich. Oxford: Clarendon, 1978.
- . *Philosophische Versuche über die menschliche Erkenntniß von David Hume*. Reprinted in *Reception of the Scottish Enlightenment in Germany: Six Significant Translations, 1755–1782*, ed. H. Klemme. 1755. Rpt. Bristol: Thoemmes Press, 2000.
- Knutzen, Martin. *De immaterialitate animi*. Königsberg, 1741.
- . *Elementa philosophiae rationalis seu logicae*. Königsberg: Hartung, 1747.
- . *Philosophischer Beweis von der Wahrheit der Christlichen Religion*. Königsberg, 1740.
- . *Systema causarum efficientium*. Leipzig: Langenhemium, 1745.
- . *Vernünftige Gedanken von den Cometen*. Königsberg: Hartung, 1744.
- Lange, Joachim. *Anmerkungen über des Herrn Hoff-Raths und Professor Christian Wolffens Metaphysicam von denen darinnen befindlichen so genannten der Natürlichen und geoffenbarten Religion und Moralität entgegen stehenden Lehren*. Reprinted in Abt. 3, Bd. 23, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1986.
- . *Ausführliche Recension der wider die Wolfianische Metaphysic auf 9. Universitäten und anderwärtig edirten sämmtlichen 26. Schriften*. Halle, 1725.
- . *Bescheidene und ausführliche Entdeckung der falschen und schädlichen Philosophie in dem Wolfianischen Systemate Metaphysico von Gott, der Welt und dem Menschen*. Halle, 1724.
- . *Causa Dei et religionis naturalis adversus atheismus*. Halle, 1723.
- . *Modesta disquisitio novi philosophiae systematis de Deo, mundo et homine et praesertim de harmonia commercii inter animam et corpus praestabilita*. Reprinted in Abt. 3, Bd. 23, of Christian Wolff's *Gesammelte Werke*. 1723. Rpt. Hildesheim: G. Olms Verlag, 1986.
- . *Nova Anatome*. Frankfurt, 1726.
- . *Placidae vindiciae modestae disquisitionis de systemate philosophiae novo*. Reprinted in Abt. 3, Bd. 23, of Christian Wolff's *Gesammelte Werke*. 1723. Rpt. Hildesheim: G. Olms Verlag, 1986.
- Leibniz, G. W. *Briefwechsel zwischen Leibniz und Christian Wolff*, ed. C. I. Gerhardt. 1860. Rpt. Hildesheim: Georg Olms Verlag, 1971.
- . Letter to Rudolph Christian Wagner, June 4, 1710. Reprinted in *Die Philosophischen Schriften*, ed. C. I. Gerhardt. Berlin, 1875–1890. Vol. 7, pp. 528–532.

- . *Philosophical Essays*, ed. and trans. R. Ariew and D. Garber. Indianapolis: Hackett, 1989.
- . *Theodicy*, trans. E. M. Huggard. LaSalle, Ill.: Open Court, 1985.
- Malebranche, Nicolas. *The Search after Truth in Philosophical Selections*, ed. S. Nadler. Indianapolis: Hackett, 1992.
- Mauvertuis, Pierre. *Recherche des loix du mouvement*. Berlin, 1746.
- Meier, Georg. *Beweis der vorherbestimmten Übereinstimmung*. Halle, 1743.
- Mendelssohn, Moses. *Philosophische Gespräche*. Berlin, 1755.
- Newton, Isaac. *Mathematical Principles of Natural Philosophy and His System of the World*, trans. A. Motte, revised by F. Cajori. Berkeley: University of California Press, 1934.
- Plouquet, Gottfried. *Principia de Substantiis et Phaenomenis*. Frankfurt, 1753 and 1764.
- Reinbeck, Johann Gustav. *Erörterung der philosophischen Meynung von der sogenannten Harmonia Praestabilita, worinnen gezeigt wird, 1. was diese Hypothesis eigentlich sagen wolle, und warum sie der menschlichen Freyheit nicht nachtheilig sey. 2. Was dieselbe vor dem Systemate influxus für einen Vorzug habe, und 3. Warum der Autor nichts destoweniger derselben nicht beypflichte, aus Liebe zur Wahrheit und zur Verhütung fernerer verworrenen Streitigkeiten, nebst einem Vorbericht herausgegeben*. Berlin, 1737.
- Review of Martin Knutzen's *Systema Causarum Efficientium*, in *Zuverlässige Nachrichten vom dem gegenwärtigen Zustande, Veränderung und Wachstum der Wissenschaften*, Part 73. Leipzig, 1746. Pp. 48–67.
- Schopenhauer, Arthur. *Die Welt als Wille und Vorstellung*. In *Sämtliche Werke*, vol. 2. Wiesbaden: Eberhard Brockhaus Verlag, 1972.
- Sendschreiben an alethophilum, darin deutlich wird, daß der Herr Probst Reinbeck die Wolffische Meynung von der Harmoniae Praestabilita in der That angenommen habe, es nur nicht Wort haben wolle*. N.p., n.p.: 1737.
- Stiebritz, Johann Friedrich. *Erläuterungen der Wolffischen vernünfftigen Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt*. Halle, 1742–1743.
- Thümmig, Ludwig Philipp. *Institutiones philosophiae Wolfianae*. Reprinted in Abt. 3, Bd. 19, of Christian Wolff's *Gesammelte Werke*. 1725–1726. Rpt. Hildesheim: G. Olms Verlag, 1982.
- Tetens, Johann Nicolas. *Philosophische Versuche über die menschliche Natur und ihre Entwicklung*. Leipzig, 1777.
- Wolff, Christian. *Anmerkungen über das Buddeischen Bedenken von der Wolffischen Philosophie*. Reprinted in Abt. 1, Bd. 17, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1980.
- . *De differentia nexus rerum sapienties et fatalis necessitatis, nec non systematis harmoniae praestabilitae et hypothesium Spinozae luculenta commentatio*. Reprinted in Abt. 2, Bd. 9, of Christian Wolff's *Gesammelte Werke*. 1723. Rpt. Hildesheim: G. Olms Verlag, 1983.
- . *Eigene Lebensbeschreibung in Ausführliche Nachricht von seinen eigenen Schriften, die er in deutscher Sprache herausgegeben*. Reprint of the 2nd ed. in Abt. 1, Bd. 9, of Christian Wolff's *Gesammelte Werke*. 1726. Rpt. Hildesheim: G. Olms Verlag, 1973.
- . *Gründliche Antwort*. Reprinted in Abt. 3, Bd. 23, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1986.

- . *Klarer Beweis, daß der Herr D. Budde die von ihm gemachten Vorwürffe einräumen und gestehen muß, er habe aus Übereilung die ungegründete Auflagen der Hällischen Widersacher recht gesprochen.* Reprinted in Abt. 1, Bd. 19, of Christian Wolff's *Gesammelte Werke*. 1725. Rpt. Hildesheim: G. Olms Verlag, 1980.
- . *Monitum ad commendationem luculentam de differentia nexus rerum sapientis et fatalis necessitatis quo nonnulla sublima metaphysica ac theologiae naturalis capita illustratur, autore Christiano Wolfio.* Reprinted in Abt. 2, Bd. 9, of Christian Wolff's *Gesammelte Werke*. 1723. Rpt. Hildesheim: G. Olms Verlag, 1983.
- . *Nöthige Zugabe zu den Anmerkungen über Herrn D. Buddeus Bedenken von der Wolffischen Philosophie.* Reprinted in Abt. 1, Bd. 21, 4, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1981.
- . *Philosophia Prime sive Ontologia.* Frankfurt and Leipzig, 1729.
- . *Rede über die praktische Philosophie der Chinesen.* Hamburg: Felix Meiner, 1985.
- . *Sicheres Mittel wieder unbegründete Verleumdungen, wie denselben am besten abzuhelpfen.* Reprinted in Abt. 1, Bd. 21, 4, of Christian Wolff's *Gesammelte Werke*. 1723. Rpt. Hildesheim: G. Olms Verlag, 1981.
- . *Vernünftige Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt.* Reprinted in Abt. 1, Bd. 2, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1983.
- . *Der vernünftigen Gedancken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt, Anderer Theil, Bestehend in Ausführlichen Anmerkungen.* Reprinted in Abt. 1, Bd. 3, of Christian Wolff's *Gesammelte Werke*. 1724. Rpt. Hildesheim: G. Olms Verlag, 1983.

SECONDARY LITERATURE

- Adams, Robert. "Berkeley's 'Notion' of Spiritual Substance." *Archiv für Geschichte der Philosophie* 55 (1973): 47–69.
- . *Leibniz: Determinist, Theist, Idealist.* New York: Oxford University Press, 1994.
- . "Things in Themselves." *Philosophy and Phenomenological Research* 57 (1997): 801–825.
- Allison, Henry. "Causality and Causal Laws in Kant: A Critique of Michael Friedman." In *Kant and Contemporary Epistemology*, ed. P. Parrini. Dordrecht: Kluwer Academic, 1994. Pp. 291–307.
- . *Kant's Theory of Freedom.* New York: Cambridge University Press, 1990.
- . *Kant's Transcendental Idealism: An Interpretation and Defense.* New Haven: Yale University Press, 1982.
- Ameriks, Karl. "The Critique of Metaphysics: Kant and Traditional Ontology." In *The Cambridge Companion to Kant*, ed. P. Guyer. New York: Cambridge University Press, 1992. Pp. 249–279.
- . "From Kant to Frank: The Ineliminable Subject." In *The Modern Subject: Conceptions of the Self in Classical German Philosophy*, ed. K. Ameriks and D. Sturma. Albany: SUNY Press, 1995. Pp. 217–230.
- . "Kant and Short Arguments to Humility." In *Kant's Legacy: Essays in Honor of Lewis White Beck*, ed. P. Cicovacki. Rochester: University of Rochester Press, 2001. Pp. 167–194.

- . "Kant on Science and Common Knowledge." In *Kant and the Sciences*, ed. E. Watkins. New York: Oxford University Press, 2001. Pp. 31–52.
- . "Kantian Idealism Today." *History of Philosophy Quarterly* 9 (1992): 329–342.
- . *Kant's Theory of Mind*. New York: Oxford University Press, 1982.
- . "Kant's Transcendental Deduction as a Regressive Argument." *Kant-Studien* 69 (1978): 273–285.
- Armstrong, David. *A World of States of Affairs*. New York: Cambridge University Press, 1997.
- . "The Identification Problem and the Inference Problem." *Philosophy and Phenomenological Research* 53 (1993): 421–422.
- . *What Is a Law of Nature?* New York: Cambridge University Press, 1983.
- Arnoldt, Emil. "Möglichst vollständiges Verzeichnis aller von Kant gehaltenen oder auch nur angekündigten Vorlesungen nebst darauf bezüglichen Notizen und Bemerkungen." In *Gesammelte Schriften*, ed. Otto Schöndörffer. Berlin: Verlag von Bruno Cassirer, 1909. Pp. 173–344.
- Beck, Lewis White. "A Prussian Hume and a Scottish Kant." In *Essays on Kant and Hume*, ed. L. W. Beck. New Haven: Yale University Press, 1978.
- . *Early German Philosophy: Kant and His Predecessors*. Cambridge: Harvard University Press, 1969.
- Beiser, Frederick. *Kant and the Fate of Reason*. Cambridge: Harvard University Press, 1989.
- Bennett, Jonathan. *Events and Their Names*. Indianapolis: Hackett, 1988.
- . *Kant's Analytic*. New York: Cambridge University Press, 1966.
- . *Kant's Dialectic*. New York: Cambridge University Press, 1974.
- Bianco, Bruno. "Freiheit gegen Fatalismus: Zu Joachim Langes Kritik an Wolff." In *Zentren der Aufklärung I, Halle: Aufklärung und Pietismus*, ed. N. Hinske. Heidelberg: Lambert Schneider Verlag, 1989.
- Broad, C. D. [Charles Dunbar]. "Determinism, Indeterminism, and Libertarianism." In *Ethics and the History of Philosophy*. London: Routledge & Kegan Paul, 1952.
- . *Kant: An Introduction*. New York: Cambridge University Press, 1978.
- Buchdahl, Gerd. *Metaphysics and the Philosophy of Science*. Oxford: Basil Blackwell, 1969.
- Carl, Wolfgang. *Der schweigende Kant: die Entwürfe zu einer Deduktion der Kategorien vor 1781*. Göttingen: Vandenhoeck & Ruprecht, 1989.
- Carpenter, Andrew. "Kant's First Solution to the Mind/Body Problem." In *Kant und die Berliner Aufklärung*, ed. V. Gerhardt, R. Horstmann, and R. Schumacher. Berlin: De Gruyter, 2001. Vol. 2, pp. 3–12.
- Carr, David. *The Paradox of Subjectivity*. New York: Oxford University Press, 1999.
- Carrier, Martin. "How to Tell Causes from Effects: Kant's Causal Theory of Time and Modern Approaches." *Studies in History and Philosophy of Science* 34 (2003): 59–71.
- Cartwright, Nancy. *The Dappled World*. New York: Cambridge University Press, 1999.
- . *Nature's Capacities and Their Measurement*. Oxford: Clarendon, 1989.
- Chisholm, Roderick. "Freedom and Action." In *Freedom and Determinism*, ed. K. Lehrer. New York: Random House, 1966.

- . "Human Freedom and the Self." In *On Metaphysics*, ed. R. Chisholm. Minneapolis: University of Minnesota Press, 1989.
- . *Person and Object*. LaSalle, Ill.: Open Court, 1976.
- . "Replies." *Philosophia* 8 (1978): 620–636.
- Corr, Charles. "C. Wolff and Leibniz." *Journal of the History of Ideas* 36 (1975): 241–262.
- Davidson, Donald. "Actions, Reasons, and Causes." In *Essays on Actions and Events*, ed. D. Davidson. Oxford: Clarendon, 1980. Pp. 3–19.
- . "Causal Relations." In *Essays on Actions and Events*, pp. 149–162.
- . "Mental Events." In *Essays on Actions and Events*, pp. 207–227.
- Dretske, Fred. "Laws of Nature." *Philosophy of Science* 44 (1977): 248–268.
- Edwards, Jeffrey. *Substance, Force, and the Possibility of Knowledge: On Kant's Philosophy of Material Nature*. Berkeley: University of California Press, 2000.
- Eells, Ellery. *Probabilistic Causality*. New York: Cambridge University Press, 1991.
- Erdmann, Benno. *Martin Knutzen und seine Zeit: Ein Beitrag zur Geschichte der Wolfischen Schule und insbesondere zur Entwicklungsgeschichte Kants*. Leipzig: Verlag von Leopold Voss, 1876.
- Ewing, A. C. *Kant's Treatment of Causality*. Hamden, Conn.: Archon, 1924.
- Fabian, Gerd. *Beitrag zur Geschichte des Leib-Seele-Problems: Lehre von der prästabilierten Harmonie vom psychophysischen Parallelismus in der Leibniz-Wolffschen Schule*. Langensalza, 1925.
- Falkenstein, Lorne. "Kant's Argument for the Non-Spatiotemporality of Things in Themselves." *Kant-Studien* 80 (1989): 265–283.
- Fogelin, Robert. *Hume's Skepticism in the Treatise of Human Nature*. London: Routledge & Kegan Paul, 1985.
- Frank, Manfred. "Is Subjectivity a Non-Thing, an Absurdity [Uning]? On Some Difficulties in Naturalistic Reductions of Self-Consciousness." In *The Modern Subject. Conceptions of the Self in Classical German Philosophy*, ed. K. Ameriks and D. Sturma. Albany: SUNY Press, 1995. Pp. 177–197.
- Frankfurt, Harry. *The Importance of What We Care About*. New York: Cambridge University Press, 1988.
- Freddoso, Fred. "The Necessity of Nature." *Midwest Studies in Philosophy* 11 (1986): 215–242.
- Friedman, Michael. "Causal Laws and the Foundations of Natural Science." In *The Cambridge Companion to Kant*, ed. P. Guyer. New York: Cambridge University Press, 1992. Pp. 161–199.
- . *Kant and the Exact Sciences*. Cambridge: Harvard University Press, 1992.
- Garrett, Don. *Cognition and Commitment in Hume's Philosophy*. New York: Oxford University Press, 1997.
- . "The Representation of Causation and Hume's Two Definitions of 'Cause.'" *Nous* 27 (1993): 167–190.
- Gawlick, Günter, and Lothar Kreimendahl. *Hume in der deutschen Aufklärung: Umriss einer Rezeptionsgeschichte*. Stuttgart-Bad Cannstatt: Frommann-Holzboog, 1987.
- Geach, Peter. *Truth, Love, and Immortality*. London: Hutchinson, 1979.
- Ginet, Carl. *On Action*. New York: Cambridge University Press, 1990.

- Guyer, Paul. *Kant and the Claims of Knowledge*. New York: Cambridge University Press, 1987.
- Harré, Rom, and Edward Madden. *Causal Powers: A Theory of Natural Necessity*. Oxford: Blackwell, 1975.
- Henrich, Dieter. *Identität und Objektivität: Eine Untersuchung über Kants transzendente Deduktion*. Heidelberg: Carl Winter Verlag, 1976.
- Heßbrüggen-Walter, Stefan. "Die Seele und ihre Vermögen: Kants Metaphysik des Mentalen in der 'Kritik der reinen Vernunft.'" Ph.D. diss., Universität Münster, 2000.
- Hitchcock, Christopher. "The Role of Contrast in Causal and Explanatory Claims." *Synthese* 107 (1996): 395–419.
- Hogan, Desmond. "Cosmological Unity and Universal Interaction in Kant's *True Estimation of Living Forces*." Manuscript.
- Hudson, Hud. *Kant's Compatibilism*. Ithaca: Cornell University Press, 1994.
- Hughes, R. I. G. "Kant's Analogies and the Structure of Objective Time." *Pacific Philosophical Quarterly* 71 (1990): 141–163.
- Kane, Robert. *The Significance of Free Will*. New York: Oxford University Press, 1996.
- Keller, Pierre. *Kant and the Demands of Self-Consciousness*. New York: Cambridge University Press, 1998.
- Kim, Jaegwon. "Causation, Nomic Subsumption, and the Concept of an Event." *Journal of Philosophy* 70 (1973): 217–236.
- . "Causes and Counterfactuals." *Journal of Philosophy* 70 (1973): 570–572.
- Kintrup, Joachim. *Das Leib-Seele-Problem in Georg Bernhard Bilfingers Buch "De harmonia animi et coporis humani, maxime praestabilita, ex mente illustris Leibnitii, commentatio hypothetica (1723)" in der geschichtlichen und philosophischen Zusammenschau*. Münster: Münstersche Beiträge zur Geschichte und Theorie der Medizin, 1974.
- Kitcher, Patricia. "Kant on Self-Identity." *Philosophical Review* 91 (1982): 41–72.
- . *Kant's Transcendental Psychology*. New York: Oxford University Press, 1990.
- Korsgaard, Christine. *Creating the Kingdom of Ends*. New York: Cambridge University Press, 1996.
- Kuehn, Manfred. *Kant: A Biography*. New York: Cambridge University Press, 2001.
- . "Kant and His Teachers in the Exact Sciences." In *Kant and the Sciences*, ed. E. Watkins. New York: Oxford University Press, 2001. Pp. 11–30.
- . "Mendelssohn's Critique of Hume." *Hume Studies* 21 (1995): 197–220.
- . *Scottish Common Sense in Germany, 1768–1800: A Contribution to the History of Critical Philosophy*. Kingston and Montreal: McGill-Queen's University Press, 1987.
- Langton, Rae. *Kantian Humility*. New York: Oxford University Press, 1998.
- Laywine, Alison. *Kant's Early Metaphysics and the Origins of the Critical Philosophy*, North American Kant Society Studies in Philosophy, vol. 3. Atascadero, Calif.: Ridgeview, 1993.
- Lee, Sukjae. "Leibniz on Individual Substances and Causation: An Account of Divine Concurrence." Ph.D. diss., Yale University, 2001.
- Lewis, David. "Causation." In *Philosophical Papers*. New York: Oxford, 1986. Vol. 2, pp. 159–213.
- . "Causation as Influence." *Journal of Philosophy* 97 (2002): 182–197.

- . "Events." In *Philosophical Papers*. New York: Oxford, 1986. Vol. 2, pp. 247–254.
- . "Humean Supervenience Debugged." *Mind* 103 (1994): 473–489.
- . "Introduction." In *Philosophical Papers*. New York: Oxford, 1986. Vol. 1, pp. ix–xvii.
- . "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 71 (1983): 343–377.
- Liebing, Heinz. *Zwischen Orthodoxie und Aufklärung*. Tübingen: Mohr, 1961.
- Lodge, Paul. "Leibniz's Commitment to the Pre-established Harmony in the Late 1670s and Early 1680s." *Archiv für Geschichte der Philosophie* 80 (1998): 292–320.
- Loeb, Louis. *From Descartes to Hume: Continental Metaphysics and the Development of Modern Philosophy*. Ithaca: Cornell University Press, 1981.
- Loewer, Barry. "Humean Supervenience." *Philosophical Topics* 24 (1996): 101–127.
- Longuenesse, Beatrice. *Kant and the Capacity to Judge*. Princeton: Princeton University Press, 1998.
- Lorenz, Stefan. *De Mundo Optimo: Studien zu Leibniz' Theodizee und ihrer Rezeption in Deutschland (1710–1791)*, *Studia Leibnitiana Supplementa* 31. Stuttgart: Franz Steiner Verlag, 1997.
- Mackie, John. *The Cement of the Universe*. New York: Oxford University Press, 1973.
- Malzkorn, Wolfgang. *Kants Kosmologie-Kritik*. Berlin: de Gruyter, 1999.
- McRae, Robert. "Time and the Monad." *Nature and System* (1979): 103–109.
- Melnick, Arthur. *Kant's Analogies of Experience*. Chicago: University of Chicago Press, 1973.
- Mercer, Christia. *Leibniz's Metaphysics: Its Origin and Development*. New York: Cambridge University Press, 2001.
- Mercer, Christia, and Robert Sleight. "Metaphysics: The Early Period to the *Discourse on Metaphysics*." In *The Cambridge Companion to Leibniz*, ed. N. Jolley. New York: Cambridge University Press, 1995. Pp. 67–123.
- Morrison, Margaret. "Space, Time and Reciprocity." In *Proceedings of the Eighth International Kant Congress*. Milwaukee: Marquette University Press, 1995. Vol. 2, pp. 187–195.
- Nagel, Thomas. *The View from Nowhere*. New York: Oxford University Press, 1986.
- Nelkin, Dana. "Two Standpoints and the Belief in Freedom." *Journal of Philosophy* 97 (2000): 564–576.
- Nierhaus, Friedhelm. "Das Problem des psychophysischen Kommerziums in der Entwicklung der Kantischen Philosophie." Ph.D. diss., Universität Köln, 1962.
- Nozick, Robert. *Philosophical Explanations*. Cambridge: Harvard University Press, 1981.
- O'Connor, Timothy, ed. *Agents, Causes, and Events*. New York: Oxford University Press, 1995.
- . *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press, 2000.
- O'Neill, Eileen. "Influxus Physicus." In *Causation in Early Modern Philosophy*, ed. S. Nadler. University Park: Penn State Press, 1993. Pp. 27–55.
- Paul, L. A. "Aspect Causation." *Journal of Philosophy* 97 (2000): 223–234.
- . "Logical Parts." *Nous* 36 (2002): 578–596.

- Pearl, Judea. *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press, 2000.
- Pierris, Graciela de. "Hume's Pyrrhonian Skepticism and the Belief in Causal Laws." *Journal of the History of Philosophy* 39 (2001): 351–383.
- Prauss, Gerold. *Kant und das Problem der Dinge an sich*. Berlin: de Gruyter, 1974.
- Rauscher, Fred. "Kant's Two Priorities of Practical Reason." *British Journal of the History of Philosophy* 6 (1998): 397–419.
- Ravier, Emile. *Bibliographie de Oeuvres de Leibniz*. Hildesheim: Georg Olms, 1966.
- Rutherford, Donald. *Leibniz and the Rational Order of Nature*. New York: Cambridge University Press, 1995.
- . "Metaphysics: The Late Period." In *The Cambridge Companion to Leibniz*, ed. N. Jolley. New York: Cambridge University Press, 1995. Pp. 124–175.
- . "Natures, Laws, and Miracles: The Roots of Leibniz's Critique of Occasionalism." In *Causation in Early Modern Philosophy*, ed. S. Nadler. University Park: Penn State Press, 1993. Pp. 135–159.
- Sans, Georg. *Ist Kants Ontologie Naturalistisch?: Die "Analogien der Erfahrung" in der "Kritik der reinen Vernunft"*. München: Verlag Kohlhammer, 2000.
- Schönfeld, Martin. *The Philosophy of the Young Kant: The Precritical Project*. New York: Oxford University Press, 2000.
- Shoemaker, Sydney. "Causality and Properties." In *Identity, Cause and Mind*, ed. S. Shoemaker. New York: Cambridge University Press, 1984.
- . "Self-reference and Self-awareness." *Journal of Philosophy* 65 (1968): 555–567.
- Sleigh, Robert. *Leibniz and Arnauld*. New Haven: Yale University Press, 1990.
- Strawson, Peter. *The Bounds of Sense*. New York: Methuen, 1966.
- . *Individuals: An Essay in Descriptive Metaphysics*. New York: Methuen, 1959.
- Strawson, Galen. *The Secret Connexion*. Oxford: Clarendon, 1989.
- Stroud, Barry. *Hume*. London: Routledge & Kegan Paul, 1977.
- . "Transcendental Arguments." *Journal of Philosophy* 65 (1968): 241–256.
- Suppes, Patrick. *A Probabilistic Theory of Causality*. Amsterdam: North Holland, 1970.
- Taylor, Richard. *Action and Purpose*. Englewood Cliffs, N.J.: Prentice Hall, 1966.
- Tonelli, Giorgio. "Die Anfänge von Kants Kritik der Kausalbeziehungen und ihre Voraussetzungen im 18. Jahrhundert." *Kant-Studien* 57 (1966): 417–460.
- Tooley, Michael. "The Nature of Laws." *Canadian Journal of Philosophy* 7 (1977): 667–698.
- van Cleve, James. *Problems from Kant*. New York: Oxford University Press, 1999.
- van Fraassen, Bas. *Laws and Symmetry*. New York: Oxford University Press, 1989.
- van Inwagen, Peter. *An Essay on Free Will*. New York: Oxford University Press, 1983.
- Warren, Daniel. *Reality and Impenetrability in Kant's Philosophy of Nature*. New York: Routledge, 2001.
- Watkins, Eric. "The Argumentative Structure of Kant's *Metaphysical Foundations of Natural Science*." *Journal of the History of Philosophy* 36 (1998): 567–593.
- . "The Development of Physical Influx in Early Eighteenth Century Germany: Gottsched, Knutzen, and Crusius." *Review of Metaphysics* 49 (1995): 295–339.

- . “Forces and Causes in Kant’s Early Pre-Critical Writings.” *Studies in History and Philosophy of Science* 34 (2003): 5–27.
- . “From Pre-established Harmony to Physical Influx: Leibniz’s Reception in Early 18th Century Germany.” *Perspectives on Science* 6 (1998): 136–203.
- . “Kant on Force and Extension: Critical Appropriations of Leibniz and Newton.” In *Between Leibniz, Newton and Kant: Philosophy and Science in the 18th Century*, ed. W. Lefevre. Dordrecht: Kluwer Academic, 2001. Pp. 111–127.
- . “Kant on Rational Cosmology.” In *Kant and the Sciences*, ed. E. Watkins. New York: Oxford University Press, 2001. Pp. 70–89.
- . “Kant’s Antinomies: Sections 3–8.” In *Kooperativer Kommentar zu Kants Kritik der reinen Vernunft*, ed. G. Mohr and M. Willaschek. Berlin: Akademie Verlag, 1998. Pp. 445–462.
- . “Kant’s Theory of Physical Influx.” *Archiv für Geschichte der Philosophie* 77 (1995): 285–324.
- . “Kant’s Third Analogy of Experience.” *Kant-Studien* 88 (1997): 406–441.
- . “Kant’s Model of Causality: Causal Powers, Laws, and Kant’s Reply to Hume.” *Journal of the History of Philosophy* 42 (2004): 449–488.
- . Review of Daniel Warren’s *Reality and Impenetrability in Kant’s Philosophy of Nature*, in *Notre Dame Philosophical Reviews* (2002). [Http://ndpr.icaap.org/content/archives/2002/2/watkin-warren.html](http://ndpr.icaap.org/content/archives/2002/2/watkin-warren.html), accessed February 2002.
- . “Transcendental Idealism and the Categories.” *History of Philosophy Quarterly* 19 (2002): 191–215.
- Watkins, Eric, and Mark Fisher. “Kant on the Material Ground of Possibility: From *The Only Possible Argument* to *The Critique of Pure Reason*.” *Review of Metaphysics* 52 (1998): 369–395.
- Wilson, Fred. *Hume’s Defence of Causal Inference*. Toronto: Toronto University Press, 1997.
- Winkler, Ken. “The New Hume.” *Philosophical Review* 100 (1991): 541–579.
- Wolf, Susan. *Freedom within Reason*. New York: Oxford University Press, 1990.
- Wood, Allen. “Kant’s Compatibilism.” In *Self and Nature in Kant’s Philosophy*, ed. A. Wood. Ithaca: Cornell University Press, 1984. Pp. 73–101.
- Wright, John. *The Sceptical Realism of David Hume*. Minneapolis: University of Minnesota Press, 1983.
- Wundt, Max. *Die deutsche Schulphilosophie im Zeitalter der Aufklärung*. 1945. Rpt. Hildesheim: Georg Olms, 1992.
- Zeller, Eduard. *Geschichte der deutschen Philosophie seit Leibniz*. München, 1873.
- . “Wolff’s Vertreibung aus Halle; der Kampf des Pietismus mit der Philosophie.” *Preussische Jahrbücher* 10 (1862): 47–72.

Index

- accidents, 80, 188, 222, 262n36;
inherence in substance, 154–155,
166, 170, 259–261; knowability of,
222; migration of, 27–28, 66, 72,
94, 159, 163; not powers, 259–260,
262
- Acta Eruditorum*, 38, 39, 40
- action: at a distance, 176; divine,
63–68, 87–89, 96, 97, 98; as
ground of change, 251; Leibniz’s
definition of, 57; mechanics of, 79,
98; and motion, 104–106, 276;
personal element of, 303,
339–343, 360; as primitive, 65–66;
and reaction, 137, 165, 181, 250,
405; smallest, 9, 80
- activity: as equivalent to
substantiality, 30, 110, 126, 252,
256n29, 351–352; Kant’s notion
of, 13–14, 18, 19, 231–232,
265–282, 396–397, 420–421;
self-consciousness as providing the
meaning of, 272–282; sphere of,
109–111, 179. *See also* “causality of
the cause”
- agency, 15, 16; atemporal, 333,
333n32, 334; uncaused, 361,
412–417, 418. *See also* causation;
free will and determinism,
problem of; freedom
- Allison, Henry, 196n20;
epistemological interpretation of
Transcendental Idealism, 191n10,
319n15; reading of the “Antithesis”
of the Third Antinomy, 35, 309,
309n4, 310, 310n6; reading of the
Second Analogy, 203, 206
- alteration, 255–256
- Ameriks, Karl, 191n10, 195n19,
318n12
- Analogies of Experience: context in
the *Critique*, 11, 185–202; as
division in *Analytic of Principles*,
186–188; general principle of, 188;
and problem of
time-determination, 11, 185,
188–191; status of the claims of, 12,
196–202; analytical model, 196,
197–198, 200, 201;
epistemological model, 197,
198–199, 200, 201–202;
metaphysical-epistemological
model, 199–202; structure of, 188,
190; and *Transcendental*
Deduction, 185, 186–188. *See also*
First Analogy; Second Analogy;
Third Analogy
- Analytic of Concepts*, 186
- Analytic of Principles*, 181, 186; task
of, 186

- Antinomies of Pure Reason, 2, 15, 181, 194, 195n18, 301, 374; First Antinomy, 304, 314–315; format of, 305, 306; Fourth Antinomy, 304; Resolution of, 313–315; Second Antinomy, 304, 314–315. *See also* Third Antinomy
- appearances: definition of, 314n8; determinism in the realm of, 15, 217, 302, 315; as distinguished from things in themselves, 15, 302, 303, 308, 311–325; grounds of, 251, 302, 325–329, 334–339, 361; indeterminacy of, 16, 128n30, 251, 264n40, 313–314, 343, 349; and principle of sufficient reason, 305, 307, 314, 316, 327
- Apperception: awareness of activities in, 14, 273–282; distinguished from inner sense, 14, 141n40, 273–274; identity of, 276; as producing “I think,” 276, 279; unity of, 195, 202, 229. *See also* self-consciousness
- appetitions, 70, 95; Leibniz’s notion of, 35, 84; as primitive active forces, 35, 66
- apprehension: role of the categories in, 197; succession in, 189, 204, 206, 209, 219, 220; synthesis in, 218, 313–314
- a priori intuition, 172, 267, 379, 420
- a priori knowledge, 378–380; of mathematics, 197, 379, 384, 388; of metaphysics, 380; *See also* synthetic a priori knowledge
- Architectonic of Pure Reason, 172, 420
- Aristotle: on causal powers, 13, 251, 394; followers of, 2, 24, 195n18
- Armstrong, David, on laws of nature, 18, 392, 401–407, 422
- Arnauld, Antoine, 39
- attention, *see* self-consciousness
- Aufklärer*, 42
- axioms of intuition, 190, 417
- Baumeister, Friedrich Christian, 45, 51, 80
- Baumgarten, Alexander Gottlieb, 45, 97, 103, 119, 146, 165; account of real and ideal influence, 75–78, 83, 86, 149; defense of pre-established harmony, 51, 73–78, 97–98; *Metaphysica*, 60, 74–75, 78, 121n22, 146, 164, 200; on substance, 13, 110–172, 259
- Beattie, James, 368
- Beck, Lewis White, 10n4, 41n40, 163n67, 169n70, 203
- Beiser, Frederick, 367, 368
- Bennett, Jonathan, 186n1, 307n3
- Berkeley, George, 3, 279n59, 280n59
- Bilfinger, Georg Bernhard, 45, 51, 51n58, 69n96
- Broad, C. D., 240n16; objection to agent causation, 18, 243n18, 409–412, 414, 419
- Buchdahl, Gerd, 203
- Budde, Johann Franz, 51, 53
- calculus, 38, 39
- Cartwright, Nancy, 393, 395, 397
- categories, 191, 205, 212, 261, 266; normative status of, 182, 187–188, 209, 277; relational, 188, 190, 191–192, 196–197; table of, 182, 186, 191, 261, 266–267; unschematized, 186, 190n8, 193, 205, 251. *See also* category of causality; category of mutual interaction
- Category of causality, 4, 188, 207, 209, 219, 232, 267; attributing a spatio-temporal meaning to, 267–268, 324; and necessity, 212, 387, 406, 427
- Category of mutual interaction, 188, 219
- “causality of the cause,” 13, 231–232, 244, 253, 347, 349; as analogous to substance-inherence relationship, 261; distinguished from cause

- itself, 244, 249; indeterminacy of, 13, 17, 231, 257, 263n40, 296; meaning of, 244–249, 265
- causal laws, 211; how Kant establishes the necessity of, 216, 232, 286–291, 297; knowledge of, 198–199, 202, 213, 290; and “strong reading” of the Second Analogy, 203–204, 215–216, 286–287, 287n63
- causal powers: Baumgarten’s account of, 13 in *Critique of Pure Reason*, 252; Crusius’s account of, 51, 81, 82–85, 164, 220, 235; distinguished from event-event accounts of causation, 245–246, 397–400; powerful particulars, 394–396, 397. *See also* “causality of the cause”; grounds
- causation: agent, 14, 18–19, 243n18, 302, 304, 359–360, 361, 409–419, 425; atemporal, 263n40, 264n41, 324–325, 333; efficient, 2, 25, 37, 70–71, 151, 159, 181, 253, 332n29; event-event, *see* event-event causation; intrasubstantial, 55, 66, 67; mind-body, *see* mind-body problem; and principle of continuity, 7n3, 217, 231, 252, 255, 255n28, 257; simultaneous, problem of, 217, 231, 252–255
- Chisholm, Roderick, 19, 409, 412, 413, 415, 416
- coexistence: knowledge of, 12, 200–201, 217, 220–223, 227, 240–241, 282; of substance, 188, 201, 221–222, 227, 282–284. *See also* principle of coexistence
- community: category of, 188, 219; immediate vs. mediate, 227–228, 240n16; meaning of “thoroughgoing”, 227–228, 246, 247n22, 250; as necessary for knowledge of coexistence, 12, 199–201, 217–229, 294, 374; and principles of succession and coexistence, 12, 136–156, 180, 423, 425; role of God in bringing about, 141, 144, 148–156, 180
- complete concepts, theory of, 33–34, 49, 81, 94, 114n13, 148n53, 164, 189
- compatibilism, 329–335; Kant’s way of reconciling with incompatibilism, 334–338
- conatus, 54, 66. *See also* appetitions
- condition-conditioned relationship, 16n5, 210, 313–315
- consciousness, *see* self-consciousness
- conservation of motion: law of, 27, 68, 99; and physical influx, 46n48, 68–72, 89n128; problem of, 8, 27, 68–72
- contradiction, law of, *See* principle of contradiction
- cosmology, 51, 82, 83, 146, 194; importance in the *Critique of Pure Reason*, 182, 194, 195; in *Metaphysica*, 74, 146; in *Sketch of the Necessary Truths of Reason*, 83, 86, 89, 89n128; in the work of Christian Wolff, 41, 46, 48, 81, 95. *See also* world
- critical turn, 6, 11, 181, 183, 232, 282; according to “epistemological interpretation” of Transcendental Idealism, 319, 320–321; continuities and discontinuities with pre-Critical period, 11–12, 181–184, 229, 255n28, 291–296
- Critique of Judgment*, 7n3, 181, 255n28
- Critique of Practical Reason*, 181, 295n67, 315n10, 335n35, 336n39, 337n43, 342, 359n67
- Crusius, Christian August, 7, 8, 21, 50, 73, 81, 99, 113n10, 141, 200, 261, 371, 383, 386; arguments for physical influx, 82, 86–87, 89; on causal powers, 51, 81, 82–85, 164, 220, 235; conception of metaphysics, 81, 82, 83; eclecticism and anti-Leibnizianism of, 9–10, 24, 82, 99; inefficacious vs.

- Crusius, Christian August (*cont.*)
 efficacious grounds, 83–85, 86, 98, 103, 145, 180, 220; influence on Kant, 10, 24, 87n123, 123, 145, 145n49, 149, 156, 160, 164, 166, 169n70, 181, 261, 424; on Knutzen, 92, 93; on the mind-body problem, 51, 82, 89–93, 99, 107; objections to pre-established harmony, 82, 87–89, 92, 92n133, 93, 158, 174; *Path to the Certainty and Reliability of Cognition*, 92–93, 275; and pietism, 81, 93, 98; on the problem of free will and determinism, 82, 85, 98; real vs. ideal grounds, 83–89, 156, 158, 164–165, 169, 180, 200, 383; *Sketch of the Necessary Truths of Reason*, 81–92, 123; on space and time, 82, 85; on the world, 83, 85–89, 90n128, 98, 146–147, 149, 153
- Darjes, Joachim Georg, 45, 73, 89n128
- Descartes, René, 27n8, 33, 46n48, 68, 278, 281; causal theory of, 2–3, 24; on mind-body interaction, 2, 3, 26–27
- Discipline of Pure Reason, 375–381, 387–388
- Doctrine of Method, 183, 375–381, 383–384, 420
- Dretske, Fred, 401
- “Dynamical Principles”. *See* Analogies of Experience; Postulates of Empirical Thought
- empiricism, 2, 3, 5; Crusius’s sympathy with, 81; possibility of a systematic alternative to, 5, 17, 419–421; Second Analogy as a refutation of, 6; Tetens’s sympathy with, 368, 373
- “*Entelechia*,” 271n46
- Essence, 123; Crusius’s account of, 81, 82, 84, 87, 88, 90; existence vs. essence, 82–85; Kant’s account of, 10, 116, 123–125, 357; metaphysical, 81, 82; of the mind and body, 91, 91n131, 99; Wolff’s conception of, 120, 120n21, 122, 123, 133
- Euler, Leonard, 73, 73n100, 272n48
- events: Hume’s conception of, 17, 230, 232–235, 236, 245, 386, 391, 393; Kant’s conception of, 17, 236; Lewis’s conception of, 398–400; in the Second Analogy, 221, 284
- event-event causation, 12, 14, 230, 231, 232, 238, 245–246, 249, 284; simple models of, 238–239, 245, 246, 249; complex models of, 239–243, 245, 249; Lewis’s account of, 398–400; Third Analogy’s incompatibility with, 231, 237, 238–243, 284, 296
- experience, 11, 424; perceptual, 195; possibility of, 4, 11, 195, 379, 387, 405, 424, 428; scientific, 195; unity of, 193–195, 202, 229, 309
- First Analogy of Experience, 192, 221, 243, 268n45, 282, 288, 289
- Fisher, Christian Gabriel, 44, 45
- force, 102; activity as criterion of, 10, 102, 104, 179; attractive and repulsive, 136n35, 182, 224, 268–272, 272n48, 281, 289, 429; derivative, 9, 36–37, 53, 58–59, 62, 71, 97; living vs. dead, 108; primitive active and primitive passive, 9, 35–37, 58–59, 62, 71, 97, 112, 125–126, 271n46; and *vis motrix*, 54, 104, 179. *See also* activity; appetite; “causality of the cause”
- Frankfurt, Harry, 357–359
- Frederick the Great, 44, 101
- Frederick William I, 44
- freedom: exercised by noumenal agents, 250, 316, 333, 334–335, 336, 337; libertarian, 82, 85, 98, 408, 415, 416, 417; practical, 332; psychological, 343; “spontaneous cause,” 308–310, 316, 333; transcendental, 309, 332, 342.

- See also free will and determinism, problem of
- Free will and determinism, problem of: compatibilism, 329–332; and control, 340–341, 342, 343, 360, 409, 411; Crusius's solution to, 82, 85, 98; incompatibilism, 330, 332; Kant's solution to (in the resolution of the Third Antinomy), 311–316, 332, 341, 355; libertarianism, 82, 85, 98, 408, 415, 416, 417; location of free will, 357–359; and personal facts, 303, 339–343; and regress of desires, 354–357, 360. See also freedom; Third Antinomy
- Friedman, Michael, 196n20, 203–204, 287n63
- geometry, 109, 111
- Ginet, Carl, 408
- God: arguments for the existence of, 34, 113n10, 141, 156, 167; as cause, 2, 25, 29–30, 134–136; as harmonizer of monads, 2, 23, 25, 34, 37, 76; knowledge of, 4, 424; *The Only Possible Argument*, 103, 164, 167; as special ground of causal relations, 88, 98, 150–156, 175, 178, 180, 181, 424; see also action, divine; Voluntaristic conception of, 82, 88
- Gottsched, Johann Christoph, 45, 52, 53, 73
- Groundwork for the Metaphysics of Morals*, 326, 328
- grounds, 200; determining, 85, 113n10, 115, 132, 133; antecedently vs. consequentially, 113n10; definition of, 114n12; inefficacious (existential) vs. efficacious (active), 83–85, 86, 98, 103, 145, 180, 220; Kant's conception of, 10, 122, 123; real vs. logical, 103, 163–167, 169; real, 10, 103, 203; Leibniz's conception of, 125–129, 134, 180, 423; real vs. ideal, 83–89, 156, 158, 164–165, 169, 180, 200, 383; Wolff's conception of, 102, 103, 113n10, 118–121, 124, 133, 171, 180
- Guyer, Paul, 146n2, 187n2, 189n7, 192n11, 197–199, 201–202, 203–204, 206n38, 210–213, 218n52, 242n17, 253n27, 309, 385n40
- Hamann, Johann Georg, 367
- Harré, Rom, 393, 394–397, 422
- Henrich, Dieter, 206
- Hollmann, Samuel Christian, 52, 73
- Hume, David, 3–5, 7, 10, 11, 12, 16–18, 45n47, 163n67, 207, 233, 233n1, 235, 274n50, 401; on causal laws, 203, 216, 286; conception of events, 17, 230, 232–235, 236, 245, 386, 391, 393; *Enquiry Concerning Human Understanding*, 10, 16, 161, 164, 167n68, 181, 235n9, 364, 365; German reactions to, 16, 364–367, 373; importance for Kant (in leading him to real grounds), 103, 161–162, 167, 167n68, 169n70, 172, 373; Kant's explicit remarks about, 374–381; on miracles, 30, 31n21; model of causality, 12, 14, 230–231, 238, 239–245, 284, 384–386; on necessary connection, 3, 161–163, 169, 170–172, 181, 203, 216, 234–235, 274, 278, 296, 384, 385; on the self and self-knowledge, 277, 277n53, 277n54, 278, 278n56, 280; and skepticism, 7, 103, 181, 216, 367, 374–378, 379–380; *A Treatise of Human Nature*, 45, 233–235, 236, 253, 274n50, 277, 286
- Huygens, Christiaan, 39, 46n48
- idealism: Leibniz and, 35–36, 37, 47, 95, 128, 129, 159; Wolff's agnosticism about, 9, 47–48, 95–96, 99, 159. See also Refutation of Idealism; Transcendental Idealism

- identity: of apperception, 276, 278;
and composition, 174, 325n23;
conditions of, 29, 94, 262n35, 288,
290n64; personal, 359n67;
relations of, 370, 372
- inertia, 69, 97, 109, 226
- imagination: as distinctive human
faculty, 198; and representation of
causal relations, 370, 371; as
source of a priori representations,
379; syntheses of, 207, 218; unable
to yield knowledge of objective
succession, 208, 219
- Inaugural Dissertation*, 11, 103, 104,
160, 170, 172, 176n74, 178, 181,
194, 288, 292
- induction, 376
- influxus physicus*, *see* physical influx,
theory of
- inner sense, 44, 45, 51, 273, 273n49,
276; distinguished from
apperception, 14, 141n40,
273–275, 276, 368, 369–373;
distinguished from outer sense,
273, 368
- Journal des Savants*, 26, 39, 40, 170
- judgment, 165, 166, 169, 170, 181,
186, 187, 202, 213n45, 266, 373
- Kane, Robert, 408
- Knutzen, Martin: defense of physical
influx, 53, 68–73; *De immaterialitate
animi*, 52; *Elementa philosophiae
rationalis logicae*, 52; *Philosophischer
Beweis von der Wahrheit der
Christlichen Religion*, 52; and
pietism, 52; *Systema causarum
efficientium*, 53–58, 60, 67n88, 73,
108n6, 159; *Vernünftige Gedanken
von den Cometen*, 52
- Kraft*, *see* force
- Lange, Joachim, 42, 43n, 43n43, 51,
53, 73, 74n101
- law of the conservation of living
forces, 49, 68–69, 70–73, 89n128,
94, 96, 99
- law of the conservation of motion, 8,
46n48, 68, 89n128, 94, 99
- law of the equality of action and
reaction, 137
- law of inertia, 69, 97, 109, 226
- laws of nature: causal efficacy of, 18,
403; as contingent on free will, 15,
302, 303, 326, 334, 337, 344, 346;
as contingent on the natures of
things, 15, 18, 344, 345, 346, 406;
distinguished from accidental
generalizations, 346; “governing”
of, 18, 401–402, 403–407; Humean
 (“regularity”) accounts of,
345–346, 398, 401–402; Leibniz’s
account of, 31–32, 97, 145n48,
338; Malebranche’s account of, 31;
necessitarian accounts of, 18, 363,
401–408; necessity of, 18, 346, 403,
406; transcendental vs. empirical,
199, 203–204, 290, 405. *See also*
causal laws; law of the conservation
of living forces; law of the
conservation of motion; law of the
equality of action and reaction; law
of inertia; necessity; Second
Analogy, “weak vs. strong reading
of”; second law of motion
- Leibniz, Gottfried W.: on bodies,
36–37, 58–60, 325; definition of
action, 57, 99; “Discourse on
Metaphysics,” 33, 39, 40; doctrine
of complete concepts, 33–34, 49,
81, 94, 114n13, 148n53, 164, 189;
doctrine of pre-established
harmony, 2, 8, 23, 24–37, 75,
94–95, 96, 161, 163; “On an
Emendation of First Philosophy
and the Notion of Substance,” 40;
on forces, *see* monads, forces of;
on grounds of change, 125–129,
134, 180, 423; and idealism,
35–36, 37, 47, 95, 128, 129, 159; on
laws of nature, 31–32, 97, 145n48,
338; *Leibniz-Clarke Correspondence*,
40; “Meditations on Knowledge,
Truth, and Ideas,” 40; on
mind-body interaction, 2–3, 25,

- 26–28; on miracles, 30–32; “The Monadology,” 28, 39, 40, 149n55; on monads, *see* monads, Leibniz’s account of; “On Nature Itself,” 30–35, 40; *New Essays Concerning Human Understanding*, 40, 370; “A New System of Nature,” 26, 27n9, 28, 34, 40, 53n63, 59, 71, 143; objections to occasionalism, 25, 28–32; principle of identity of indiscernibles, 33, 49; real vs. ideal influence, 75–78, 83, 86, 88, 98, 143n45, 149; reception of, 7, 8–10, 23–24, 38–100; on relations, 131, 131n33, 133, 142, 142n42, 143n43, 145, 149, 353n58; relationship with Wolff, 9, 38–50, 95–96; “Specimen dynamicum,” 35, 36–37, 40, 71; on substance, 32–33, 94, 351; *Théodicée*, 39, 40, 75; universal expression thesis, 35, 46; “world apart” doctrine, 56–57, 98, 143, 423
- Lessing, Gotthold Ephraim, 367, 372
- Lewis, David, 18, 398–400, 401, 402, 403, 422
- Locke, John, 2, 3, 24, 35, 53n62, 251, 261, 272n48, 368, 370, 381
- Madden, Edward, 393, 394–397, 422
- Malebranche, Nicholas, 2, 32n25; conception of laws of nature, 31; definition of miracles, 31; doctrine of occasionalism, 2, 28, 31n23, 152; reason for accepting, 28–29
- Marburg, 44, 101
- “Mathematical Principles,” *see* axioms of intuition
- Maupertuis, Pierre, 44, 272n48
- Meier, Georg, 7, 8, 9, 10, 24, 45, 103, 118, 119, 180; defense of pre-established harmony, 9, 24, 51, 74, 77n107, 78–80, 98; *Proof of Pre-established Harmony*, 78
- Mendelssohn, Moses, 7, 81, 367, 372, 386
- metaphysical deduction, 13, 181, 182, 186, 187, 193n14, 209, 266, 267n44, 281–292, 296
- Metaphysical Foundations of Natural Science*, 7n3, 198, 226, 228, 250, 255, 264n42, 268, 284, 296, 315n9, 326n25, 405, 428–430
- Metaphysics Mrongovius Transcripts*, 257, 262n36, 265, 292, 294, 295, 326, 347
- mind-body problem: Crusius’s solution to, 51, 82, 89–93, 99, 107; formulation of, 26–27; Kant’s pre-Critical solution to, 102, 106–108, 166; Knutzen’s solution to, 64–65; Leibniz’s solution to, 2–3, 25, 26–28
- miracles, 25, 31, 46n48, 65, 94; different definitions of, 30–31, 31n21
- modern philosophy, traditional narrative of history of, 1–4
- monads: atemporality of, 126n28; and bodies, 36, 50, 58–60, 62, 69, 70–72, 93, 95–96, 97, 111–112, 325; as final causes, 35–36, 37, 70; forces of, 9, 35–37, 58–59, 62, 71, 97, 112, 125–126, 271n46; Knutzen’s account of, 53, 53n63, 58–60, 70–73, 96–97, 111; Leibniz’s account of, 25, 35–37, 47–48, 58–60, 60n77, 62–63, 70–72, 75, 95, 97; “The Monadology,” *see* Leibniz; perceptions and appetitions of, 35–36, 46–48, 50, 70, 75, 95; *Physical Monadology*, 102, 103, 109–112, 128, 129, 179; as mathematical points, 59; as metaphysical points, 26, 35, 47, 58–60, 97, 128; as physical points, 35, 50, 58–60, 62–63, 93, 97, 109–112, 128, 129n31, 179; as points of view, 35, 59, 325; as principles of unity, 35, 47, 59, 71, 102, 109, 179; spatiality of, 54n66, 58–59, 109–112, 179, 325; Wolff’s account of, 47–48, 95–96

- moral law, 328, 335
 mutual interaction, *See* community
- necessary connection, *see* Hume, on
 necessary connection; necessity
- necessity, 203, 329, 381, 426; causal,
 18, 171, 406, 407; logical, 18, 119,
 125, 171, 370, 382, 407; natural,
 18, 203, 332, 406, 407;
 Armstrong's account of, 401–403;
 Hume's denial of, 3, 203, 216;
 Kant's explanation of, 345, 346,
 382, 393–394, 406, 407, 427;
 metaphysical, 171, 290; role in
 Second Analogy, 211–213
- Newton, Isaac, 2, 109, 189, 192n13,
 198, 320, 429; and the calculus,
 39; on forces, 109, 268, 269, 271,
 272n48, 273; on inertia, 109; law
 of the equality of action and
 reaction, 137; second law of
 motion, 272; universal attraction,
 227
- noumena, *see* substance; things in
 themselves; Transcendental
 Idealism
- Nova dilucidatio*, 102–103, 112,
 113n10, 118, 121–125, 141–145,
 150–153, 154, 157–160, 172, 175,
 178, 179–181, 224, 251, 291, 411
- Nozick, Robert, 408
- occasionalism, 25; Leibniz's
 arguments against, 25, 29–32;
 Malebranche's motivation for
 accepting, 28–29
- O'Connor, Timothy, 409, 410–411,
 413–416, 417–419, 422
- outer sense, 273, 276, 369–373
- paralogisms, 181, 183n2, 279, 282,
 283, 296, 301, 359n67, 374
- phenomena, *see* appearances;
 substance; Transcendental
 Idealism
- physical influx, theory of:
 Baumgarten's arguments against,
 75; Crusius's defense of, 82, 86–87,
 89; Kant's qualified acceptance of,
 156–161, 177–178; Knutzen's
 defense of, 53, 68–73; Leibniz's
 arguments against, 26–28; Meier's
 arguments against, 78–80;
 migration of accidents model of,
 27–28
Physical Monadology, 102, 103,
 109–112, 128, 129, 179
 pietism, 42, 51n57, 95; and Crusius,
 9, 81, 93, 98; in German
 universities, 42; and Knutzen, 52;
 and Wolff, 42–44, 50, 51n57, 52
 points, *see under* monads
 Postulates of Empirical Thought, 190
 pre-established harmony, theory of:
 Baumgarten's defense of, 51,
 73–78, 97–98; Crusius's arguments
 against, 82, 87–89, 92, 92n133, 93,
 158, 174; Kant's arguments
 against, 102, 113–129; Leibniz's
 doctrine of, 2, 8, 23, 24–37, 75,
 94–95, 96, 161, 163; Meier's
 defense of, 9, 24, 51, 74, 77n107,
 78–80, 98, and politics, 44, 45,
 101; Wolff's limited acceptance of,
 9, 24–45, 47, 48–67, 81, 95–96
 principle of continuity, 7n3, 217, 231,
 252, 255, 255n28, 257
 principle of contradiction, 41, 49, 82,
 84, 85, 95, 98, 100, 102, 113n10,
 119, 119n15, 161, 165, 166, 167,
 169, 170–370, 372, 373, 382, 403
 principle of coexistence, 141–156,
 166, 176, 180, 220, 294
 principle of determining reason, 85,
 98, 115–116
 principle of identity of
 indiscernibles, 33, 49, 113n10
 principle of succession: arguments
 for, 113–125, 129, 132, 139, 180,
 225, 229; and pre-established
 harmony, 102, 113–129, 135,
 143–144, 158, 180; statement of,
 112–113, 131, 137, 225, 226, 294
 principle of sufficient reason, 46, 85,
 119, 199, 314, 316
Prolegomena, 170, 230, 352, 381, 384

- psychology, 41, 45, 51, 81, 95, 146;
 empirical, 41, 45; rational, 41, 46,
 48, 91, 95, 97, 278, 279, 283
- Putnam, Hilary, 319n16
- quantity, category of, 192
- Rationalism, 2, 44, 183, 373
- rational psychology, *see* psychology
- real vs. ideal influence, 75–78, 98,
 149, 292
- realism, 260n32, 319, 319n16,
 320–321. *See also* Transcendental
 Realism
- reason: causality of, 328, 332,
 334n34, 355–356; faculty of, 314,
 314n8, 323, 342, 355, 373, 378;
 practical, 182, 282–283, 295,
 336n37, 376; skeptical,
 dogmatic, and critical uses of,
 375–381
- Refutation of Idealism, 182, 206,
 282–283
- Reid, Thomas, 243n18, 417
- Reinbeck, Johann Gustav, 73
- Reusch, Johann Peter, 45, 73
- Rutherford, Donald, 32n25, 54n66,
 148n53
- Schematism, 181, 183, 266, 267,
 267n44, 281
- schemata, 190n8, 285, 324, 327
- Schopenhauer, Arthur, 212,
 238
- Schulz, Friedrich Albert, 52
- Second Analogy of Experience:
 argument of, 4, 5, 11–12, 203–217,
 223; claim of, 8, 188, 203, 221, 229;
 “Weak” vs. “Strong reading,”
 203–204, 215–216, 286–287,
 287n63; context of, 185–195;
 model of causality presupposed by,
 13–14, 284–286; *see also* causal
 powers; Third Analogy of
 Experience; role of necessity in,
 211–213; traditional view of Kant’s
 intentions in, 4, 5, 6–7, 13
- second law of motion, 272
- self: bundle theory of, 277;
 constitution of, 358–359;
 empirical, 334, 336; noumenal,
 330, 333, 334, 337; as represented
 in inner sense, 14, 45, 273, 274,
 276; as synthetic activity, 280
- self-knowledge, 277–278, 280;
 Hume’s denial of, 274, 277,
 277n53, 278; paradox of,
 273–274, 276
- self-consciousness, 196, 232, 266,
 272–281; distinguished from inner
 sense, 14, 273–274; as involving
 consciousness of synthesis, 14,
 273–282; as providing meaning of
 “activity,” 14, 272–282, 297;
 transcendental, 276. *See also*
 apperception
- Shoemaker, Sydney, 280n60, 395
- skepticism, of Hume, 7, 103, 181, 216,
 367, 374–378, 379–380
- souls, 34, 37, 93n140, 127; Crusius’s
 account of, 87, 89–93, 99;
 immortality of, 4, 34; Kant’s
 account of, 107–108, 283; Wolff’s
 account of, 48, 91
- space, 201, 270, 284, 295; Crusius’s
 account of, 82, 85; filling a space,
 109–111, 232, 268, 289; as
 generated by causal interaction,
 135, 176n74, 179; infinite
 divisibility of, 102, 109–111, 179,
 234; Leibniz’s account of, 49, 97,
 320; Newton’s picture of, 189, 320;
 and monads, 54n66, 58–59,
 109–112, 179, 325; as subjective
 principle of the sensible world, 12,
 172, 181, 191, 312, 338, 420; unity
 of, 176, 176n74
- Spinoza, Benedict, 2, 2n2, 367
- spontaneity: in causality, 308–310,
 316, 333; Kant’s account of, 306,
 308–310, 316, 330, 333, 340, 341,
 342, 348, 425; Leibniz’s account
 of, 30, 51n57, 94
- Stiebritz, Johann Friedrich, 80
- Strawson, Peter, 196, 196n21, 198,
 203, 204, 205, 206

- substance: activity as criterion of, 351;
 Baumgarten's account of, 13,
 110–172, 259; identity conditions
 of, 94; Knutzen's account of, 66;
 Leibniz's account of, 32–33, 94,
 351; permanence of, 221, 288;
 phenomenal, 191n10, 200, 222,
 229, 292, 293, 295, 303, 344,
 349–354 (*see also appearances*); not
 a bare particular, 222; noumenal,
 229, 293, 295, 296, 344, 349–354;
see also things in themselves; as
 self-sufficient, 94, 143, 147, 162,
 259; Spinoza's account of, 2;
 Wolff's account of, 110, 351. *See
 also* accidents; monads
- Sulzer, Johann Georg, 16, 363–368,
 372, 383
- synthetic a priori judgments, 186,
 187, 383
- synthetic a priori knowledge, 183,
 229, 312, 379–381, 384, 388
- synthetic a priori truths, 229,
 384
- systematic metaphysics, 19, 419–421
- Taylor, Richard, 19, 243n18, 409,
 412, 418
- Tetens, Johann Nicolas, 14, 273–275,
 368, 371; on causality, 276,
 369–373, 374, 383, 386; reaction
 to Hume, 16, 273–275, 363,
 368–373
- things in themselves, 128n30, 200,
 201, 292; as causes of appearances,
 302, 326–329, 334–339, 355–356;
 complete determinacy of, 222,
 237, 245–250, 305, 313, 314;
 knowledge of, 181, 183, 295, 323,
 326. *See also* causation; substance;
 Transcendental Idealism;
 Transcendental Realism
- Third Analogy of Experience:
 argument of, 12, 191, 217–229,
 237, 245–250; claim of, 8, 185, 201,
 217, 227, 231, 243, 293; context of,
 185–195; model of causality
 presupposed by, 12, 231, 237,
 238–243, 284, 296; why restricted
 to spatial substances, 226–227,
 228, 282–284, 293, 297
- Third Antinomy, 6, 8, 302, 304–316,
 328n26, 330, 335n35, 343, 347,
 355, 425; antithesis argument, 305,
 308–316; resolution, 311, 332, 341,
 355; thesis argument, 305–308,
 310
- Thoughts on True Estimation, see True
 Estimation*
- Thümmig, Ludwig Philip, 44, 45, 51,
 273, 273n49, 276
- time: subjective vs. objective, 189;
 unity of, 186, 191, 194n16, 195,
 202, 229, 286
- time determination: naïve picture of,
 188; problem of, 11, 185, 188–191,
 200, 210, 215, 216, 219, 220, 221,
 222, 249, 250, 263, 263n39;
 subjective vs. objective time, 189
- Tooley, Michael, 401
- Transcendental Aesthetic, 82, 181,
 190
- Transcendental Analytic, 205, 301
- Transcendental Deduction, 141n40,
 181, 182, 185, 186–188, 198, 206,
 209, 273, 274, 276–277, 340, 374
- Transcendental Dialectic, 41n41,
 170, 181, 183, 301, 374
- transcendental freedom, *see* freedom
- Transcendental Idealism, 170, 190,
 250, 295, 302, 310n6, 312n7, 373,
 378; different interpretations of,
 191n10, 302, 317–325;
 epistemological interpretation,
 191n10, 265n43, 317–323, 328;
 ontological interpretation,
 191n10, 317, 323–325, 329; as key
 to the resolution of the
 antinomies, 305, 311–313
- Transcendental Dialectic, 41n41,
 170, 181, 183, 301, 374
- Transcendental Realism, 15, 302,
 305, 307–308, 310, 310n6, 311,
 315, 348

- True Estimation*, 102, 104, 107n5, 109, 112, 147n51, 166, 174n73, 179, 254, 271n46, 384
- unconditioned, 128, 128n30, 314n8
- understanding: Crusius's account of, 82, 83, 87, 261; divine, 88, 98, 141, 150, 176, 294; Kant's account of, 170, 172, 181, 186, 198, 261, 274, 373, 378; Tetens's account of, 370–371
- unity: of apperception, 195, 202, 229; of experience, 193–195, 202, 229, 309; relation between unities, 194, 195, 229; of time, 186, 191, 194n16, 195, 202, 229, 286; of the world, 98, 173–175, 193, 194n16, 195, 199, 229
- Universal Natural History and Theory of the Heavens*, 101, 136n35
- universals, 401–405
- Verificationism, 198
- vis motrix, *see* force
- Van Inwagen, 336–337, 337n43, 338, 344
- Van Fraassen, Bas, 403
- Walch, Johann Georg, 51, 53
- Wolf, Susan, 338n45
- Wolff, Christian, 7, 43n43, 53, 98, 123–125; *Anmerkungen zur Deutschen Metaphysik*, 47–48; current reputation of, 38; eighteenth-century influence of, 40–41, 44, 95; expulsion from Prussia, 44, 95; on force, 121, 179; on grounds, 102, 103, 113n10, 118–121, 124, 133, 171, 180; and idealism, 9, 47–48, 95–96, 99, 159; and Leibniz, 9, 38, 38n35, 49, 49n55, 50, 125–126, 129; methodology of, 40–41, 41n40, 81, 95; on mind-body problem, 46n48; pedagogical achievements of, 40–41; and Pietists, 42–44, 50, 51n57, 52; on pre-established harmony, 9, 24–45, 47, 48–67, 81, 95–96; on principle of contradiction, 102, 372; *Rational Thoughts on God, the World, and the Soul of Human Beings, and All Things in General*, 24–45, 47, 48–67; on substance, 110, 351
- Wood, Allen, 324n22, 333, 333n32, 334, 334n33, 336n39, 350n51
- world, 11, 19, 195n18, 304; Crusius's account of, 83, 85–89, 90n128, 98, 146–147, 149, 153; perfection of, 9, 46, 74–78, 97; as a real whole, 12, 85–89, 248n23, 291; Wolff's account of, 85