

Kant's Legacy
edited by Predrag Cicovacki

KANT'S LEGACY

Rochester Studies in Philosophy

ISSN: 1529–188X

Senior Editor: Wade L. Robison

Rochester Institute of Technology

1. *The Scottish Enlightenment: Essays in Reinterpretation*

Edited by Paul Wood

2. *Kant's Legacy: Essays in Honor of Lewis White Beck*

Edited by Predrag Cicovacki

**Kant's Legacy:
Essays in Honor of Lewis White Beck**

**Edited by
Predrag Cicovacki**



University of Rochester Press

Copyright © 2000 Contributors

All Rights Reserved. Except as permitted under current legislation, no part of this work may be photocopied, stored in a retrieval system, published, performed in public, adapted, broadcast, transmitted, recorded or reproduced in any form or by any means, without the prior permission of the copyright owner.

First published 2000

The University of Rochester Press is an imprint of Boydell & Brewer, Inc.
668 Mt. Hope Avenue, Rochester, NY 14620, USA
and of Boydell & Brewer, Ltd.
P.O. Box 9, Woodbridge, Suffolk IP12 3DF, UK

ISBN 1-58046-053-4

ISSN: 1529-188X

Library of Congress Cataloging-in-Publication Data

Kant's legacy : essays in honor of Lewis White Beck / edited by
Predrag Cicovacki.

p. cm. — (Rochester studies in philosophy, ISSN 1529-188X ; 2)
Includes bibliographical references and index.

ISBN 1-58046-053-4

1. Kant, Immanuel, 1724-1804. I. Beck, Lewis White.
II. Cicovacki, Predrag. III. Series.

B2798.K226 2000

193—dc21

00-064926

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Typeset by Isis-1 Corporation

Printed in the United States of America

This publication is printed on acid-free paper

For my sister

***“The only way to speak the truth
is to speak lovingly.”***

—Thoreau

Contents

List of Contributors	xi
Method of Citation	xv
Predrag Cicovacki: Introduction: Man as a Creator	xvii
Part One: Kant's Copernican Revolution	1
Stanley Rosen: Is Thinking Spontaneous?	3
Graham H. Bird: Lewis White Beck's Account of Kant's Strategy	25
Predrag Cicovacki: Paths Traced through Reality: Kant on Commonsense Truths	47
Gordon G. Brittan Jr.: The Anti-Reductionist Kant	71
Carsten Held: Analyticity and the Semantics of Predicates	93
Robert Howell: Kant, the 'I Think', and Self-Awareness	117
Gerold Prauss: The Problem of Time in Kant	153
Karl Ameriks: Kant and Short Arguments to Humility	167
Part Two: Kant's Rousseauistic Revolution	195
Ralf Meerbote: Which Freedom?	197
Robert L. Holmes: Consequentialism and Its Consequences	227
Rüdiger Bubner: Another Look at Maxims	245
Allen W. Wood: Kant versus Eudaimonism	261
Yirmiyahu Yovel: Kant and the History of the Will	283
Joseph P. Lawrence: Moral Mysticism in Kant's Religion of Practical Reason	311
Susan Meld Shell: Kant as Educator: Reason and Religion in Part One of the <i>Conflict of the Faculties</i>	333
Henry E. Allison: The <i>Quid Facti</i> and <i>Quid Juris</i> in Kant's Critique of Taste	369

John H. Zammito: Kant in the 1760s: Contextualizing the “Popular” Turn	387
Index	433

Contributors

Henry E. Allison is Professor of Philosophy at Boston University. He is a member of the Norwegian Academy of Arts and Sciences and was President of the Pacific Division of the American Philosophical Association (1996–97). His books include *Lessing and the Enlightenment* (1966), *Benedict de Spinoza* (1975), *Kant's Transcendental Idealism: An Interpretation and Defense* (1983; Spanish translation 1992), *Kant's Theory of Freedom* (1990), and *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (1996). Allison is the translator and editor of *The Kant-Eberhard Controversy* (1973).

Karl Ameriks is Professor of Philosophy at the University of Notre Dame. He is the author of *Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason* (1982; 2nd rev. ed. forthcoming), and *Kant and the Fate of Autonomy* (forthcoming). He is coeditor of *The Modern Subject* (1995), the editor of the *Cambridge Companion to German Idealism* (forthcoming), and cotranslator of Edmund Husserl's *Experience and Judgment* (1973), and of Kant's *Lectures on Metaphysics* (1997).

Graham H. Bird has recently taught at Manchester University (England). He is the author of *Kant's Theory of Knowledge* (1962) and *William James* (1986), and the editor of *William James: Selected Writings* (1995). Bird is coeditor of the recently founded *Kantian Review*, the first philosophy journal in English devoted to Kant's philosophy.

Gordon G. Brittan Jr. is Professor of Philosophy at Montana State University. He is the author (with Karel Lambert) of *An Introduction to the Philosophy of Science* (1970), of which six translations have been published, and *Kant's Theory of Science* (1978), and the editor of *Causality, Method, and Modality: Essays in Honor of Jules Vuillemin* (1991). Brittan is a member of the editorial board of *Kant-Studien*.

Rüdiger Bubner is Professor of Philosophy at the University of Heidelberg, and also a member of the Heidelberg Academy of Sciences. His areas of research cover ancient philosophy, German idealism, hermeneutics, social philosophy, and aesthetics. Bubner's books include *Handlung, Sprache und Vernunft* (1976; 2nd ed. 1982), *Modern German Philosophy* (1981), *Geschichtsprozesse und Handlungsnormen* (1984), *Essays in Hermeneutics and Critical Theory* (1988), *Ästhetische Erfahrung* (1989), *Antike Themen und ihre moderne Verwandlungen* (1992), *Welche Rationalität bekommt der Gesellschaft?* (1996), *Innovationen des Idealismus* (1995), and *Drei Studien zur politischen Philosophie* (1999).

Predrag Cicovacki is Associate Professor of Philosophy at the College of the Holy Cross. His interests cover Kant, epistemology, early modern philosophy, and philosophy of psychoanalysis. He is the author of *Anamorphosis: Kant on Knowledge and Ignorance* (1997) and the editor of *Essays by Lewis White Beck: Fifty Years as a Philosopher* (1998). Cicovacki has published numerous articles on Kant and the problem of truth, and is currently writing a book on Kant's conceptions of truth and illusion.

Carsten Held teaches philosophy at the University of Freiburg. His research interest is divided between philosophy of science (especially philosophy of quantum mechanics) and philosophy of language. He is the author of *Die Bohr-Einstein-Debatte: Quantenmechanik und physikalische Wirklichkeit* (1998), and is currently preparing a study of Frege's semantics.

Robert L. Holmes is Professor of Philosophy at the University of Rochester. His research is primarily focused on philosophy of war; the ethics of nonviolence; contemporary ethical theory; and normative problems in social, political, and international affairs. Holmes is the author of *On War and Morality* (1989), *Basic Moral Philosophy* (1993; 2nd ed. 1998), and the editor of *Nonviolence in Theory and Practice* (1990). He coauthored the second edition of *Philosophical Inquiry* with Lewis White Beck (1968), a colleague for 36 years.

Robert Howell is Professor of Philosophy at the State University of New York at Albany. He is the author of *Kant's Transcendental Deduction: An Analysis of the Main Themes in His Critical Philosophy* (1992), a study of the Transcendental Deduction of the Categories, and of a number of

papers on Kant's theoretical philosophy. He has also published essays on pictorial representation, the theory and ontology of fiction, and other issues in aesthetics.

Joseph P. Lawrence is Associate Professor of Philosophy at the College of the Holy Cross. He is the author of *Schellings Philosophie des ewigen Anfang* (1989), and has recently completed a book manuscript "The Silence of Socrates." Lawrence has published articles on Greek philosophy, philosophy from Kant to Heidegger, eastern philosophy, and philosophy of religion.

Ralf Meerbote is Professor of Philosophy at the University of Rochester, where he succeeded Lewis White Beck in 1980. His research is focused on Kant's philosophy, seventeenth- and eighteenth-century philosophy, aesthetics, and epistemology. Besides publishing numerous articles on Kant in leading philosophical journals, Meerbote is the editor of *Kant on Causality, Freedom and Objectivity* (with William L. Harper, 1984) and of *Kant's Aesthetics* (1992). His works in progress involve two book manuscripts: "The Systematicity of Nature: Kant's Antinomy of Teleological Judgment and Its Resolution," and "Kant's Theory of Human Knowledge and Its Object."

Gerold Prauss is Professor of Philosophy at the University of Freiburg. His numerous books include *Platon und der logische Eleatismus* (1966), *Erscheinung bei Kant* (1971), *Kant und das Problem der Dinge an sich* (1974), *Erkennen und Handeln in Heideggers "Sein und Zeit"* (1977), *Einführung in die Erkenntnistheorie* (1980), *Kant über Freiheit als Autonomie* (1983), and *Die Welt und wir* (Band I/1 1990, Band I/2 1993, Band II/1 1999). Prauss is the editor of *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln* (1973) and of *Handlungstheorie und Transzendentalphilosophie* (1986).

Stanley Rosen is Professor of Philosophy at Boston University. He has lectured and held visiting professorships in many of the leading universities of North America and Europe. Rosen is the author of thirteen books, some of which include *Plato's Symposium* (1969; 2nd ed. 1987), *Nihilism: A Philosophical Essay* (1969), *G.W.F. Hegel: An Introduction to the Science of Wisdom* (1974), *Plato's Sophist* (1983), *The Limits of Analysis* (1985), *Hermeneutics as Politics* (1987), *The Ancient and the Modern*:

Rethinking Modernity (1989), *The Question of Being: A Reversal of Heidegger* (1993), *Plato's Statesman* (1995), and *The Mask of Enlightenment: Nietzsche's Zarathustra* (1995). Some of his books have been translated into French, Italian, Catalan and Japanese. In 1997 Rosen was awarded an honorary doctorate by the New University of Lisbon.

Susan Meld Shell is Professor of Political Science at Boston College. Her publications include *The Right of Reason: A Study of Kant's Philosophy and Politics* (1980) and *The Embodiment of Reason: Kant on Spirit, Generation, and Community* (1996). She is also the author of essays on Rousseau, German Idealism, and theories of liberalism.

Allen W. Wood is Professor of Philosophy at Stanford University. He is the general editor (together with P. Guyer) of the monumental fifteen-volume edition of Kant's works, currently published by Cambridge University Press. He is the author of several books: *Kant's Moral Religion* (1970), *Kant's Rational Theology* (1978), *Karl Marx* (1981), *Hegel's Ethical Thought* (1990), and *Kant's Ethical Thought* (1999). Wood has also edited and translated several books, two of which include Kant's *Lectures on Philosophical Theology* (with Gertrude S. Clarke, 1978) and Kant's *Critique of Pure Reason* (with P. Guyer, 1998).

Yirmiyahu Yovel divides his teaching duties between the Hebrew University in Jerusalem and the New School University in New York. He is the author of *Kant and the Philosophy of History* (1980; 2nd ed. 1989), *Spinoza and other Heretics* (two volumes, 1989), and *Dark Riddle: Hegel, Nietzsche, and the Jews* (1998). Yovel is the editor of *Nietzsche as Affirmative Thinker* (1986), *Maimonides and Philosophy* (1987), *Kant's Practical Philosophy Reconsidered* (1989), *God and Nature in Spinoza's Metaphysics* (1991), and coeditor of *Spinoza on Knowledge and the Human Mind* (1994).

John H. Zammito is Associate Professor of History at Rice University. He is the author of *The Genesis of Kant's "Critique of Judgment"* (1992). Zammito is currently working on the relationship between Kant and Herder.

Method of Citation

In the following essays, citations to Kant's texts are usually given parenthetically, and two forms of citations are used. Citations to the *Critique of Pure Reason* are given in the customary manner by reference to the pagination of the first ("A") and second ("B") editions. Where both A and B page numbers are provided, the passage cited is included in both editions; otherwise the passage occurs only in the cited edition. Passages from all other works by Kant are located by volume and page number, given in Arabic numerals separated by a colon, in the standard edition of Kant's works, *Kants gesammelte Schriften*, edited by the *Königlichen Preußischen* (later *Deutschen*) *Akademie der Wissenschaften* (Berlin: Georg Reimer [later Walter de Gruyter], 1900–). These references are preceded by a full or an abbreviated title for the work in question, unless the context obviates the need for that. Several contributors have followed the *Akademie* edition citation with a citation of an English translation of the work; when that is the case, their notes provide information about the translations used.

The following lists of the abbreviated titles of Kant's works is used in this collection, with the original titles and the references to the *Akademie* edition of Kant's works in parenthesis:

<i>Anthropology</i>	<i>Anthropology from a Pragmatic Point of View</i> (<i>Anthropologie in pragmatischer Hinsicht</i> ; 7:117–333)
<i>Conjectures</i>	<i>Conjectures on the Beginning of Human History</i> (<i>Mutmaßlicher Anfang der Menschengeschichte</i> ; 8:107–23)
<i>Conflict</i>	<i>Conflict of the Faculties</i> (<i>Der Streit der Fakultäten</i> ; 7:1–116)
<i>Groundwork</i>	<i>Groundwork of the Metaphysics of Morals</i> (<i>Grundlegung zur Metaphysik der Sitten</i> ; 4:385–463)

<i>Idea</i>	<i>Idea for a Universal History from a Cosmopolitan Point of View (Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht; 7:15–31)</i>
<i>Judgment</i>	<i>Critique of Judgment (Kritik der Urteilkraft; 5:165–485)</i>
<i>Metaphysical Foundations</i>	<i>Metaphysical Foundations of Natural Science (Metaphysische Anfangsgründe der Naturwissenschaft; 4:465–565)</i>
<i>Morals</i>	<i>Metaphysics of Morals (Der Metaphysik der Sitten; 6:203–493)</i>
<i>Nova dilucidatio</i>	<i>A New Exposition of the First Principles of Metaphysical Cognition (Principiorum primorum cognitionis metaphysicae nova dilucidatio; 1:385–416)</i>
<i>Practical Reason</i>	<i>Critique of Practical Reason (Kritik der praktischen Vernunft; 5:1–163)</i>
<i>Progress</i>	<i>What Real Progress Has Metaphysics Made in Germany since the Time of Leibniz and Wolff? (Über die Fortschritte der Metaphysik seit Leibniz und Wolff; 20:253–351)</i>
<i>Prolegomena</i>	<i>Prolegomena to Any Future Metaphysics Which Will Be Able to Come Forth as a Science (Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können; 4:253–383)</i>
<i>Religion</i>	<i>Religion within the Limits of Reason Alone (Religion innerhalb der Grenzen der bloßen Vernunft; 6:1–202)</i>

Introduction: Man as a Creator

Predrag Cicovacki

I

There is no better account of the rise of German philosophy, from its humble beginnings to its magnificent peaks at the end of the eighteenth century, than that written by Lewis White Beck. His *Early German Philosophy: Kant and His Predecessors*¹ is a superb synthesis of systematic scholarship, deep insight, and clear presentation. Beck initially planned to extend his account to the entire history of German philosophy. While preparing the second volume, he realized how powerfully attracted he was to Kant's legacy and abandoned the completion of the project to immerse himself even more deeply in the study of Kant.

What attracted Beck so powerfully to Kant? What did he see as the most significant element of Kant's rich, complex, and controversial legacy? Perhaps more than anything else, it was the concern with what Kant considered the ultimate philosophical question: What is man? While the significance of this question is all too clear, Kant never framed his response in a direct and unambiguous manner. It is far clearer that Beck takes Kant's most important answer to this question to be that Man is a creator.² This answer has a double edge. On the one hand, it dignifies man and elevates him high above the blind mechanical forces of nature. Man is not a mere spectator, nor a puppet on the stage set by a god. Man is a legislator and actor; he is a free designer of his own life and the world-scene. On the other hand, the role of a creator imposes a difficult burden on the shoulders of man, the task of providing a unitary world-view and an immanently grounded system of values and norms. Just two centuries after Kant, that burden may appear too heavy. Our "post-modern" age has witnessed the collapse of confidence in human thought in general and the fragmentation—even disappearance—of a unitary world-view with established values and norms.

However controversial this aspect of Kant's philosophy may be, Beck found his conception of man as a creator most important and fascinating,

whether he was writing on Kant's theoretical or practical philosophy. Most of the contributors to this volume, no doubt under Beck's influence, also consider this theme to be of central significance for the proper understanding and evaluation of Kant's legacy.

II

The first part of this collection comprises eight papers which deal with the issues of subjectivity and objectivity in Kant's theoretical philosophy. The authors of these papers would agree with Beck's claim that the central thought of Kant's *Critique of Pure Reason* is the creative activity of the mind.³ The human mind imposes upon the phenomena the spatial, temporal, and categorial relations which not only make it possible for the cognitive subject to know something of nature a priori, but make nature itself (with respect to its formal structure, i.e., its laws) possible. The essence of Kant's so-called "Copernican revolution" is that "man is the creator of nature as phenomena under the law."⁴

Even if we grant to Kant that this view leads to a satisfactory answer of the question concerning the possibility of objectively valid synthetic a priori judgments, his view invites pressing questions concerning the objective validity of empirical judgments. Some of these misgivings are masterfully articulated in Stanley Rosen's article, "Is Thinking Spontaneous?" which opens this collection. Rosen asks: Given the Copernican turn, i.e., given the constructive and spontaneous role of thinking, how is it possible to preserve the integrity of the distinction between the process of thinking and the matter of thought? Put differently, given Kant's premises, how is the objectivity of empirical knowledge of the world possible?

While Rosen does not believe that the Kantian framework allows enough room for a satisfactory answer,⁵ papers by Graham Bird, Predrag Cicovacki, and Gordon Brittan Jr., point in the opposite direction. According to Bird, a proper answer to the challenge of accounting for the possibility of objective knowledge will depend on what we make of Kant's famous statements that "without sensibility no object would be given to us, without understanding no object would be thought. Thoughts without content are empty, intuitions without concepts are blind" (A51–2/B75–6). Bird furthermore emphasizes Kant's attempt to avoid and obliterate the traditional metaphysical and epistemological dualisms between 'sub-

ject' and 'object', 'empirical' and 'transcendent', and 'realism' and 'idealism' as too narrow and inappropriate.

Like Bird, Cicovacki also insists on the interdependence of understanding and sensibility, but he develops this theme in the direction of an interactive conception of truth; one where both subject and object provide important contributions to, and constraints on, the possibility of having objectively valid judgments.

Brittan makes a further contribution to our recognition of the limitations of the subject's creative role in cognition. While he would gladly subscribe to Kant's claim that "thought is the act of referring given intuitions to an object" (A247/B304), he insists that there is a fundamental distinction between 'objects' and 'constructs'; Kant's account of the possibility of objective cognition of the external world ultimately refers to the former and not the latter.

Kant's Copernican revolution does not only lead to questions regarding the objectivity of our knowledge of the external world. Recognition of the subject's creative role in cognition also raises new issues of intersubjectivity, and—no less importantly—those of the subject's understanding of himself as a cognitive being. The Kantian standpoint has to account not only for the possibility of (a priori and empirical) synthetic judgments, but also for the possibility of analytic judgments. Their validity depends, partially or fully, on the shared meanings of the words used; but—given the subject's active role in the formation of concepts—we must explain how we come to form and share concepts with the same content and meanings? This issue is tackled by Carsten Held, who interprets concepts as the "meanings of predicates," and argues that the conception of analyticity based on this view makes a fruitful contribution to philosophical semantics.

The difficult problems pertaining to the thinker's knowing and understanding himself as a cognitive subject are addressed by Robert Howell, Gerold Prauss, and Karl Ameriks. Howell's paper is focused on the issues regarding self-awareness and Kant's famous "I think," which "must be capable of accompanying all my representations" (B131). "I think" does not seem to fit into the framework of Kant's transcendental idealism (e.g., it is neither a concept nor an intuition), and Howell argues that it should be understood via the contemporary idea of "direct reference."

Prauss is more directly interested in the problem of time and its difference from space; but after proposing an ingenious dynamic model for understanding time, Prauss intends to indicate its relevance for under-

standing the subject of cognition. The subject is an entity and yet not a *res*, as Descartes assumed; this entity is not a static substance but—like time—a dynamic carrier of constant motion.

Ameriks is similarly concerned about the role of the traditional view concerning substances in Kant's philosophy, not only with respect to the subject but also in connection with the notoriously problematic things in themselves. While it is not clear how to give a positive account of how things in themselves might affect us and what they are (e.g., whether they have intrinsic, noncontingent, and nonrelational properties), Ameriks attempts to show that they cannot be simply dismissed either. They are crucially important for Kant's account of freedom and, more generally, for his insistence of the primacy of practical over theoretical reason.

This brings us to the second part of the book and to papers dealing mostly with Kant's practical turn, where issues related to man's creativity may carry even more weight than they do in the theoretical realm. This is certainly what Beck believed,⁶ and he always stressed that besides the Copernican revolution, where the cognitive subject imposes the form of lawfulness on events in nature, an analogous revolution occurs in Kant's moral philosophy. Beck called this revolution the "Rousseauistic revolution," because it was based on Kant's admiration for Rousseau's view that "Freedom is obedience to a law which man gives to himself." Kant applied this view not only to political philosophy, as Rousseau did, but primarily to ethics, and to a significant extent to the philosophy of religion and aesthetics.

The Rousseauistic revolution leads to as many questions as its theoretical counterpart. As Ralf Meerbote shows in his paper, "Which Freedom?," one of the most basic of these questions deals with the proper understanding of freedom, in connection with the natural causality (as defined in the Second Analogy of the *Critique of Pure Reason*). One interpretation, championed by Beck, demotes the principle of natural causality to the status of a merely regulative principle, and holds that there must be uncaused events for there to be free actions. Another interpretation, defended by Allen Wood, opts for nontemporal agency and must defend Kant's problematic noumenal realm. Henry Allison affirms natural causality and freedom existing side by side. Meerbote himself defends a conception of freedom based on Davidson's understanding of "anomalous monism." There are apparently many ways to interpret Kant's understanding of freedom; each conception has its advantages, but each has also a price to pay. Which, then, is the right one?

Although the problem of compatibilism of freedom with natural cau-

salinity remains a bone of permanent contention among Kant scholars, papers by Robert Holmes, Rüdiger Bubner, Allen Wood, and Yirmiyahu Yovel turn our attention to other issues of importance for the proper understanding and evaluation of Kant's moral philosophy. While Holmes' central goal is to show that consequentialism (one of Kant's main competitors for the moral throne) fails as an ethical theory, his paper also offers a terminology that may provide a way around Kant's compatibilistic maze. He distinguishes 'effects', which depend upon a causal connection between an action and its outcome, from 'consequences' and 'results'. Consequences include effects, but there are some candidates for consequences to which the notion of effects is inappropriate (e.g., outcomes that involve the actions of others). Holmes understands results even more broadly than consequences.

A broader perspective (than that of compatibilism) on Kant's moral philosophy is also taken by Bubner. His paper discusses the morality of actions in connection not with effects, consequences, or results, but with maxims. Moreover, it reminds us of their (frequently forgotten) role in Kant's moral philosophy. Bubner's paper relates Kant's view to the French moralists, from whom he inherited the idea of maxims as principles that guide our actions, and also to Wittgenstein and contemporary philosophers concerned with the nature and validity of rules in language.

Wood's paper, by contrast, turns our attention back toward Socrates and other Greek moralists; he attempts to demonstrate the advantages of Kant's moral approach over the traditional views on happiness [*eudaimonia*]. His view is that happiness is a self-contradictory tendency in human beings; while it is based on the pursuit of the satisfaction of desires, the very nature of desires prevents this pursuit from being realized and realizable. This may be why Kant shifts the emphasis of his moral theory from the idea of happiness to that of the highest good.

This topic of the *summum bonum* is pursued further by Yovel. He masterfully situates Kant's moral theory in the context of the socioeconomic, religious, and philosophical developments of that time, and gives an extensive analysis of Kant's theory of the self-determining will. Yovel furthermore displays the metamorphosis of Kant's moral theory, which ends with the view of the highest good as a regulative idea of history achieved by the human will.

Yovel's discussion opens a door for the consideration of the role of religion in Kant's practical philosophy, and this difficult and important topic is discussed by Joseph Lawrence and Susan Meld Shell. Lawrence

deals with the aging Kant's desire to interpret Christianity in the light of his critical philosophy. Despite Kant's insistence on restraining religion "within the limits of reason alone" (which apparently clashes with Christianity's reliance upon historical revelation, especially through the unique person of Jesus Christ), Lawrence argues that a moral mysticism lies at the core of Kant's philosophy. The emptiness of the moral law leaves ample room for purity of heart; this state can lead to a mystical fusion with the divine essence.

Shell's paper looks at the clash of theological and philosophical faculties at the German universities of Kant's time, the broader conflict of religion and philosophy, and an even broader question of the possibility of moral education within the confines of Kant's philosophy. Must the possibility of moral improvement lead Kant to assume a kind of noumenal or intelligible conversion, outside of space and time, and involving something like a "new creation" and a "Christian rebirth"? While this may be one part of the answer, Kant saw its other part to consist in the proper organization of universities which, functioning simultaneously as centers of free intellectual activity and indirect alliances between the powers of reason (philosophy) and state power, could lead the way toward a universal transformation and moral improvement.

Henry Allison's paper takes us into different waters, those of Kant's deduction of pure judgments of taste. Although a claim of taste involves a demand for universal agreement that is analogous to the demand built into a cognitive judgment, the subject's claim is based on a feeling. Allison's thesis is that Kant's frequently depreciated deduction of taste achieves the goal it sets to establish, and that is the claim to normativity of such judgments. What remains undetermined (and perhaps undeterminable) is whether the conditions of a pure judgment of taste have been met in a particular instance. Allison argues that this conclusion does not go against Kant's deduction; on the contrary, it reveals what must be the case, given the specific nature of the judgments of taste.

Finally, Kant's well-known critical achievements may be contrasted to his early intentions and ambitions through John Zammito's paper, "Kant in the 1760s: Contextualizing the 'Popular' Turn." This paper is inspired by Beck's way of historically contextualizing Kant in the *Early German Philosophy*. A careful historical and biographical analysis on Kant's life and thought in the decade when he was deeply impressed by Rousseau and popular philosophy paints a picture of the pre-critical Kant, a philosopher struggling to find his identity and secure his place in the German

intellectual world. While this paper aptly concludes the book, it certainly does not terminate our critical reflection on Kant's, and Beck's, legacy.

III

While the papers collected here address issues concerning creativity in all aspects of human experience—from knowledge of the external world to self-knowledge, from moral to religious dilemmas, from judgments of taste to the art of living—they also bring to our attention a theme of utmost significance for both Kant and Beck: the limitations of human creativity. Beck never tired of stressing that although man is a creator, “he is a creator within narrow limits, and a creator whose acts are incomprehensible to us.”⁷ Nor did Beck mind repeating that “the ‘world’ man creates is a perspective on a world he did not create . . . while Kant’s man exercises many functions previously assigned to God, man is not a god.”⁸

While this “reminder” may appear trivial, it is far from being so. In the form of science and technology, for instance, while the sweet fruit of creativity has opened our eyes to distant and unknown worlds, it has also redirected our sight away from what is close and important, from ourselves. Liberated human creativity has produced not only the tools that make our life longer and more convenient, but also the tools that may at any moment lead to a complete self-destruction of humanity.⁹

It may be that his awareness of these limitations, together with his recognition of their exceptional importance, turned Beck away from continuing his writing of the post-Kantian history of German philosophy. Beck came to believe that Kant’s followers among the German idealists failed to acknowledge that man’s creativity is exercised in a world that he himself did not make. In contrast to the German idealists, Beck was powerfully drawn equally to Kant’s image of man-creator and Kant’s image of “the land of truth.” For Kant, this “land of truth” was an island, “surrounded by a vast and stormy ocean, the native home of illusion, where many a fog-bank and many a swiftly melting iceberg give the deceptive appearance of farther shores, deluding the adventurous seafarer ever anew with empty hopes, and engaging him in enterprises which he can never abandon and yet is unable to carry to completion” (A235–6/B294–5).

Kant’s *Critique of Pure Reason*, together with the two other *Critiques* and his book *Religion within the Limits of Reason Alone*, was intended to show not the temporal and changeable limits of human creativity, but to

uncover and mark its permanent boundaries. In Kant's critical philosophy, insofar as it aimed at delineating the borders of the land of truth within which human creativity can (and should) develop and blossom, Beck recognized an accurate vision of man's position in the world. The "land of truth" is man's realm, the realm in which man is a creator, legislator, and actor. Yet the land that man orders, or seeks to order, "is only the known part of the unknown all."¹⁰ Beck held this to be one of the most important insights for the proper understanding of our own nature and role in reality. As he explained in one of his last public lectures, "We could hardly understand our history, or ourselves as products of history, without appreciating the nature and impact of this belief."¹¹

While the idea of man's creativity has become an indispensable aspect of our modern heritage, the idea of an unknowable beyond was born and cultivated in the ancient world. Beck associated it with Plato and thought that, whether or not we accept it, ever since Plato's time it has formed the *nervus probandi* of Western philosophy and civilization. Our general frame of orientation, our sense of identity, and our system of values are to a significant degree determined by our stand toward this Platonic idea of the world of ultimate but hidden reality.

Beck furthermore believed that only three principal attitudes are possible with respect to this idea. First, it is possible to deny that this world exists, and take it to be "a product of human imagination seeking solace which it cannot find in this world, which is the only real world and the only object of knowledge." Beck identified this view as the standpoint of Marx, Nietzsche, positivistic philosophers, and many scientists.

The second attitude is that of philosophers, such as Plato and Hegel, who believed that that other world not only exists, but can be known—either by mystical insight, or philosophical reasoning, or both.

The third view is that while this other world exists, we cannot know in any way what it is. In Beck's words:

We are indefeasibly ignorant of it. But that transcendent world is of such paramount importance to human beings that we cannot be indifferent to it as we are to most things of which we are ignorant; the attitude we should take to it is not one of claiming knowledge, but of professing faith. Because of our indefeasible ignorance of the most important things—God, the state of the soul, the destiny of the human race—we have need of an act of faith that is not just a poor substitute for an act of knowledge, but is an act which is laden with an existential significance lacking in knowledge.¹²

Beck believed that this view, in various forms, has been developed by philosophers as different as Aquinas, Pascal, Kierkegaard, and William James. Not surprisingly, Beck was convinced that this stand was also taken by Kant. What about Beck's own view? Those who knew Beck—that always wonderful and modest man and equally profound and clear thinker—could have anticipated his answer:

It is not my place to tell you whether there is indefeasible ignorance of ultimate reality. I am ignorant of whether there is or is not. But you should think of these things because there are no things more important, though there are no questions more difficult or less answerable. But one's whole life may be changed if one changes his mind about these questions.¹³

This is Kant's legacy, and it is also Beck's. They have left us a rich inheritance, one that only may be grasped through continuing their explorations along the exciting and dangerous trail of human creativity.¹⁴

Notes

1. Cambridge: Harvard University Press, 1969; reprint, Bristol: Thoemmes, 1996.

2. See, for instance, "What Have We Learned from Kant?" in *Self and Nature in Kant's Philosophy*, ed. A. Wood (Ithaca: Cornell University Press, 1984), 17–30. Reprinted in *Essays by Lewis White Beck: Five Decades as a Philosopher*, ed. P. Cicovacki (Rochester: University of Rochester Press, 1998), 199–211; cited 209. All the subsequent references to this paper will be to the latter edition. See also *Kant Selections*, ed. and intro. L.W. Beck (New York: Macmillan, 1988), Introduction, 20–2.

3. "What Have We Learned from Kant?" 202.

4. *Kant Selections*, 21.

5. One of Rosen's central worries deals with the relationship of spontaneity and necessity, which Kant assumed always go together, while Rosen sees a tension between them. According to Rosen, if our rules and laws are the results of spontaneous creative thinking, why would they be necessary and not replaceable by other spontaneous rules and laws? Beck voiced a similar criticism in the context of acting in particular and practical philosophy in general: "Kant says of a rational being that it is one that can act according to a conception of law. This may be accomplished by obedience to the law, by breaking the law, or by acting so as to exempt oneself from the application of law"; *The Actor and the Spectator* (New Haven: Yale University Press, 1975), 129.

6. For Beck's views on Kant's moral philosophy, see his incomparable *Commentary on Kant's "Critique of Practical Reason"* (Chicago: University of Chicago Press, 1960).

7. *Kant Selections*, 22.

8. *Ibid.*

9. Writing about the exploration of space in the search for extraterrestrial life, but obviously having a larger picture in mind, in his presidential address delivered before the Sixty-eighth Annual Eastern Meeting of the American Philosophical Association (New York City, December 12, 1971), Beck warned his colleagues: "We are now suffering from technological shock, destroying by radiological and chemical, if not moral, pollution the only abode of life we know. Are we not enough like our ancestors to respond with the same desperate hope they did? *Exobiology recapitulates eschatology*. The eschatological hope of help from heaven revives when the heavens of modern astronomy replace the Heaven of religion. That we can learn from more advanced societies in the skies the secret of survival is the eschatological hope which motivates, or at least is used to justify, the work of exobiologists. But somewhat like people who object to spending money needed in the ghettos on exploring the moon, I think the best hope for our survival is to be based on understanding human predicaments here on earth, not on expecting a saving message from super-human beings in the skies"; "Extraterrestrial Intelligent Life," in *Essays by Lewis White Beck*, 117.

10. "What Have We Learned from Kant?" 210.

11. The lecture in question was the Kearse Distinguished Lecture, entitled "Praise of Ignorance," delivered at the College of Liberal Arts, Rochester Institute of Technology, on May 2, 1983, on the occasion of Student Honor Awards Ceremony. This lecture was recently published in *Diotima* 1:2000, 3–9; citing 7.

12. *Ibid.*, 9.

13. *Ibid.*

14. I am grateful to Nalin Ranasinghe for his comments on an earlier version of this introduction.

Part One

Kant's Copernican Revolution

Is Thinking Spontaneous?

Stanley Rosen

I

The English word “spontaneous” is derived from the Latin noun *spons*, which occurs almost exclusively in the ablative, *sponte*; Varro glosses the term as follows: *id est voluntate*. To act *sponte sua* is thus to act freely or voluntarily; and this is the sense of the first use of “spontaneous” cited in the *Unabridged Oxford English Dictionary* from Hobbes’ “Essay on Liberty,” written in the seventeenth century. But we also refer to an act as spontaneous to indicate that it was not prepared by antecedent deliberation or a conscious exercise of the will. In the first sense, spontaneity is obviously connected to autonomy; in the second sense, it is not. Spontaneous acts of the second sort are similar to nonhuman occurrences for which no natural cause can be assigned, or which are due to chance.

The relation between natural and spontaneous occurrences is complicated. An event like a fire can occur spontaneously in the sense that no one intended or consciously caused it, yet it remains in accord with natural laws. Even more complicated is what we mean by the expression “a spontaneous person.” Such a person no doubt acts easily and without considerable advance deliberation, but this is not in itself a proof of free will. Spontaneous acts may arise on the basis of extremely rapid processes of calculation, but this in itself does not distinguish them from the operations of computers.

If there is a difference between human beings and calculating machines, it must lie in a difference as well as a connection between spontaneity and autonomy. It is not the mere manifestation of laws that is spontaneous, but the freedom with which they are given. And freedom is a characteristic of human beings when it is associated with consciousness and intentionality. Differently stated, I will call behavior “spontaneous” if there is no rule for the application of rules that govern part or all of the behavior in question. Whether or not this is Kant’s official definition of spontaneity, it is certainly Kantian in spirit. Kantian or not, however, it is

my contention that thinking is both spontaneous and autonomous. It is rule-governed but not algorithmic. I believe that this is a sound approach to the analysis of thinking, but I have some modifications to introduce into the Kantian formulation.

The main point can be stated as follows. Freedom for Kant means, of course, not arbitrary action in accord with the vagaries of the will, but rather action, whether cognitive or practical, in accord with rules. Nevertheless, the freedom of thinking is spontaneous because it is not derived from or regulated by an external source, whether this source be God or nature. This means inevitably that thought is free because it makes its own rules. Yet Kant argues that thought is bound by these rules in two senses. First, it cannot make rules other than those in which it necessarily actualizes when thinking the world of objects of experience. Second, our conception of the dignity of rationality binds us to obey the categorical imperative in the domain of morality.

In this paper, I will be concerned primarily with the first or theoretical sense of spontaneity. My intention is not to provide a philological or historical exegesis of Kant's doctrine of spontaneity, but to show how a fundamental philosophical problem is introduced by his teaching. Stated with introductory succinctness, the problem lies in the relation between spontaneity and necessity. As rules, the results of thinking are necessary; but if they are spontaneous, they can be replaced by other spontaneous rules or laws. In this case the rules are contingent, and there is no rule that prevents us from replacing one set of contingent rules with another.

By way of transition to the main section of this paper, I call your attention to an important distinction. The term "thinking" is itself ambiguous or leads directly to an ambiguity. It is easy to confuse thinking as a psychological process with the ontological status of what one thinks about. Under ordinary circumstances, no one would be likely to say that the perception of a tree is the same as the tree that we perceive. Not even Kant would go this far; but he does hold that the tree as an "object of experience" is constructed within the unity of consciousness from various acts of pure thinking applied to sensations. As I am about to argue, it is virtually impossible on Kantian premises to preserve the integrity of the distinction between the process of thinking and the ontological independence of what one thinks about.

We may also distinguish between a subjective representation, such as the perception of a tree or the psychological process by which we construct a concept of trees, and the objective meaning of that concept. Such

a distinction was in fact drawn by thinkers like Bolzano, Frege, and Husserl.¹ The question that was previously addressed to the difference between the perception or concept of the tree and the tree itself, is now transferred to the difference between the subjective and objective representation. If neither of these two distinctions can be drawn, then the tree in the first case and the meaning of the concept “tree” in the second case are themselves products of the psychological process of thinking.

This process may be conceived as either transcendental or empirical. The ostensible superiority of the transcendental thesis is that it avoids psychologism, i.e., the reduction of things and meanings to the status of empirical and subjective constructions or particulars, which cannot account for universality or necessity, and thus not for objectivity or truth. On the other hand, the transcendental thesis has the defect that it places an apparently impenetrable screen between the activity of thinking and the beings to which the meanings of concepts refer. Even if these meanings are “objective” in the sense of universal and necessary, the truth that they manifest is a spontaneous production of the activity of thinking. With no access to a natural or ontologically independent world of beings about which we are thinking, thought is inevitably transformed into the thinking of its own products.²

The question then arises as to how to justify the validity of any particular analysis of the laws of transcendental thinking. I would restate this as the question of how to distinguish between the necessity of rationality and the contingency of imagination. In more concrete terms, the problem is as follows. There seem to be only three possible relations between thinking and what we think. Suppose that I am thinking about a tree. Either I think the tree itself, or some perceptual and conceptual representation of the tree, or else I think a semantic entity which I will call the meaning of the concept “tree.” If I think the tree itself, then there seems to be no distinction between being and thinking; I become the tree by thinking it. As Aristotle puts it, “the soul is somehow the beings.”³ “Somehow” allows for a distinction, but it is not very informative. If I think a perceptual and conceptual representation of the tree, then it must be explained how this representation corresponds to the tree; but the attempt to explain this leads to an infinite regress. Finally, if thinking about the tree is thinking the sense or meaning of its concept, the same problem arises as in the case of representations. The only difference here is that the meaning may be regarded as “objective” or distinct from psychological processes. But I must now explain a three-term relation, namely, that in which the representation mediates between the meaning and the tree.

In short, the problem of the spontaneity of thinking is this: unless we are prepared to accept the absurd thesis that we become what we think, or to give a satisfactory account of the Aristotelian “somehow” (*pos*), it seems that we must admit that the activity of thinking separates us from what we think about. In the Kantian version, thinking separates us from the thing in itself, or gives rise to transcendental philosophy. It then remains to be seen whether transcendental cognition can be prevented from deterioration into the very psychologism it was intended to overcome. This is the problem of the spontaneity of thinking.

The problem becomes less severe, but is by no means dissolved, when we see that thinking has no structure. Whereas thinking produces structure, it is not a machine but an activity; and this is the sense in which I suggest that thinking is spontaneous. We are free to make contact with items in the world because of the openness of our cognitive power, and this is the basis for all productive work, including that of the imagination and the will. This is what Aristotle means by the elliptical statement that the soul is somehow the beings. The progress of epistemology, or the attribution of structure to the intellect, results in a steady separation of intellect from beings, or the transformation of *onta* into objects, that is, artifacts of cognition. When the rules for the production of these artifacts are conceived as spontaneous, our epistemologies collapse into contingency.

Thinking can be spontaneous in at least the following two senses: (1) as independent of any regulation whatsoever, whether from an external world or the inner necessity of the activity of thinking; (2) as independent of external regulation by a world of things in themselves, although bound by the rules of its own activity. It will be my contention that the assertion of sense (2) deteriorates because of internal inconsistencies into the assertion of sense (1). But sense (1) leads to chaos, not to an explanation of what we actually do when we think. As is the case with all philosophical problems, our only basis for avoiding chaos is to cling to the facts of everyday life, and to accept no explanation that contradicts them. Our goal must be to come as close as possible to an account of thinking things rather than concepts or meanings, but in such a way as not to become identical with them.

My thesis could also be stated as follows: the reason why we can distinguish between subjective and objective representations, or in other words between spontaneous creations of the imagination, transcendental or empirical, as regulated by the will, and meanings that are the same for

anyone who is capable of grasping them, is because this distinction is rooted in a world that is not produced by our thinking. The meaning of the concept “tree” is objective only because there exist trees to which we have access by perceiving and thinking. But the same is true of any concept, whether of a moral property, an aesthetic value, a logical relation, a mathematical function, and so on. These concepts have objective meanings only because we have understood something about the world, including ourselves as thoughtful residents in it, that is, not merely a product of the activity of thinking.

In short, I shall follow the procedure of distinguishing between thinking as a process and the “objects” of thought, by which expression I mean entities like sensations, perceptions, concepts, arguments and intuitions of form. But the objects of thought will be held to be distinct from and to refer to or to derive their meaning from beings in a world that is not produced by thinking, or from properties or relations of those beings. Thinking could then be spontaneous in two distinct senses: (1) the process could be autonomous or neither reducible to nor derivable from other processes, e.g., neurophysiological processes; (2) whatever the sources or causes of the process of thinking, the objects of thought could be produced spontaneously within that process, rather than caused by or derived from beings external to the process. I am here interested in the question of spontaneity in the second sense. In this sense, the question is whether the process of cognition produces its objects, which are themselves the beings that we think; or in other words whether the maxim that we know only what we make is true because thinking is making.

II

In turning now to Kant, I will begin as cautiously as possible, by wrapping myself in the protective mantle of my friend and colleague, Henry Allison. In his book, *Kant's Theory of Freedom*,⁴ Allison points out that the function of spontaneity in Kant's doctrine of rational agency is to be understood not via the moral theory but in “his views on the spontaneity of understanding and reason in their epistemic functions” (36). Allison then provides a convenient summary of those views, which may itself be summarized as follows. Spontaneity is immediately evident in apperception, or the awareness of ourselves as possessing two faculties, understanding and reason. In the exercise of the first faculty, we take some-

thing as something, or make predicative judgments, and as Allison puts it, we also know "in some sense" what we are doing. ("In some sense" sounds suspiciously Aristotelian.) In the second case, we take premises as justifying the conclusion or, at a higher level, as Kant says at A548/B576 of the first *Critique*, "reason . . . frames for itself with perfect spontaneity an order of ideas." Allison glosses this passage as follows:

[R]eason has the capacity to form ideas and to regulate enquiry in accordance with these ideas. Such formation and regulation involve spontaneity because, rather than simply reflecting a pre-given order of nature, reason projects an order of its own "in accordance with ideas," that is, it generates a set of norms and goals of explanation in terms of which scientific enquiry must proceed (38).

Guided by Allison's formulation, we may venture forward a step or two into the first *Critique*. The fundamental distinction is of course between spontaneity and receptivity. At B132, Kant says that pure or original apperception is "that self-consciousness which, while producing the representation 'I think,' which must be able to accompany all [other representations], and is one and the same in all consciousness, cannot be accompanied by any further representation." Kant calls this apperception "an act of spontaneity, i.e., it cannot be understood as pertaining to sensibility."

Since it is via sensibility or *Empfindungen* that we come into contact with the things in themselves, or with what was called in the pre-Copernican terminology the order of nature, it is obvious that apperceptive spontaneity frees us from subordination to such an order. We become the masters and possessors of nature in a sense related to that intended by Descartes. Nature is constituted by the activity of apperception, that is, by understanding and reason. I note in passing that we also free ourselves from subordination to a transcendent God by the spontaneous formulation of moral rules which unite rationality and autonomy.⁵

The spontaneity of understanding plays the active role in the constitution of objects, whereas the spontaneity of reason projects the unity and totality of the world as the theater within which rational agents carry out their theoretical and practical roles. Kant also speaks in the first *Critique* of transcendental freedom as the absolute spontaneity of a cause that is not part of the natural sequence of causes within the world of appearances that is constituted by the synthesis of sensibility and thinking (A446/

B474, A473/B501). Since the spontaneity of thinking is not itself part of that sequence, which latter is instead a result of the synthesis of sensibility and thinking, the only obstacle to the transcendental freedom of our faculty of thinking would seem to be its limitation by and dependence upon sensibility.

Despite his insistence to the contrary, Kant himself takes a crucial step toward this emancipation of thinking by the very doctrine of sensibility as receptive. Sensibility is rooted in sensations (*Empfindungen*) which arise from the impingement of what is in itself unknowable onto the pure forms of intuition. The resultant sensations are thus not strictly speaking received, but are rather produced in response to the unknown and unknowable outer stimuli.⁶ It has often been observed that, by the terms of his own doctrine, Kant has no right to posit the unknowable things in themselves as causal agents of our sensations. Whether for this reason or for some other, Kant himself, as is evident in the *Opus postumum*, tended in his later years to think of sensations too as produced by the cognitive activity.⁷

History teaches us that it did not take long for philosophers to dissolve the limitation of sensibility on the absolute freedom of spontaneous thinking. Somewhat more difficult, but not beyond the scope of the human imagination, was the dissolution of what I will call an “internal” restriction on spontaneous thought by categories and regulative Ideas. The essential point of the argument can be restated as follows: we know that the world is intelligible, and intelligibility includes knowledge of necessity and causality. But necessity and causality can be known only if they are produced by thought itself. Since the being and the knowability of objects depend upon and are consequences of the activation of these transcendental products, it follows that “the world” of objective experience is a product of thinking. At this point the question arises: why could not thinking produce some other structure of intelligibility, e.g., some other table of categories or some other manner of conceiving orderly succession in phenomena? Why is knowing a world not synonymous with a way of world-making?

Kant in effect rejects this alternative by pointing first to the world as we experience it, and second to his theoretical conception of the nature of rationality. But the first procedure is inefficacious, since Kant denies that the world as we experience it is self-explanatory; in fact, Kant follows the early modern scientifically oriented philosophers in insisting that we disregard the everyday or apparent world, precisely in order to understand it

as a world of appearances (in the technical sense of that term). Note that this is already tantamount to denying that thinking puts us into direct touch with an independent world, of which we ourselves are members; instead, we are in touch with conceptual constructions. Kant in effect substitutes for everyday experience a theoretical explanation of the inner or genuine nature of experience.

Differently stated, we experience the world in many ways, and Kant is attempting to articulate the principles of rational experience. Reliance upon a certain conception of rationality that is derived from actual scientific knowledge, is at least empirically or historically unpersuasive, since the progression of knowledge produces new categories and, even worse, new doctrines according to which categorial structure is a product of the imagination. It could also be said to be circular, since the rejection, for example, of continuous creation is certainly motivated by the rationalist desire to render the foundations of the world accessible to human intelligence.

Kant's doctrine of the spontaneity of thinking was plainly intended by him as an essential ingredient in the reconciliation of rationality and freedom. The spontaneity of thinking, as conceived by Kant, frees us from an incomprehensible deity but also from an independent and finally inscrutable nature. It saves rationality from skepticism as well as from determinism. Skepticism is conquered by the ostensible demonstration that thinking produces or spontaneously actualizes as the rules of world-order. Mankind is preserved from determinism because the rules bind only the world as produced from the synthesis of thinking and sensing; they do not, in other words, bind the spontaneous thinker.

Kant may never have noticed the extraordinarily dangerous implications, or possible consequences, of his doctrine of spontaneity. I believe that this is because of what may be called a certain "old-fashioned" scorn for the imagination understood as fantasy rather than in its transcendental sense. It can also be explained as a consequence of his extreme confidence in scientific rationalism. Perhaps one could summarize the complex situation in Kant in something like this manner: By asking us to conceive the world "as if" it were constituted in accord with the Kantian hypothesis, Kant comes dangerously close to subordinating reason to the imagination. By virtue of the same capacity to conceive the world "as if" it obeyed Kant's rules and regulations, I can also conceive the world "as if" it did not obey those rules and regulations. This tendency is encouraged by the inner link between spontaneity and the will. In addition, the

will and the fantasy are natural allies; one suggests and the other demands. In the Twentieth Century, the emancipation of the imagination is as visible in the metaphysics of modal logic, with its reliance upon possible worlds, as it is in postmodernist deconstruction, for which the world is a text that cannot be deciphered, since to read is to write, and thus to erase what one is attempting to read. One hardly needs to emphasize the concomitant emancipation of the will.

I have already stated my opinion that the fundamental weakness underlying Kant's doctrine is his assumption that the spontaneous manifestation of rational order is not itself contingent. Kant made the daring assumption, no doubt carried away by Newtonian mechanics on the one hand and the apparent indubitability of mathematics on the other, that mankind was in possession, not merely of the paradigm of absolute knowledge, but of absolute knowledge itself. This is a later version of what I take to be the Cartesian assumption that once we are in possession of the true method, we are in effect wise, even though the actual accumulation of positive knowledge of particulars, or the continuous application of the method to new cases, may be an unending task.

Kant's root assumption can be restated as follows: he took it for granted that, to the extent that everyday experience is intelligible, it must accord with the rules and methods of the mathematical and experimental sciences. Kant subordinated common sense or everyday experience to science in this sense of the term. He did not take his bearings, as in the pre-Copernican epoch of philosophy was still possible, by the commonsense world of everyday life in its pre-scientific character. This world is for Kant just a swarm (*Gewühl*: A111), heap (*Haufen*: A122), or rhapsody (A156/B195) of sensations, at best governed by empirical laws of association and at worst indistinguishable from illusions and dreams.

Of the many consequences of this assumption, I mention just one. Since it is thinking that synthesizes the swarm of sensations into the perceived object, or in other words since the perceived object is the same as the object of cognition, this raises the problem of false perceptual identification, and with it, the famous problem of the distinction between judgments of perception and judgments of experience. A false perceptual identification must be a judgment of perception; in other words, judgments of experience must always be correct.⁸ The judgments of perception "require no pure concept of the understanding, but only the logical connection of perceptions in a thinking subject." "The judgment of experience must therefore add to the sensible intuition and its logical connection in a

judgment (after it has been rendered universal by comparison) something that determines the synthetic judgment as necessary and therefore as universally valid.”⁹

Pace Beatrice Longuenesse, I see no difference between the judgment of perception and the perceptual identification via the association of sensations as described by British empiricism. But whether or not this is so, such a judgment cannot be true for Kant in any objective sense; it is only what seems to me to be so at the moment of the perception. Truth enters into the picture only with judgments of experience. But there is no room here for falsity, and this makes the appeal to truth problematic. Furthermore, the perceived or phenomenological “looks” of sensory objects, e.g., of a dog, consist of nonperceptual categories and rules for schematization which are sufficiently general to be universally applicable, and sensory material that is sufficiently amorphous to leave the empirical object without an intrinsic nature. In short, the analysis or explanation of how we perceive a dog is equally valid for the perception of any four-footed animal, and hence it is no explanation of any particular perceptual act whatsoever.

Kant is thus faced with the difficulty of explaining the unity of the transcendental structure of every perceived object with its amorphous empirical content, or in other words of producing a dog as an object of experience, i.e., of empirical science, and so as constructed by the full transcendental machinery, that is also a contingent, historically changing being, and more specifically, of deriving a definite phenomenological look, the look of the dog, from rules on the one hand and chaotic sensation on the other.

In short, the object of experience is an interpretation from the standpoint of experience, and a transcendental construction from the standpoint of the critical philosophy. Kant transformed the view, common to the early modern philosopher-scientists, that the world of everyday experience is an illusion or epiphenomenon of the genuine, scientifically accessible world, by imposing transcendental order onto the illusion. But if order is transcendental, then it is of course not immanent, i.e., it is not a characteristic of experience apart from thinking. By making order a product of rational thought, Kant prepared the ground for the transformation of theory, or the discovery of rational order, into interpretation, or the imaginary creation of a model of rationality.

But an interpretation of necessity as the spontaneous actualization of thinking is itself a spontaneous actualization of thinking. If the laws of

order arise spontaneously, they do not lose their spontaneous nature from the fact of their actual authority. If it is a matter of the spontaneity of thinking that $2+2=4$, then another surge of spontaneity could produce non-standard integers, or still more radically, a world in which the principle of non-contradiction no longer holds. Kant cannot block this possibility merely by insisting that his model is final; and there is no way in which to demonstrate the finality of any interpretation of natural order, if natural order is itself an interpretation, i.e., if there is no access to nature apart from or in advance of the choice of an interpretation.

Let me now summarize my preceding remarks. If thinking is spontaneous, then all doctrines of order and structure, hence of rationality and intelligibility, are contingent. In the language of modal logic, necessity is possible, not necessary; that is, what counts as necessary is relative to one's modal logic; and this choice in turn is a function of the spontaneous exercise of thinking, or in other words, of the imagination and the will. In the language of semantics, the objectivity of meanings is defined by the will, not by beings in the world. The next step is to replace semantics by syntactics, or meanings by rules for the use of terms designating concepts. But rules, as spontaneous, are arbitrary, not transcendental.

This situation arises when order, structure, rule-governed thinking, or more comprehensively, rational thought, is separated from everyday experience—when everyday experience is defined as a *Gewühl* or swarm of deteriorated antecedent interpretations. Kant's critical philosophy does not resemble a formal calculus, but it is a *model* of experience in exactly the same sense that mathematicians construct models as interpretations of formal calculi. Kant's critical philosophy is a model constructed to explain the formal structure of experience. The fact that Kant's model is not written in the language of mathematics, or that it cannot itself be formalized, is in the present context irrelevant.

The transformation of philosophy into model-theory, or in other words of theory into interpretation, is rooted in the previously noted unofficial axiom of modern philosophy that we know only what we make. This axiom is understood to mean that we make the objects of our thought *within* the process of thought and then project these objects "outward," whether onto extension in the void, unknowable things in themselves, or sheer nothingness. What we are no longer taken to do is to apprehend beings that exist and subsist externally to and independently from the process of thinking, i.e., that we apprehend them *as* they exist or as *what* they are.

Experience is so to speak "formalized" or defined within a model. But the relativity of experience to the model means that experience is contingent, since models are themselves not necessary or imposed onto us, but contingent products of thinking. The models are themselves spontaneous in the sense that they depend upon a mixture of human intentions, perspectives, traditions or antecedent scientific doctrines, and so on. They are of course not spontaneous in the sense of coming into existence *ex nihilo*. Models are obviously responses to experience and their adequacy is determined by "comparing" them to experience. But this is a circular procedure, since what we mean by experience is itself a derivative of our models.

In fact, model-builders proceed inconsistently, since they assume the stability or independence of experience in the process of checking the adequacy of the model. But this stability or independence cannot itself be due to some other model; in this case the process of checking models would be like checking the consistency of one formal theory within, and so relative to, another. Science would then be surrealism or perhaps a kind of literary criticism in which one imaginative interpretation of the text of the world is checked in the light of another imaginary interpretation. This process is also known as axiomatic mathematics.

This was of course not Kant's intention. Kant assumed that the spontaneity of thinking served as the guarantee of the complete accessibility or intelligibility of the structure of the world to its maker; he also assumed that arbitrariness was excluded by the fact that spontaneous thinking must actualize in the shape of concepts, categories, Ideas, or rules. Aristotle defines man as the animal who possesses *logos*. In Kant, we see the ancestor of Heidegger's definition of man as the animal who is possessed by *logos*. What Kant did not anticipate was the contingent status of the efficacy of *logos*, or what Heidegger calls its "historicity" (*Geschichtlichkeit*).

This is the background of the problem of spontaneity as it emerges from the philosophy of Kant. If my sketch has been accurate, it would seem to be the case that, if there is any chance of avoiding or resolving the problem, it depends upon our reversing the so-called Copernican revolution, and returning to the pre-Kantian situation, in which thought conforms to the object, or, as it should then be called, the entity or being, rather than the object to thought. In my opinion, something like this is indeed necessary; but as I have just stated the alternative, it is too general and in fact, merely a shift from one extreme to another. Without entering

into elaborate historical detail, let us briefly consider the major implications of the pre-Kantian hypothesis, namely, that thinking conforms to the object of thought (more precisely, to the thing or entity).

III

On the Kantian hypothesis, to say that the object conforms to thinking is to say that the object of cognition, perceptual as well as conceptual, is a product of the cognitive process. To be sure, Kant insists that the object is, as Allison puts it, one aspect of the independently existing entity or thing in itself.¹⁰ However, the object or *Erscheinung* is the only aspect that is accessible to cognition, as distinct from speculative and hence unscientific thinking.

In my opinion, even if we accept the “two-aspect” interpretation of the “phenomenon–noumenon” distinction, the thing in itself has the status of a hypothesis or, to employ a Hegelian expression, a *Gedankending*, an entity of thought, and so, of that which is immediately open to interpretation in two different ways, either as itself produced by thinking, and so no longer a thing in itself, or in the more extreme case, as a phantom of the imagination. In short, conformity of the object to thinking, even if it is initially asserted to rest upon the independent subsistence of the thing that underlies the thought, leads directly to the disappearance of the independently subsisting being. There are no more beings but only objects. Thinking is now spontaneous in the sense that it may think whatever it wills.

On the reverse hypothesis, namely, that thinking conforms to the object (which is of course not an object in the sense of a projection of a subject), it seems at first that there is no reason to deny spontaneity, and so freedom, to thought. In order for thinking to conform to what it thinks, it must be initially independent. One conjures up the scenario of a thinker approaching a candidate for thought and directing his or her attention to it, conforming to it in order to apprehend it, but not thereby sacrificing the integrity of the activity of cognition. The thinker can turn away from the candidate, cease to consider it, or indeed, imagine it in a variety of contexts or identities that deviate from those which are initially given.

This scenario is defective for more than one reason. In the first place, it confuses the process of thinking with the status of the things thought. I may be free to think or not to think, even though, once I begin to think, I

am compelled to think this rather than that. This gives rise to a second difficulty, which turns upon what we mean by "conformity" to the object, i.e., to the being. The word "conformity" suggests a common form; in other words, the process of cognition is one in which thought not simply shares, but receives, the form of the being, and so is "informed" in the sense of being *enformed* or shaped by it. In this case, the process of thinking resembles that in which there is a continuous changing of forms within the intellect, something like an animate and sentient slide-projector being operated by a god who changes the slides and thereby alters the show, not on an external screen, but within the slide-projector itself.

If conformity means literally "having a common form," then there is no spontaneity in thinking but rather comprehensive subordination to, and definition by, the object of cognition. Whether man possesses *logos* or *logos* possesses man, we think as we are told. Or rather, we do not think at all, except in the restricted and highly artificial sense of becoming the form of the thing that we cognize. A tree does not think by virtue of exhibiting the look or *eidos* "tree." If therefore my intellect takes on the form of the tree through its conformity with the being of the tree, it would seem to be as thoughtless as the tree. Indeed, it *is* a tree, in the decisive sense of "is."

It is also worth noting that the doctrine of strict conformity shares the defect of Kantianism in being unable to distinguish between cognition and perception. To perceive a tree is to know that it is a tree that I perceive. But to know a tree is to conform with or assume its form. In other words, the distinction between recognition and knowledge of essential or defining properties is inexplicable. For what does it mean to recognize or identify something as a tree, if not to think "this is a tree"? And how am I to think that, except by a conformity of thinking to the being, i.e., the tree? There are not here two processes, one in which I am shown a photograph of a tree, and another in which I am asked to identify it. Or rather, if there are two processes, then thinking is not conformity with the thing, but detachment from it.

Let me turn directly to this alternative interpretation of the pre-Kantian thesis, in order to see whether it brings us closer to the truth, or at least to a plausible account of our actual experience. Suppose then that thinking is not assuming a form, i.e., taking on a definite shape or look, but rather trying to identify more precisely something seen at a distance. This is of course not intended as an exact description but merely as an introductory metaphor. If it is helpful, we can think of this step as a shift back from

Aristotle to Plato. Let us note first that in this metaphor, thinking conforms to, in the sense of approaching, an independent being rather than leaping spontaneously into action by constituting the object of thought. On the other hand, the form of the candidate for thought does not itself leap spontaneously into our thoughts with a tyrannical determination that leaves no room for reflection.

To think something is not to become it but to consider it. This is the sense in which I take Husserl's observation that intuition is not thinking but supplies evidence to thinking.¹¹ If the strict-conformity explanation is adopted, it becomes immediately necessary to introduce a supplementary hypothesis of two intellects, one that conforms to the form and another that reflects upon the state of the first intellect. But this leads to an infinite regress, since the only way in which the second intellect can consider or think about the status of the first intellect is by conforming to its assumed form. Thus a third intellect is required to reflect upon the result, and so on forever. I might also mention a related difficulty. If thinking is strict conformity, then we can only think of one thing at a time. But in this case, reasoning, inferring, or arguing, i.e., consecutive thought, becomes impossible. We cannot compare two thoughts, but only change forms in a discontinuous and so thoughtless manner.

These problems are not removed, but occur in slightly different form as a consequence of the Kantian thesis, i.e., on the assumption that thinking is constituting the object. We must, I think, reject both the constitution thesis as well as the strict-conformity thesis and look more closely at the thesis of detachment from but approach toward the candidate for thought. On this thesis, thinking is not spontaneous in the sense that it does not produce its objects, or at least not all of them. But it is spontaneous in the sense that it must engage in activity that is independent of control by the things about which it is thinking. The control on thought must come from the purposes for which we think, and only secondarily from the things or objects of thought. As a corollary, I note that freedom of thought is now not only possible but a necessary feature of the activity of thinking. I cannot know the truth unless I am free to err. To know something is not to be it; the thesis that being and thinking are the same is just unthinkable, not to say absurd.

It does not however follow that the thesis of detached approach is free of all difficulties. To mention what will be for many its most disappointing limitation, it does not tell us what thinking is like. But this is because thinking is *sui generis*. It is not like anything else nor can it be analyzed

into its constituents, because it has none. The partisan of detached approach, in other words, rejects epistemology or the attribution of cognitive structure to consciousness, because structure is not conscious and is incapable of cognition. It blocks access to what we wish to think about and replaces these items with artifacts. In so doing, it replaces spontaneity with rules.

None of this is to deny that thinking identifies, analyzes, or describes structure, or that it produces rules, methods, forms, arguments, and so on in an apparently limitless richness. But that is an entirely different matter from the attempt to construct understanding from the operation of cognitive machines. It is all very well to talk about precise conceptual analysis. But we have to begin with that which is to be analyzed, not with a set of extraneous analytical tools which may or may not fit what is to be analyzed. Furthermore, it is one thing to analyze something with a set of tools, and something else again to replace what is to be analyzed with a construction made up of the tools themselves. Nor can I construct the process of thinking from a collection of anecdotes about conforming to, or approaching, or for that matter constituting, trees, dogs, or persons. If I am going to construct a "model" that purports to explain thinking, the model must itself *conform* to thinking; it begs the question to force thinking to conform to the model.

Shocking as it must sound to Kantians, I believe that the myths and metaphors employed by Plato to describe the process of thinking are more accurate than the quasi-scientific epistemologies of modern philosophy. That is, they are truer to the phenomena because they are not precise, or let us say that their precision is poetic. Kant of course regarded Plato's dialogues as *Schwärmerei*. "*Die Philosophie des Aristoteles ist dagegen Arbeit*." Perhaps so, but there is also no epistemology in Aristotle. And I defend poetry over epistemology, odd as it may sound, on Kantian grounds, because of the antagonism between epistemology and spontaneity. We Kantians must never tire of asking ourselves the simple but indispensable question. What do I do when I think? Do I generate trees, dogs, and human beings by operating a series of interconnected machines? Is thinking like Charlie Chaplin caught on the assembly line in *Modern Times*? Do I turn into trees, dogs, and human beings in a series of discontinuous and surrealist episodes, as though I were a peculiarly amorphous being at a costume ball who constantly changes his disguises?

Allow me to repeat the deep truth, made famous by Kant and Wittgenstein but first enunciated by Plato in his dialogue, *Statesman*, that

there are no rules for the application of rules. Thinking is spontaneous because it must determine continuously what comes next, what it is to say now, which of its constructions are here applicable, and so on. No model of the mind can tell us what comes next, and so no model can verify itself. But there is no model of thinking, where by thinking is meant the activity of following rules in order to manufacture intellectual commodities. When looking for causal relations or when making perceptual identifications, I must be engaged in an *unruly* activity, one to which no rules apply, because I am in the process of applying rules, or of deciding which rules to apply.

So the question is not simply one of the difference between the “phenomenological” or empirical stream of consciousness on the one hand, and on the other, the scientific or philosophical analysis of the underlying structure of the stream of consciousness. Thinking is in the first place not a stream of consciousness, or not a stream of consciousness alone, but the intelligible and intelligent consideration of everything under the sun. In the second place, the analysis of structure is not the uncovering of the structure of the intelligent consideration of everything under the sun, because considering has no structure. There are no doubt rules, methods, and procedures to facilitate consideration, but these are not the same as the act of considering itself. It is at this point that Kant’s doctrine of apperception becomes necessary.

The spontaneity of thought consists in its ability to produce rules, and so to consider the circumstances in which they may be applicable, and hence when they are irrelevant or superfluous or misapplied. This may also be expressed by saying that we are always in a position to change our minds. If this expression sounds too chaotic to the proponent of rules, I note that it accords with the Aristotelian analysis in the *De Anima*, according to which the intellect, as the “form of forms” or *topos* of forms, cannot have a determinate form itself. One can bring Aristotle into harmony with Kant by saying that the unity of apperception has no more structure than does *nous*.

If the mind, or more specifically the intellect, i.e., thought, had a determinate structure, whether Kantian or of some other variety, then it could never entertain thoughts of entities other than itself. For this reason, the faculties or machinery attributed to thought by Kant and other model-theoreticians cannot be understood as the actual structure of thought, in the sense that the shape of a tree or a dog in fact sets the limit to the essential being and acting of the tree or the dog. The rules may bind in the

sphere of their applicability, but the use of the rules is spontaneous. And this means that our thinking about the rules, and so too about which objects are covered by the rules, is spontaneous.

I am not arguing here that it is entirely a question of will or choice on my part as to which objects I shall think or perceive. Sometimes it is, but more often it is not. And the cases in which to think is to produce *objects* are all derivative from, and so dependent upon, cases in which thinking (and perceiving) is the thinking (and perceiving) of beings that I have not produced. My point is not that I produce what I think or perceive; I am only partly a Kantian, as I hope to have made clear. My point is rather that in coming to grips with the elements of the continuum of experience in which I am situated, I neither produce nor identify with these elements. I approach them, or attempt to approach them, and in this way to discover the rules, if any, by which their nature and behavior is determined. But in making this approach, I do not engage in a rule-governed process of switching on engines, applying tools, or following a manual. How I approach the elements of my experience is determined by the nature of the element, or rather by what I apprehend of that nature when it first attracts my attention. The codifiable procedures of thinking, and so the rules which may be formulated to express the activity of thought within a given domain, are discovered only by reflection upon the activity of thinking about things of a certain kind. They are not discovered by emptying the mind of all specific content other than its own activity and then inspecting that empty activity itself.

Thinking is thus both spontaneous *and* rule-governed. But it could not be rule-governed if it were not spontaneous in the sense just noted, because an inability to detach oneself from, and to reflect upon, rules, is here equivalent to being unconscious or not thinking at all. And this is the explicit contention of those who insist that the mind is a machine. As I employ the term, spontaneity refers to thinking, but not primarily to the production of the object of thought by thinking. Thinking is regulated, not by its inner rules of manufacture, but in the first instance by the purposes for which it thinks and in the second by the natures of what it thinks about. It is true that many objects of thought are produced by thinking. But these objects, if they are anything in particular, themselves have natures that regulate what we think about them. And their possible creation is itself rooted in the actual nature of the beings of everyday experience, from which we deviate or concerning which we improvise, or which we attempt to analyze or explain.

If this is granted, there is no danger of the domination of thinking by the imagination, since the imagination is also held in check by the objects of thought. On the other hand, since thinking is detached from what it thinks about, we are able to explain error and success, as well as reasoning in a sequential or deductive manner; and we are also able to explain why we have no luck in acquiring perfect knowledge about the objects of thought, or in an older idiom, about why we are not wise.

This is because we can never close the gap between thinking and the elements about which we think. This gap, incidentally, is not an abyss. We are often, but not always, free to consider or not to consider beings of such-and-such a kind. But once we decide to consider such beings, we are bound to consider them as they are; otherwise, we are considering something else. It therefore turns out that the freedom of the will, the passport by which we cross the boundary between one set of rules and another, is restricted by a requirement analogous to that referred to by Pierre Bayle, who said that he changed his religion at the border.

IV

In closing, I want to emphasize a point to which I have just alluded. I have denied that we manufacture what we think about entirely from the process of thinking itself. But I am far from denying that the activity of thinking contributes to or modifies what we think about. It does in fact make sense to talk about perceptions, thoughts, concepts, categories, logical rules, and so on. The experience of thinking itself, and not a transcendental, analytical, or scientific theory, forces us to acknowledge two facts.

First: we are capable of identifying and understanding what we think about in the ordinary or pre-theoretical senses of “identifying” and “understanding.” Everyday experience is not a *Gewühl* or a chaos of random and arbitrarily associated sensations. Second: we are often mistaken in our identifications or presumed understanding of this or that element or sector of everyday experience, but we are never entirely mistaken, and the errors we make are corrigible by access to other elements or sectors of everyday experience.

I infer from this undeniable basis that we produce neither the beings nor the process of thought that underlies all subsequent cognitive production, but that we are capable of an initial apprehension of what presents itself which is not sufficient to be called “conformity with” in the sense

of either being or knowing perfectly the elements we apprehend, but which is sufficient to license progress and decline. Equally if not more important, we make a sufficient approach to things to engender reasonable activity. All this means that perceptions, concepts, categories, and so on, are not the initial medium of contact with the elements of or candidates for thinking, in the sense of artifacts manufactured by thinking in order to furnish us with, or to place us into contact with, whatever we think.

Thinking is like moderately good vision for which spectacles are required. If thinking were blind, spectacles, however strong, would be useless. But neither can we explain thinking as the *use* of spectacles, nor is it permissible to explain how thinking functions by applying to it the principles by which we manufactured the spectacles. On the contrary, we manufactured the spectacles by applying the principles of thinking, or in the metaphor, of optics, together with a variety of other kinds of knowledge, all of them consequences of thought, not thought itself.

Notes

1. For the history of this distinction, see J. Alberto Coffa, *The Semantic Tradition from Kant to Carnap: To the Vienna Station* (New York: Cambridge University Press, 1991).

2. "But we can only understand and communicate to others what we ourselves can *make*"; letter to Beck, July 1, 1794 (11:514), translated in Kant's *Philosophical Correspondence, 1759–99*, ed. and trans. A. Zweig (Chicago: University of Chicago Press, 1967), 216. Kant is referring here to synthesis as opposed to intuition. See also Bxiii in the first *Critique*.

3. *De Anima* III.8, 431b21 ff. See also *De Anima* III.7, b16: "*holos de ho nous estin ho kat'energeian ta pragmata*."

4. Cambridge: Cambridge University Press, 1990.

5. Cf. Bernard Rousset, *La doctrine kantienne de l'objectivité* (Paris: J. Vrin, 1967), 409: Autonomy means that "*nous construisons l'objectivité du réel par nous seuls, sans appui sur aucune transcendance, ni extérieure, ni intérieure, sans Dieu, ni âme, comme simples consciences humaines*."

6. Compare here Beatrice Longuenesse, *Kant and the Capacity to Judge*, trans. C.T. Wolfe (Princeton: Princeton University Press: 1998), 300. On her interpretation "matter of appearances" corresponds to sensation. The latter is given; the former is "apprehended" via a figurative synthesis.

7. See Vittorio Mathieu, *Kants Opus postumum*, ed. G. Held (Frankfurt/M.: Vittorio Klostermann, 1989), for details. Compare Michael Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992), 238: the *Opus*

postumum is designed to supply a priori foundations to the forces of empirical physics.

8. Among the various discussions of this point, I mention Gerold Prauss, *Erscheinung bei Kant* (Berlin: Walter de Gruyter, 1971); Hansgeorg Hoppe, *Synthesis bei Kant* (Berlin: Walter de Gruyter, 1983), and Henry Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983). Allison (72 f.) follows Prauss' suggestion with respect to the impossibility of equating objective judgment with truth, since then every judgment of perception would be true. Prauss suggests that objective validity for Kant means simply the capacity to be true or false. But this capacity depends upon the determinateness of the perceptual judgment as possibly true or false of this particular object. Otherwise put, Prauss' solution falls prey to the criticism by Hoppe, which can be summarized as follows: private perceptual judgments are objectively determined prior to the functioning of the unity of apperception. A private perception is of some definite object; it is not a *Gewühl* (A111) of random sensations (Hoppe, 109, 149). In other words, we do think of objects by association alone. The form must already be present, actually or *in potentia*, in order for the content of a private perception to refer to some particular object. If it does not so refer, then there is no private perceptual judgment at all. A further difficulty is that categories cannot be determinate concepts; that is, they cannot account for the perceived form of the object. In sum, Kant has not explained the precategorical or conceptual shape of associated perceptions. And he explains error as the subsumption of an object under the wrong concept, without explaining how we can identify the object we are subsuming, rightly or wrongly, as a particular object. If the private perception is pre- or non-objective, then it becomes so via subsumption, in which case the judgment is always correct. But if the perception is pre-objective, this gives rise to an infinite regress. And neither the category nor the schema can explain the perceived form of the particular object, e.g., a dog as distinct from all other four-legged animals (see A141/B180). As a coda, I note that it is curiously Aristotelian to speak of sensation as possessing forms *in potentia*.

9. Cited from the *Prolegomena*, §19 (4:298–9), in B. Longuenesse, *op. cit.*, 194 n. 47.

10. Allison, *op. cit.*, 3–4.

11. *Logische Untersuchungen* (Tübingen: Max Niemeyer Verlag, 1968), II/I, 168.

Lewis White Beck's Account of Kant's Strategy

Graham Bird

Lewis White Beck's formidable reputation among Kant commentators was not confined to the major contributions he made to moral philosophy. He contributed also and significantly to Kant's theoretical philosophy, and in "Kant's Strategy"¹ gave an account, part historical part philosophical, of the central themes of the *Critique of Pure Reason*. It is a tribute both to his vision and his scholarship that he was able in that paper to combine so successfully a very general theme with careful attention to the Kantian detail. It is not my intention to query the account he gave there, for it seems to me both illuminating and correct. What I want to do instead is to push his account one stage further in two different but related directions. He himself began the paper with a military metaphor of a war conducted by Kant on two fronts, and one way of putting my aim would be to say that I would like to advance on two sectors of the battlefields he identified. In section I, I summarize the salient features of his account; in section II, I argue that his account can be extended in one important way; and in section III, I comment on a recent philosophical discussion by John McDowell² of the central theme Beck identifies in Kant, namely the integral relationship between sense and understanding famously expressed at A51–2/B75–6: "Without sensibility no object would be given to us, without understanding no object would be thought. Thoughts without content are empty, intuitions without concepts are blind."

I

Beck's central theme in "Kant's Strategy" is that Kant, in waging a war on two fronts against empiricists and rationalists, identifies and rejects an assumption common to both opposing forces. That common assumption is that knowledge has "but one ultimate source." The two opposing positions identify that ultimate source differently; it is typically identified by empiricists as sense experience and by rationalists as reason. Kant rejects

both views not only in his explicit appeal to a necessary connection between sense and understanding in the passage at A51/B75, but also in the still neglected and misunderstood section *The Amphiboly of Concepts of Reflexion* (A260–92/B316–49).

Kant's strategy, so understood, has two noteworthy features. First it makes room for a proper understanding of both science and metaphysics. It argues, for example, that mathematics cannot contain only analytic truths, and that metaphysics cannot aspire to establish truths which transcend experience. Beck provides a clear summary of Kant's developing thought on these issues from the 1760s to the 1770 *Inaugural Dissertation*, but I shall not comment on those historical antecedents to the first *Critique*. Second the strategy exemplifies a formal principle associated with Ramsey or Wittgenstein, namely that of locating an assumption shared between two opposed positions so that there is no genuine forced choice between them. For if the two positions share a common assumption, then to reject that assumption is to reject an exhaustive opposition between them and so to allow the possibility of a third position. That formal technique runs through the *Critique*, but receives perhaps its clearest expression in the Preface (Bxvi) and the *Dialectic*, especially the *Antinomies*. In the passage at A502–8/B530–6 Kant is quite explicit about the formal strategy in which he distinguishes between 'dialectical' and 'analytical' opposition (A504/B532). The formal principle is outlined at A503/B531:

If two opposed judgments presuppose an inadmissible condition, then in spite of their opposition, which does not amount to a contradiction strictly so-called, both fall to the ground inasmuch as the condition under which alone either of them can be maintained, fails.

At its most general the strategy is reflected in Kant's advertised attitude to previous metaphysics in which a futile and endless oscillation between apparently opposed positions has led to only mock disputes (Bxv). Kant represents his own project as turning away from the path of earlier metaphysics and striking out in a quite new direction. The so-called 'Copernican Revolution'³ is both a revolt against previous philosophical positions with their common assumptions, and a 180 degree turn away from those earlier paths.

Beck decorates this theme with illuminating detail, and is careful to avoid the over-simple idea that Kant's philosophy is an eclectic mixture of fragments from the rejected positions. He adds also an important dimension to the account in noting, quite correctly, that the strategy pro-

vides an essential resource for the transition to Kant's moral philosophy. In limiting the domains of science and metaphysics to experience it explains, and to some degree validates, our tendency to travel conceptually beyond the bounds of sense experience. In both disciplines the tendency may issue in a Platonist ontology which for Kant is an illusory hypostatization of our ability to abstract from sense experience. Our conceptual capacities enable us to fabricate such abstract expressions and so to outrun sense experience in a way which has both a good and a bad side. The good side reflects our imaginative ability to construct theories which may then be exposed to experiential test; the bad side reflects an arrogant presumption that we can gain genuine knowledge of noumenal reality whether in science or metaphysics. Reason has two faces and Kant's account enables us in principle to distinguish them; to value the acceptable but to limit and distrust the unacceptable face. Beck is right to see Kant's defense of transcendental freedom in the Third Antinomy as an example of this view; it enables us to secure ordinary, practical, freedom but not to provide any knowledge of a noumenal reality.

II

Although Beck correctly identifies a common assumption in both empiricist and rationalist philosophers it indicates only one, admittedly central, application of Kant's formal strategy. There may be *other* common assumptions which Kant wished also to reject in the background of the same philosophers. There is after all no doubt of the extent of Kant's conviction that all previous philosophy had committed fundamental mistakes in its underlying assumptions, and a passage from the Preface of the *Prolegomena* (4:255) testifies to the vehemence of his attitude:

My intention is to convince those who value their interest in metaphysics that it is absolutely necessary for them to lay aside their previous work and to treat all earlier developments as non-existent so that they can first ask whether metaphysics is even possible.

If there are other common assumptions in the earlier disputes then Kant's revolution might also be directed at them.

There is no difficulty in locating at least one other such assumption. For implicit in both empiricist and rationalist philosophers is a more gen-

eral assumption about the essentially subjective basis for experience. Empiricists typically take that subjective basis to consist of individual, private, sense experiences; rationalists typically take it to consist of a personal exercise of reason. Despite those differences, however, both respond to that subjectivist belief, to what McDowell⁴ calls the 'interiorization of the space of reasons'. One might say, if the terms were not typically elusive, that a common assumption in the opposed positions is precisely that any ultimate source of knowledge must inevitably be subjective and personal rather than objective and public. Moreover, that common assumption sets a familiar central problem for modern philosophy, namely how on such a subjective basis we can achieve objective experience. As Beck rightly points out empiricists typically canvass ways of bridging the gap between personal sense experience and external reality, or else admit failure and lapse into skepticism; while rationalists typically invoke principles such as that of a pre-established harmony, or appeals to God, in order to warrant the best exercises of a personal reason. Although such an account is very general there is no doubt that the opposed positions which Beck identified subscribe to that deeper assumption.

It may seem less obvious that Kant himself would have wished to identify and reject such an underlying assumption. Kant's central problem in the *Critique* might be identified as that of explaining the very possibility of objective experience, and he has commonly been interpreted as attempting to provide such an explanation within the traditional framework and its subjectivist assumption. Other commentators,⁵ however, have noted his evident wish to escape from the previous tradition, to revolutionize it and turn it upside down, and there are many other indicators of that radical change of direction. The project of explaining how objective experience is possible can be understood less as acceptance of the traditional subjective/objective distinction than as a prior attempt to elucidate it. Kant's opening remarks in the Introduction (B1–2) and *Prolegomena* §21a (4:304), which distinguish immediately between the origin of knowledge and its status,⁶ indicate an appreciation of one source of misunderstanding in the traditional positions. But although these considerations may suggest a rejection of the previous tradition they do not point specifically to a rejection of the subjectivist assumption. They are compatible with an acceptance of the unassailably subjective character of our basic experiences and with the project of justifying an objective experience on their basis. They are, so far, compatible with an engagement in the traditional disputes which Kant himself characterized as mock combats (Bxv). It is

true that such an engagement would be at odds both with Kant's evident recognition of the failures of previous philosophy and with his equally evident intention to mark out a radically new direction for philosophy, but thinkers do not always fully realize, or fully understand, their own revolutionary goals. If Kant is to be represented as identifying and rejecting that subjectivist assumption it is necessary to locate features of the *Critique* which point more explicitly towards such a position. I therefore draw attention to four passages which seem to me to point more clearly towards a rejection of the assumption and make it impossible to represent Kant as simply endorsing it.

The passages in question are in the B-edition Preface (especially Bxxviii-xxx), the initial conclusions (A26–30/B42–6, A32–42/B49–59) and the General Observations (B59–73) in the Aesthetic, and the Fourth Paralogism (A366–72). In these passages, in somewhat different ways, Kant draws a clear distinction between an empirical and a transcendental conception of an object, and of the associated contrast between objects and their appearances. He thus distinguishes empirical and transcendental conceptions of the contrast between what is subjective and what is objective, between representation and what is represented, and so marks that prior elucidation of the 'subjective/objective' distinction which the tradition took for granted. The B-edition Preface justifies our taking the distinction not as one we can make materially between two distinct but identifiable types of object, but as one made formally between two points of view, or two ways of conceiving the term 'object'.⁷ In the Fourth Paralogism of the A-edition the point is made by contrasting empirical and transcendental senses of 'external' in a more restricted conception of an 'external object'. In the Aesthetic's General Observations it is made by separating the empirical contrast between objects (raindrops) and appearances (rainbows)⁸ from the transcendental contrast between things in themselves and appearances. Finally in the earlier Aesthetic (A26–30/B42–6, A32–42/B49–59) and again in the Fourth Paralogism in the A-edition Kant identifies a four-fold classification of philosophical theories depending on combinations of empirical/transcendental idealism and empirical/transcendental realism in place of a simple traditional dualism.⁹ What is real, or objective, might be construed either empirically or transcendently, and by the same token what is apparent, or subjective, is susceptible to the same ambiguity. Kant is, in these passages, quite explicit in distinguishing the different senses for 'object' and 'appearance'. He is not as explicit in distinguishing corresponding senses for 'idea' or 'representa-

tion' but the passages commit him to that parallel contrast which is formally captured in the explicit distinction between empirical idealism and transcendental idealism.¹⁰

In these passages transcendently external objects are things in themselves, rational or intelligible objects, conceived as totally distinct from any sensory access to them (A369, A373, and A380), while empirically conceived objects are things related essentially to our senses.¹¹ Empirically 'external' objects are spatial objects essentially related to outer sense. Similarly the empirical distinction between objects and appearances divides the content of experience into, respectively, items like raindrops or planets on one side, and items like rainbows or scents on the other. Items of all those kinds are, as empirical, available within experience, but this is not true of the transcendental distinction. Transcendentally, in the philosophical context, appearances cover all items of those former types, while things (in themselves) include neither. Kant's transcendental term 'appearance' is thus not to be confused with our ordinary, empirical, term, as Kant is at pains to point out (B69–70). It is, first, a technical, philosophical, term devised in response to philosophical pressures, and, second, it draws no distinctions within experience, in the way that the colloquial, empirical, term does. The transcendental conception of appearances is on the contrary designed to mark a philosophical contrast between the world of our sense experience, including physical objects, and what may lie beyond it.

This account seems to me indisputable so far as it goes, but it occurs within a more complex theory which may sometimes seem to point in a different direction. Notoriously, for example, Kant's occasional use of the terms 'transcendental object (subject) = X' brings in additional considerations which are controversial. Similarly it is well known that Kant often expresses the central transcendental distinction by speaking of (transcendental) objects, or noumena, and 'their' appearances. This has been thought to imply that we know 'mere appearances' to be appearances 'of' things in themselves, or that the transcendental distinction should be modeled on the familiar empirical contrasts between a rose and its scent, or between a raindrop and a rainbow. Both such implications are at odds with Kant's account; the latter flagrantly violates the very distinction drawn between the empirical and the transcendental, and the former is at odds with Kant's undeniable official view that we have no knowledge of things in themselves. I shall, therefore, consider not whether Kant is wholly consistent,¹² but only the strict implications of the account I have given.

Two lessons can be learned from this account in relation to the subjectivist assumption. First is that the traditional contrast between what is subjective and what is objective can no longer remain a simple univocal dualism. That simple dualism is now replaced by a *four-fold* scheme in which the subjective and the objective are each provided with two distinct senses. The empirically subjective covers those empirical items which we think of as personal and mental, that is representations of raindrops as opposed to the raindrops themselves for example; while the transcendently subjective will cover all items of experience available to sensory experience, that is raindrops as well as rainbows. The empirically objective will, in contrast to the empirically subjective, cover those items which we do not think of as personal or mental, that is such things as raindrops as opposed to their mere appearances or representations; while the transcendently objective will range over those purported items which are not available to our senses, are indeed wholly independent of them. That is the conception of rational, intelligible, things in themselves. Kant's scheme already shows that we can no longer assume a simple traditional dualism, for there is no such simple contrast. Instead there is a more complicated set of distinctions the terms of which are all extensionally different, and whose differences are concealed by the ambiguous use of the same terms, 'object', 'appearance', and 'representation' in the common assumption.

Kant adds a second dimension to his analysis of the common assumption by showing that the ambiguities turn crucially on quite different motivations. The empirical contrasts are motivated by their use and serviceability in everyday and scientific contexts, while the philosophical, transcendental, contrasts are not. The latter are the result of specifically philosophical pressures in responding to a traditional framework with its unthinking assumptions and the ambiguous problems which they generate. In the previous empiricist/rationalist tradition there was, I think, no room for Kant's central idea that we are concerned not simply with what exists but rather with how we, in various postures empirical, or transcendental, scientific or metaphysical, conceive what exists.¹³ For the tradition, as I have outlined it, what basically exist were subjective mental representations. For Kant this is dangerously oversimple: We can not only demonstrate its ambiguity, but also assign its different senses to different contexts in philosophy or science.

There evidently is an empirical way in which the subjectivist assumption can be accepted and used as the basis for a developmental psychol-

ogy. There is no serious prospect of doubting that cognitive development begins with our subjective experiences, from which we develop an adult, objective, experience. But such a project is an empirical, scientific program in cognitive psychology and is not philosophical. One difference is that the empirical project operates in a context where the character of that adult objective experience is taken for granted. It is assumed, and not called into question, that such an experience is of objective items such as physical objects in causal interaction and so on.¹⁴ The task is not to justify such assumptions but only to map a normal development from an empirically subjective experience to that objective understanding. The task in traditional philosophy is different; it does not take objective experience for granted but aims precisely to offer a specifically philosophical justification for it. In this way traditional philosophy understands the subjectivist assumption not in an empirical but in a transcendental way. It is not that our empirical ideas occur as one part of an accepted experience which includes what those ideas represent. It assumes an adequate general characterization of the content of our experience in terms of subjective representations, and then faces the challenge of justifying the step outwards from those representations to the things represented.¹⁵ Kant's own clear distinction between his transcendental, philosophical, task and that of empirical psychology (Aix, A86–8/B119–20) shows further a potential confusion in the traditional position. The tradition accepts the subjective starting point as unassailable on the basis of an empirical project in developmental psychology, but then tacitly and confusedly abandons one of the central principles of that project, namely its acceptance of the contrast with an adult objective experience. Kant's complex four-fold scheme shows the inadequacy and ambiguity of the traditional assumption; his insistence on the nature of the two different projects, empirical psychology and philosophy, shows how a failure to mark that distinction yields confusion as well.

It may be objected that these admitted aspects of Kant's account do not establish the correctness of his own scheme, but this rests on a misunderstanding. What Kant can show at this stage is not that his scheme is correct and the alternative incorrect, but only that the traditional scheme and its assumption are not obligatory. They may seem to be so, for that is the strength of the traditional position, but in the light of their ambiguities and potential confusions they can no longer be unthinkingly assumed. What Kant shows is that the apparent unassailability of the assumption is an illusion, and one perhaps motivated by a failure to distinguish the dif-

ferent roles of empirical psychology and philosophy. If this seems a disappointingly weak conclusion it is worth recalling that this is also McDowell's position. There he identifies the same tradition as his own target:

A typical form of modern philosophy confronts a familiar predicament. It takes itself to be called on to explain how, starting from independently available data of consciousness, we work out into a justified confidence that there is an objective world (*MW*, 110).

My talk of openness is a rejection of the traditional predicament, not an attempt to respond to it (*MW*, 12).

The aim here is not to answer skeptical questions, but to see how it might be intellectually respectable to ignore them (*MW*, 113).¹⁶

McDowell's own attitude to that traditional predicament is the same as Kant's attitude to the subjectivist assumption in it, namely to show that they are at least not obligatory, and that we cannot simply take them for granted.

III

I have suggested so far that Kant follows in the first *Critique* a strategy similar to that of McDowell. McDowell's central aim is to find an alternative to a traditional conception of the contrast between mind and world in which the subjectivist assumption plays an essential part. His aim is not to accept the traditional assumptions and seek a traditional solution to the problems within that framework, but instead to show that the framework is not obligatory. McDowell like Beck cites Kant's A51–2/B75–6 connection between sense and understanding in order to acknowledge the relevance of Kant's strategy to his own thesis, but unlike Beck he focuses on the underlying subjectivist assumption which I have also emphasized. McDowell's own immediate targets are the more recent views of Davidson's 'coherentism'¹⁷ and Evans' 'realism'¹⁸ rather than the traditional empiricists and rationalists, but there are evident connections between these positions. A coherentism which abandons any constraint on thought from 'outside', that is, from the senses, has evident similarities to a rationalism which at least devalues the senses in favor of knowl-

edge through reason alone. Evans insists on a non-conceptual constraint from the senses on thought and so adopts a typical empiricism, but, according to McDowell, he then falls into the erroneous 'Myth of the Given'.¹⁹ McDowell thinks that without exploring the tacit assumptions on which these opposed positions rest we are condemned to an endless oscillation between them; and this is, of course, exactly the same as Kant's advertised attitude to his philosophical antecedents (Bxiv-xv).

In these ways McDowell acknowledges, as I have,²⁰ a strategy in Kant which identifies and rejects an assumption common to traditional philosophy, but his account of Kant differs from mine in two ways. One arises from McDowell's clinging to the traditional view that Kant subverted his own insights with a 'transcendental'²¹ appeal to a wholly independent noumenal reality (MW, 41–5, 51). For that reason McDowell believes that Kant's insight into the dependence of sense on understanding at A51/B75 was lost and replaced with a 'profoundly unsatisfactory' (MW, 45) reference to a supposed transcendent reality quite independent of our cognitive powers. McDowell accepts a traditionalist commitment in Kant to the reality of a noumenal world. I have argued that this account of Kant should be given up, and I say no more about it here.²² It may, therefore, seem as if my account of Kant, which dispenses with that traditional 'transcendent' story, places him even closer to McDowell; and that conclusion may seem to be reinforced by the additional account I offered in section II. For McDowell either did not notice, or did not acknowledge the relevance of, the passages considered there for the same Kantian strategy. If they are understood in the way I have suggested, then they strengthen the connection with McDowell's own view.

That my account of Kant does not, however, represent him as even closer to McDowell is a consequence of the second difference between his account and mine. For I believe that at a central point Kant is closer to Evans' position, which McDowell rejects. That central point is McDowell's belief that "sense and understanding cannot be *even notionally* separated" (MW, 9, 41, 51; emphasis added). McDowell formulates the two positions in the following way:

Evans, then, identifies perceptual experiences as states of the informational system possessing content which is non-conceptual (MW, 48).

According to the picture I have been recommending the content of perceptual experience is already conceptual. A judgment of experi-

ence does not introduce a new kind of content, but simply endorses the conceptual content, or some of it, that is already possessed by the experience on which it is grounded (MW, 48–9).

For McDowell, and for McDowell's Kant, the understanding has to be already incorporated into the passive deliverances of the senses, so that the idea of even a notional separation of sense from understanding is impossible. Evans, by contrast, endorses the idea of a 'non-conceptual' content due to the senses, and so makes room for a constraint on thought which has at least some independence from understanding. According to McDowell, to retain that degree of even notional independence for sense from understanding falls into the 'Myth of the Given'. To avoid that Myth it is necessary to endorse the impossibility of even a notional separation of sense and understanding.

The issue, as I shall suggest, is not quite straightforward, since McDowell's denial of even a notional separation remains obscure, but I want initially to indicate one passage in which Kant's position seems to be closer to Evans. After he has introduced the link between sense and understanding at A51/B75 Kant continues in the following way:

The understanding can intuit nothing, the senses can think nothing. Only through their union can knowledge arise. But that is no reason for confounding the contribution of either with that of the other; rather is it a strong reason for carefully separating and distinguishing the one from the other (A51–2/B75–6).

The passage indicates clearly enough that despite their intimate connection the two faculties, or their respective contributions to knowledge, have to be separated in some way. That claim seems to be at odds with McDowell since Kant plainly recommends some, even if only notional, separation of the two. Kant, moreover, has some apparatus at his disposal to begin to explain what kind of separation is necessary. For, as I indicated in summarizing Beck's paper, Kant employs his device of 'abstraction' to indicate how it is possible for us to fabricate concepts *as if* they were completely separate from any sense experience.²³ It is such a device which misleads us, in the unacceptable face of reason, into thinking that we can have cognitive access to a noumenal reality totally independent of our sense experience. Kant appeals to such a diagnosis throughout the Dialectic, and, contrary to McDowell's ascription of the 'transcendental' story to him, resolutely regards any such appeal as a fundamental mistake. It is

true that Kant also, for example in the positive appeal to transcendental freedom in the Third Antinomy, allows a regulative use of such abstracted concepts as Beck rightly points out. But that important proviso continues to be hedged about with qualifications which demonstrate Kant's appreciation of its dangers.

That it remains unclear how far that passage in Kant is in conflict with McDowell reflects some uncertainty in McDowell's own position. McDowell could not think that two distinguishable faculties, or their contributions, cannot be notionally distinguished, since that seems inconsistent. The issue turns rather on the more substantial acceptance or rejection of what Evans calls 'non-conceptual' content, so perhaps the best way to resolve the unclarity is to consider how McDowell rejects the three arguments in Evans for such non-conceptual content (*MW*, 52–65).²⁴ Those arguments rely first on the idea that non-conceptual content is finer-grained than conceptual; second that perceptual experience is 'belief independent'; and third that we share perception with animals which do not have conceptual capacities. I shall concentrate on the third case, where the appeal to non-conceptual content seems strongest, but it is useful to note one feature of the structure of McDowell's argument in all three cases. Characteristically his objections take the form of showing that Evans' arguments are not logically binding, so that McDowell's alternative position is not shown to be untenable. This is a weak conclusion which falls short of offering any positive reason for accepting either position; as far as the philosophical arguments go it may be that we have no logically binding arguments to establish either position. This shows up a characteristic limitation of philosophy in such a context, namely that we cannot establish the truth of some conclusion merely by showing that a contrary argument is unsuccessful. In this case we may be tempted to look for some relevant facts, for example in the admitted neurological and behavioral similarities between human and non-conceptual animals, but McDowell does not consider these.²⁵ Indeed he resists the intrusion of any such material in his refusal to admit as relevant any considerations from cognitive psychology at all (*MW*, 55). This seems to me inevitably to reinforce the inconclusiveness of his position.

In the third argument about animal perception McDowell offers two principal considerations both of which share the above weaknesses. The first is an objection to a corollary in Evans' account, namely a claimed parallel between the animals' informational system and the operation of what Evans calls the 'testimony system'. This first line can be set aside

on the grounds that even if Evans has made some error in drawing a parallel between the two systems, the non-conceptual informational and testimony systems, that does not show that his account of the former is unacceptable. We could drop the comparison and retain the former account.²⁶ It is true that this may leave that account short of some possible support, but, as I noted earlier, McDowell himself rules out any considerations relating to the facts of animal perception.

The second line begins with the claim that Evans' position is at least not obligatory:

If we share perception with mere animals then of course we have something in common with them. . . . But it is not compulsory to attempt to accommodate the combination of something in common and a striking difference in this factorizing way: to suppose our perceptual lives include a core that we can also recognize in the perceptual life of a mere animal and an extra ingredient in addition (*MW*, 64).

McDowell then recommends the alternative of saying that "we have what mere animals have, perceptual sensitivity to features of our environment, but we have it in a special form." The recommendation is made on the grounds that without accepting it we are forced to endorse either Davidson's coherentism or the Myth of the Given.

This line of argument is equally inconclusive and open to the following difficulties. First it offers no immediate positive ground for the recommendation, but only claims that Evans' option is not obligatory; it therefore leaves McDowell's conclusion as no more than a possible option along with Evans'. Second its central argument rejects a form of 'factorization' which remains obscure. It is unclear, for example, in what way to factor out 'perceptual sensitivity' in two different forms avoids the factorization of a common core to human and animal perception where humans have an extra ingredient. Why should not the 'special form' of our perceptual sensitivity precisely be an extra ingredient in our repertoire? The point is not helped by a bare appeal to terminological differences, for example between 'perceptual sensitivity' and 'perceptual experience' (*MW*, 114) or between inhabiting a mere 'environment' or a 'world' (*MW*, 115–18).²⁷ Finally to recommend McDowell's alternative on the ground that it alone avoids the errors of coherentism and the Myth of the Given is at least in danger of circularity. For it seemed as though the arguments against Evans' position were intended to support the claim

that the Myth of the Given was unacceptable. If it now appears that we have already to be persuaded of the errors of that position in order to support McDowell's alternative then there is a circularity.

It may be objected that these complaints, even if correct, concern only the form of McDowell's argument and do not provide positive support for Evans' alternative. That is true, but, as I suggested, McDowell's own position has exactly the same character. It is not possible here to address head-on the whole complex issue of non-conceptual content, but I end with a somewhat more positive comment on it. For the issue itself has some underlying ambiguities even in the limited context of the debate between Evans and McDowell. Evans plainly assumes that non-conceptual content is licensed by a perceptual content not describable by the subject of the perception. In the discussion, however, it sometimes seems as though what is at issue is the identification of such a perceptual content without *any* reference to concepts. In one way it is, trivially, absurd to suppose that a content can be identified without any reference to concepts. The idea of 'content' is the idea of some conceptually discriminable item which can be related to cognitive faculties such as thought, memory, or perception.²⁸ It would, therefore, be easy but unsatisfying to support McDowell's thesis that sense and understanding cannot be even notionally separated by insisting on that trivial truth. There is no possibility of identifying perceptual content without some reference to concepts, and so, in that sense, no possibility of separating sense from understanding even notionally. But the argument about animals' perception is not settled by such a consideration, for that requires, as Evans assumes, not that sense should be identified without *any* reference to concepts, but that it should be identified without reference to concepts available to the subject of perception. The two claims can be expressed in the following way:

- (1) It is impossible to identify perceptual content without using (any) concepts.
- (2) It is impossible to identify perceptual content without the subject's using concepts (or without using concepts available to the subject of perception).

McDowell offers still another account of the blindness which attaches to 'intuitions without concepts' when he says: "How should we cash out the image of blindness? To say that an experience is not blind is to say that it is intelligible to its subject as purporting to be awareness of a feature of

objective reality: as a seeming glimpse of the world" (MW, 54). This is to build into non-blindness a very heavy commitment not just to concepts but to some exceedingly abstract and even philosophically contentious concepts. It construes McDowell's central thesis in a third way as:

- (3) It is impossible to identify perceptual content without using concepts (or without the subject's using concepts) which express the relation between perception and objective reality.

It seems to me, however, that while (1) is trivially true, and perhaps philosophically uninteresting for that reason, (2) is arguably false, and (3), in some interpretations, is patently false. Even for human language users it is surely possible to identify the content of perception without expressing any relation between perception and objective reality. For non-language users the claim seems even more plausible. (3) undoubtedly contains other ambiguities which deserve closer scrutiny,²⁹ but the central problem of non-conceptual content surely involves (2) rather than (3). But nothing has yet been done to show that (2) is true, and if that is so then non-conceptual content remains a viable option.

These comments indicate that Evans' position is defensible, and that therefore there is an acceptable sense in which sense and understanding may be at least notionally separable. But that conclusion says nothing about Kant's commitment to such a position, and I therefore finally put forward additional grounds for ascribing some such position to Kant. I noted earlier that even in the A51–2/B75–6 passage Kant insists on separating the respective roles, or contributions, of the two faculties, but there are other passages which carry the same message. These are passages where Kant admits the possibility of an experience quite unlike ours, so that although certain features can be said to be necessary in our experience they could not be said to be necessary for *any* experience. The best known and most quoted of these are those where Kant refers to an 'intuitive understanding' (B138–9). Kant's conception of an 'intuitive understanding' is one "through whose representation the objects of the representation at the same time exist." Kant goes on to say that we "cannot form the least conception" of such an understanding, but I take that to mean that we cannot grasp how such an understanding might work, rather than that we cannot consistently describe such a possibility. That reference, obscure though it is admitted to be, nevertheless exemplifies a typical Kantian position in which although certain features of human experi-

ence are necessary for it we cannot conclude that those features belong to every possible experience. We do not have an intuitive understanding, and this necessitates that connection between intuitions and concepts which Kant outlines at A51–2/B75–6. We might say that the relevant features do not hold in every possible world, even though they hold in every possible world with sufficient similarity to ours. These features are not logically true of experience generally, though they are necessarily true of our experience.³⁰

That same position is exemplified in another more relevant set of claims at A89–90/B122–3. There Kant says:

The categories of understanding . . . do not represent the conditions under which objects are given in intuition. Objects may, therefore, appear to us without their being under the necessity of being related to functions of understanding. . . . For appearances can certainly be given in intuition independently of functions of the understanding. . . . Appearances might very well be so constituted that the understanding should not find them to be in accordance with the conditions of its unity. . . . But since intuition stands in no need whatsoever of the functions of thought appearances would nonetheless present objects to our intuition.

Such a passage evidently reinforces the requirement Kant insisted on at A51–2/B75–6 that sense and understanding must be properly distinguished. The passage claims explicitly that appearances can be given independently of the functions of the understanding and so endorses, contrary to McDowell, an at least notional separation of the faculties. Such a claim has nothing whatsoever to do with the one proviso which McDowell himself makes on Kant's behalf, namely that alleged commitment to an independent noumenal reality in what he calls the 'transcendental' story. Kant's claim here that sense has to be properly distinguished from, should be treated as independent of, understanding is not a claim about some transcendent reality but about the immanent reality of our own experience. It is a claim which echoes Evans' own belief in non-conceptual content, rather than McDowell's rejection of such content, and it consequently places Kant closer to Evans than to McDowell. It might be said that the penalty of such a location is that it falls into error, into a Myth of the Given, but the arguments designed to establish that penalty are at best inconclusive. Kant evidently rejected McDowell's claim that sense and understanding cannot be even notionally separated.

The resulting picture is one which endorses Beck's original account but also goes beyond it in identifying a further common assumption in the previous tradition which Kant aimed to reject. The denial of that common assumption about subjectivity is one which McDowell and I would both want, sometimes for different reasons, to ascribe to Kant. But I have argued that McDowell's further claim, that sense and understanding cannot be even notionally separated, can be neither accepted nor ascribed to Kant. It cannot be accepted in part because it remains obscure; and it cannot be ascribed to Kant because his own text points to a rejection of at least some of its non-trivial interpretations. Some advance towards McDowell can be made on the two fronts which Beck identified in Kant's war against the previous tradition, but the advance stops short of McDowell's own position and of the ascription of that position to Kant.

Notes

1. Originally published in the *Journal of the History of Ideas* 28:1967, 224–36.

2. J. McDowell, *Mind and World* (Cambridge: Harvard University Press, 1994); henceforth *MW*.

3. Kant himself does not use this popular expression in the Preface of the *Critique*, but he does describe Copernicus' 'hypothesis' as a 'change in point of view' (*Umänderung der Denkart*) associated with the attempt to revolutionize metaphysics (Bxxii and note).

4. J. McDowell, "Knowledge and the Internal," *Philosophy and Phenomenological Research* 55:1995, 877.

5. Among the works which express such views I would include H. Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983); G. Prauss, *Erscheinung bei Kant* (Berlin: Walter de Gruyter, 1971), and *Kant und das Problem der Dinge an Sich* (Bonn: Bouvier Verlag, 1974); and G. Bird, *Kant's Theory of Knowledge* (London: Routledge & Kegan Paul, 1962; henceforth *KTK*), and "Tradition and Revolution in Kant," in *Proceedings of the Eighth International Kant Congress*, ed. H. Robinson (Milwaukee: Marquette University Press, 1995), vol. 1, pt. 3, 1119–36.

6. The *Prolegomena* passage (4:304) goes: "To put all this in a nutshell it is above all necessary to remind the reader that what is at issue here is not the origin of experience but of what lies in it."

7. There has been considerable controversy over the so-called 'two worlds' or 'two aspects' accounts of things in themselves. The controversy, however, has been generated largely not by what Kant says but by the labels which commenta-

tors have given to his views. If we speak of 'two aspects', for example, the question immediately arises what these are aspects of, and this creates difficulties. There is, however, no way of avoiding Kant's own provisos about the distinction which amount to a clear recognition of the difference between our identification of items in experience and our philosophical conceptions of those items. See G. Bird, *KTK*; H. Allison, "Kant's Critique of Berkeley," *Journal of the History of Philosophy* 11:1973, 43–64, and *Kant's Transcendental Idealism*; P. Guyer, *Kant and the Claims of Knowledge* (New York: Cambridge University Press, 1987); K. Westphal, "Affinity, Idealism and Naturalism: The Stability of Cinnabar and the Possibility of Experience," *Kant-Studien* 88:1997, 139–89; and R. Meerbote, "Allison on Freedom and Idealism," *Kantian Review* 1:1997, 136–50.

8. In the Aesthetic Kant draws this distinction in a variety of ways. He contrasts rainbows and raindrops, roses and their scents, tastes, colors, sounds and heat (A26–9/B42–5), moral properties (A42–4/B60–2), the way things are and the way they seem (B69–70), and Saturn's rings and its supposed 'handles' (B70 n.) as an example of the contrast between genuine and illusory properties. This variety, some of the ways in which Kant characterizes these contrasts, and a tacit association with Locke's distinction between primary and secondary qualities, have made it difficult for commentators to recognize the significance of these contrasts.

9. In line with this account it might be expected that Kant would also mark a distinction between empirical and transcendental dualism, and he does so at A379 and A389–91.

10. In his excellent comparison between Kant and Berkeley ("Kant's Critique of Berkeley"), Allison acknowledged Kant's commitment to two senses of 'representation' for the same reasons.

11. Of course it is the claim that empirical objects are *essentially* related to our senses which encourages the ascription of a dualist subjectivism to Kant. This is to assume that there is only one way, namely the traditionalist subjective way, in which the essential relation can be expressed. Even twentieth century reductive phenomenism addressed the traditional problems although it invoked a nontraditional logic. Whether other theories, such as those involving intentionality, assume a traditional framework will depend on the status accorded to 'intentional objects'. See R. Aquila, *Matter in Mind: A Study of Kant's Transcendental Deduction* (Bloomington: Indiana University Press, 1989), and his "Unity of Apperception and the Division of Labor in the Transcendental Analytic," *Kantian Review* 1:1997, 17–52, for the latter point.

12. I have elsewhere expressed the view that Kant's account of things in themselves is not inconsistent, and that it appears so mainly because it is assumed that Kant is working within a framework of traditional dualism. See Bird, *KTK*, "Tradition and Revolution in Kant," and "Kantian Myths," *Proceedings of the Aristotelian Society* 96:1996, pt. 2, 245–51.

13. It is not completely far-fetched to see in this Kantian move a step towards a Quinean relativist ontology (W.V.O. Quine, "On What There Is," in *From a Logical Point of View* [Cambridge: Harvard University Press, 1953], 1–19, and *Ontological Relativity* [New York: Columbia University Press, 1969]). That ascription is, however, complicated by Kant's evident belief that human experience is necessarily governed by the synthetic a priori principles, for this seems to conflict with, or qualify, any such relativism. See G. Bird, "Kant and Contemporary Epistemology," *Kantian Review* 1:1997, 1–16.

14. William James makes this point in the most explicit way when he says in *The Principles of Psychology* (London: Macmillan, 1890), Preface, v-vi: "Psychology assumes as its data (1) thoughts and feelings and (2) a physical world in space and time with which they co-exist, and which (3) they know. Of course these data are discussible, but the discussion of them is called metaphysics and falls outside the scope of this book."

15. It is not my intention here to endorse the tradition, and I have elsewhere ("Tradition and Revolution in Kant") argued that Kant did not endorse it.

16. The same point is made in McDowell's "Knowledge and the Internal": "[L]et me remark that my move is not well cast as an *answer* to the skeptical challenges; it is more like a justification of a refusal to bother with them" (888 n 19.); and: "[T]he thing to do is not to answer the skeptic's challenges, but to diagnose their seeming urgency as deriving from a misguided interiorization of reason" (890).

17. D. Davidson, "Thought and Talk," and "On the Very Idea of a Conceptual Scheme," both in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), 155–70, 183–98.

18. G. Evans, *The Varieties of Reference*, ed. J. McDowell (Oxford: Clarendon Press, 1982).

19. W. Sellars' seminal articles—"Empiricism and the Philosophy of Mind," in *Science, Perception and Reality* (London: Routledge & Kegan Paul, 1963), 127–97, and "Scientific Realism or Irenic Instrumentalism," in *Philosophical Perspectives: Metaphysics and Epistemology* (Atascadero: Ridgeview, 1979), 156–89—outlined his account of the Myth of the Given, but he identifies and rejects a *number* of such myths. Among the claimed myths are: (a) the belief that reports of inner experiences, such as "X looks F to S," can be analyzed so as to quantify existentially over 'looks'; (b) the belief that philosophical accounts of inner experiences, as foundations for knowledge or for language learning, are viable; (c) the belief that reports of inner experiences have a unique priority over reports of public states of affairs; (d) the belief that sense experience uniquely determines an observational vocabulary. These beliefs may be connected, but they are all different, and may even be in conflict. For example rejection of (a) seems to deny the existence of inner experiences altogether, while rejection of (b) denies their existence only as conceived in certain philosophical theories. This suggests that

we should speak of Sellars' 'Myths of the Given' in the plural rather than of 'the Myth', and recognize that some of the alleged myths may not be mythical at all. It is, for example, much easier to accept (b) as mythical than (a).

20. I would not want to give the impression that I have derived the given account of Kant from McDowell's interpretation in the *Mind and World*. I had already given the account in *KTK*.

21. This is the term McDowell himself uses (*MW*, 41): "But Kant has also a transcendental story, and in the transcendental perspective there does seem to be an isolable contribution from receptivity. In the transcendental perspective receptivity figures as a susceptibility to the impact of a supersensible reality, a reality that is supposed to be independent of our conceptual activity in a stronger sense than any that fits the ordinary empirical world." It would have been better to use Kant's term 'transcendent'.

22. Bird, "McDowell's Kant: *Mind and World*," *Philosophy* 71:1996, 219–43.

23. It is this terminology that Kant uses to summarize his account of the dialectical errors at A619/B647, and A669–702/B697–730.

24. There has been an extensive discussion in recent times of the scope of animals' cognitive powers. Among the most influential have been Davidson's "Thought and Talk," but the issue surfaced in the Kantian context earlier in Bird, *KTK*, 127–8, in relation to the discussion of S. Hampshire's *Thought and Action* (London: Chatto and Windus, 1959).

25. For an interesting survey of empirical material relating to animals' cognitive powers, see Andrew Whiten, ed., *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading* (Oxford: Blackwell, 1991); and Marian Stamp Dawkins, *Through Our Eyes Only: The Search for Animal Consciousness* (Oxford: Oxford University Press, 1998).

26. Evans himself made a general proviso about the parallel between testimony and the senses ("Reference and Contingency," *The Monist* 62:1979, pt. 2, 123 n. 5): "I am aware that the parallel between testimony and the senses needs defense, and there are important differences . . . the senses yield non-conceptual information whereas language embodies conceptual information." This reinforces the policy of setting the argument aside, since Evans' concession is precisely the point of McDowell's objection.

27. The latter terminology of 'environment' and 'world' is taken from Hans-Georg Gadamer's *Truth and Method*, trans. J. Weinsheimer and D. Marshall (New York: Crossroads, 1992).

28. In his *Kant's Intuitionism* (Toronto: University of Toronto Press, 1995), 54–9, 223–5, Lorne Falkenstein emphasizes the difficulties of Kant's notion of 'blindness' for the account of intuitions in the Aesthetic. But Falkenstein does not explore the variety of interpretations that might be given of Kant's doctrine, and in particular does not clearly distinguish the two positions noted here under (1) and (2).

29. McDowell's account discloses other ambiguities in the relation between 'mind' and 'world' which contaminate the traditional account. Three different ways of capturing that relation are involved in the relations between: (i) Mental representations (appearances or thoughts) and external objects. (ii) Thought (understanding) and reality. (iii) Thought (understanding) and sensibility. (i) might be regarded as a special case of the relation, on the assumption that external objects are only *one* aspect of reality; (ii) might be regarded as a more general relation involving also the relation between thought and mental experiences as part of reality; and (iii) involves a relation between two mental powers typically characterized as, respectively, active or spontaneous and receptive and passive. (i) represents the classical traditional problem of epistemology within a framework of idealism; (ii) is primarily a semantic problem, and (iii) primarily a psychological issue. McDowell (*MW*, 18–20, 36–8) gives the impression that he is really concerned with (ii) rather than either (i) or (iii) in his remarks on inner sense, and on the parity between inner and outer sense which he correctly ascribes to Kant. But he also sometimes gives the impression that his principal target is the tradition captured in (i).

30. Such complex modalities need further explication, and would be in conflict with either an exhaustive positivist distinction between analytic and synthetic truths, or a Quinean rejection of that distinction. Since Kripke's more complex account of such modalities in "Naming and Necessity," in *Semantics of Natural Language*, ed. D. Davidson and G. Harman (Dordrecht: D. Reidel, 1972), 253–355, however, there is scope for such an explication of these Kantian claims. That scope has been explored in a Kantian context by G. Brittan, in *Kant's Theory of Science* (Princeton: Princeton University Press, 1978); in a linguistic context by Evans, in "Reference and Contingency"; and in a wider formal context by M. Davies and L. Humberstone, in "Two Notions of Necessity," *Philosophical Studies* 38:1980, 1–30.

Paths Traced through Reality:

Kant on Commonsense Truths

Predrag Cicovacki

I

Just as Aristotle argued that ‘being’ has many senses, Kant could have made a similar claim with regard to ‘truth’. Kant’s two conceptions of judging, determinative and reflective, suggest that we need a different conception of truth for each of them. Determinative judgments themselves are further divided by Kant into cognitive and moral. While the former deal, roughly speaking, with ‘what is’, the latter concern ‘what ought to be’, and there are clear indications in Kant that the truths of such judgments are of different kinds. There are, moreover, further distinctions within the subdivision of cognitive determinative judgments. Kant has no doubt that the truths of ordinary commonsense judgments are significantly different from the truths of scientific judgments, and even more so from our metaphysical judgments about reality.

Kant thus believes the concept of truth to be complex and rich, and that each kind of truth deserves a detailed inquiry before we can offer a systematic and comprehensive theory of truth. While the completion of this project would require an essay of larger scope than the present one, my goal here is to contribute to this project by examining Kant’s insufficiently developed yet original and deep understanding of commonsense truths.¹

Kant begins his discussion by recalling that truth is frequently taken to consist in the agreement (or harmony; *Übereinstimmung*) of our cognitions with their objects. He tells us that this is a ‘nominal’ definition of truth, a mere explication of the meaning of the word ‘truth’ (A58/B82), but he could have also said that this is the way in which we commonsensically understand truth. For example, my claim that you were at home this afternoon is true if, indeed, you were at home at that time.

Or, to take another ordinary example, my recognition that your home is comfortable and spacious is true if that is really the case.

The mentioned definition and the given examples indicate some of the major strengths and weaknesses of the commonsense understanding of truth. The truths of common sense are frequently based on the direct and uncontroversial evidence of our senses, and they emerge as a result of our practical interactions with the middle-size objects in our surroundings. I claim that you were at home because you opened the door of your house for me, shook my hand, showed me around your home, we played a chess game and afterwards had an enjoyable conversation while slowly sipping your delicious jasmine tea. When we base our claims on such uncontroversial evidence, our cognitions are right on target and it appears that we can rarely go wrong. Such commonsense truths are based on our experience and practice, and they in turn serve as a reliable foundation for future experience and practice. They can even serve as a starting point for our more complex and subtle scientific or philosophical inquiries.

There are, nevertheless, serious problems with the commonsense conception of cognition in general and truth in particular. One is that, without noticing it, we easily slip away from these reliable grounds and base our claims on dubious hearsay and hasty generalizations. As a result, reliable and dubious claims find themselves assimilated together into a common view of reality. This lack of discipline and self-criticism is where science and philosophy can improve upon common sense.

A more important problem in this context concerns the low value of commonsense explanations with respect to the nature of truth. Like a craftsman who is uncomfortable explaining what he so skillfully and routinely does with his hands every day, common sense becomes confused when challenged to account for the formation and nature of its numerous truths about reality. Common sense has a shallow and sometimes even inconsistent understanding of how, and why, it happens that our cognitions hit their targets and are true to reality. In Kantian terms, the problem for common sense consists in moving beyond the nominal and finding a 'real' definition of truth.² As a mere explication of that word, the acceptance of the nominal definition of truth does not make a commitment to any specific theory of truth. The nominal definition is compatible with various interpretations and theories, which would depend on further clarification of the terms contained in it. Thus, to move from the nominal toward a real definition of truth, or toward developing a theory of truth, we need to answer questions such as follows: How is it possible that our

cognitions agree with their objects? How should we understand this relation of '*Stimmung*' between them? How is it that cognitions could refer to objects in the first place?

When pressed against the wall, common sense will offer two lines of answers to these questions, which correspond to the two distinct senses in which we use the word 'common' in 'common sense'. In the next section we will consider the 'realist' line of response to our questions. The idea is the following: we live in a world of common sense because we live in a world of common objects. According to this view, truth depends on the way the world is; what is true is such because of what is the case, regardless of whether we perceive and are aware of it. Accordingly, commonsense cognitions are true when we are capable of grasping the reality that surrounds us. If you were at home, then that was true regardless of whether I, or anyone else, knew anything about that.

Although predominant, this is not the only commonsense explanation of what makes our cognitions true. In the blink of an eye common sense can switch from this 'realist' perspective to what can be called an 'idealist', in some cases even an utterly 'relativist', account of truth. According to this alternative view, truth depends on the way we are, that is, on the way we think, perceive, and approach reality. We live in a common world because, and only insofar as, our needs, interests, cultures, and previous experiences make it possible for us to perceive the world in the similar, or common, way. When any of these elements change, our perceptions of the world vary as well. Whether a house is comfortable and spacious depends on what we understand by that, and such an understanding varies from one culture to another, from one individual to another. The 'subjective' factors affect the way we cognize reality, and thus affect the nature of truth. In section III we will consider Kant's analysis of the relevant subjective factors of our cognition of reality.

It will turn out that a proper account of commonsense truths must involve some elements that belong to both of these opposed points of view. Following Kant, I will argue in section IV that commonsense truths depend both on how the world is and on how we think of that world. The nature of the interaction between the objective and subjective elements will be explained in terms of Kant's conception of determinative judgments and the so-called 'craftsman' metaphor. In our cognition of reality we do not simply notice and take for granted what we find in front of us. Instead, we shape and determine the material we find in accordance with our conceptual frameworks, goals, intentions, and background knowl-

edge. Like craftsmen who apply their skills to the material of nature to manufacture desired artifacts, in cognition we clear the paths that help us orient ourselves in reality. Commonsense truths are like paths traced through reality.³

II

When pressed to move beyond the nominal definition and explain how truth depends on the way the world is, a realistically oriented proponent of common sense is likely to introduce the following theses:⁴ 1) Truth consists in some sort of agreement or correspondence between our cognitions and the external reality. 2) There can be only one true and complete description of that reality. 3) Reality itself consists of a fixed and determined totality of mind-independent objects.

This kind of explanation has been defended for more than two thousand years in the Western tradition, and it is hard to deny that there is something appealing about it.⁵ It insists on the thoroughly objective character of truth and claims that since truth is the truth of reality, it must thus depend on the way reality is. Despite its initial appeal, the account has always had its opponents as well. In a nutshell, they have argued that the difficulties that emerge when we attempt to develop and defend this explanation are so great that we are better off in rejecting it altogether. In this section we will try to peel off the different layers of this account and distinguish between what in it should be preserved and what is untenable.

The first problem with words like 'agreement' and 'correspondence' is that they are both vague and ambiguous.⁶ It is difficult, if at all possible, to assign them a precise meaning and still preserve the initial intuitive appeal of the broader account. 'Agreement' and 'correspondence' are taken to refer to some kind of 'fit' and 'congruence' between our cognitions and the states of affairs to which they refer. But what fit and what congruence? Could we not move beyond this pictorial and metaphorical level in our account of the nature of truth? Moreover, how could there be any fit or congruence between elements so heterogeneous, i.e., between something mental, such as our cognitions, and something physical, such as objects to which our cognitions are supposed to refer? 'Agreement', 'correspondence', 'fit', and 'congruence' all seem to lean on the relation of reference, yet this relation itself demands an account. The nature of reference is frequently explained in causal terms,⁷ but this view faces serious

problems. Even if the causal aspect is necessary, it is not sufficient for the reference relation. Reality by itself does not determine for what our words stand. A sign-relation is not built into reality, and we cannot simply assume that reality consists of self-identifying objects. The identification of ordinary objects around us is always dependent on our frames of reference which involve, among other things, pragmatic considerations and decisions. For instance, if the liquid we drink every day continues to have all the functional properties it used to have in the past, we are going to continue calling it 'water', even if its chemical analysis turns out to be different than what every schoolboy now knows as H_2O (cf. A728/B756). There are, moreover, problems with the role of causality in the reference relation as well. Think about generalizations; when it is claimed that an average American family has 3.5 children, what could this number causally refer to? In what kind of correspondence could it stand with any one real family? Furthermore, opponents of the correspondence account of truth and the causal account of reference are quick to point out that this relation of reference is problematic not only in the case of affirmative claims, but even more so in the case of negative claims and counterfactuals; it is simply difficult to conceive what causal and referential relations take place when we formulate negative and counterfactual claims. It would be premature to claim that these problems are insurmountable. Yet they are serious and require significant and detailed elaboration of the first thesis.

While the first thesis certainly needs better defense, the second is dubious and untenable. It claims that there can be only one true and complete description of reality. There are, no doubt, cases which appear as clear-cut as the second thesis suggests. You either were or were not at home at the particular time; our chess game either ended in your victory, or in my victory, or in a draw; we either had jasmine tea or we did not. Nevertheless, there are as many cases that are not so straightforward and decisive. Not only our scientific and philosophical practice, but even our commonsense experience suggests that there are alternative, yet equally correct, descriptions of reality. Consider the following examples. You and I may disagree with respect to whether your house is spacious and comfortable and yet both be correct. We could both be correct because we use different standards when estimating what counts as spacious and comfortable. Our standards may be different because of our previous experiences; you grew up in a small family that owned a mansion and I was surrounded by numerous siblings in a crowded apartment. Moreover, even you yourself can correctly offer opposing appraisals of your house in

different contexts. To your parents, who are accustomed to a much larger space, you can complain how small this house is. Or, with different intentions, you can also proudly announce that your present home is much larger and more comfortable than the apartment in which you lived before.

These examples show that our descriptions and evaluations of a given situation depend on our previous experience, and our present intentions and goals. As I will argue in section III, they also depend on the nature of the cognitive framework with which we approach the world. The same state of affairs can be conveyed by different, yet nonsynonymous, truth claims. This is possible because they describe reality from different points of view.

It is even more doubtful whether we can give a complete description of any situation. Think about describing a game of chess. While it is clear what it means to say that you and I played a game (given that it is true that we did so), what would a complete description involve? Certainly more than the outcome of the game, and the time and place where it was played. But how much more is needed for a complete account? The actual moves of the game? An analysis of the game, with a list of good moves, bad moves, missed opportunities for both sides, and an explanation of the main strategic ideas? But then why not the psychological states of the players, the relative significance of the game for our friendship, and perhaps also the past history of our games? And would not a complete description require mention of the spectators, what they thought of the game, and so on?⁸ It is clear that the list can continue *ad infinitum*. Giving a complete description of any situation resembles the Leibnizian attempt to provide a complete concept. This task would require knowledge of the present, past, and even future, a mission impossible for finite minds.

The second thesis is unnecessarily strong. There can be more than one true description of the relevant situation, and no complete description is possible. Fortunately, commonsense experience shows that a complete description is never really needed. On the one hand, the contextual parameters determine what must be described, i.e., what information is required for making an appropriate and understandable truth claim.⁹ Describing our chess game to an experienced player and describing it to someone who does not even know the rules of the game would involve quite different elements of the situation, as well as significantly dissimilar language. On the other hand, even our ordinary language includes talk about 'approximate' truths, 'partial' truths, and so on. We would not seem to be able to accommodate the widely accepted view that there are ap-

proximate truths—indeed, are not most of our truths approximate?—if we demand a complete description of reality. Why, then, would anyone expect that there can be only one true and complete description of reality? The second thesis finds its motivation, and at least partial support, in the third thesis; namely, that there is one true and complete description of reality because reality itself is one and completely determined, regardless of what we know or think about it. This third thesis, then, is the one in which this account of truth finds its ultimate support. Yet if this last thesis is significantly challenged, this will cast doubt over the entire account.

The third thesis is problematic as well. One reason for this is that it is based on an equivocation of the words such as ‘outside us’ and ‘external to the mind’ on the one hand, and ‘mind-independent’ on the other. The world that surrounds us and that we are trying to grasp is, in at least one sense of these words, outside of the mind and external to it. Yet, as Kant shows, to argue that this reality is thereby unrelated to the mind is to jump from commonsense experience to a metaphysical level (A373–5); it is to shift from ‘empirical’ to ‘transcendental’ realism (A264/B320).¹⁰

Of course, in principle there is nothing wrong with making metaphysical claims; the questions to ask are whether we are entitled to such claims and whether they are needed in the particular context. And common sense is neither entitled to such claims, nor does it need any.¹¹ The existence of the external world, the world of commonly perceivable and recognizable objects and situations, can be taken for granted without thereby leaving the grounds of common sense. Our common practice and experience entitle us to accept the existence of the outside world, without postulating the complete mind-independence of the world, or its ontological priority over our thinking about it. What is more, our common practice and experience do not entitle us to postulate the ontological independence and priority of the world. To do that is to assume an Olympian ‘God’s eye’ point of view, which is cognitively closed to us.

In Kant’s terms, partisans of the correspondence account of truth and metaphysical realism face the so-called ‘*dialelle*’ (*Durcheinander*) problem. This problem, known since the time of ancient sophists but also important in early modern philosophy and the early stages of logical positivism, is the following: if truth is conceived as a relation of correspondence between cognitions and independently existing objects (in the sense of metaphysical realism, i.e., as things in themselves), we are not in a position to judge and determine, in any non-circular way, whether our cognitions are indeed related to these objects in the appropriate way. Since

we cannot step outside our cognitive point of view, any effort to confirm the correspondence between our cognitions and the independently existent objects could only lead to further cognitions. In that case, our cognitions can only have a meaningful relation to other cognitions, but not to objects themselves. Cognitions can only verify, and in turn be verified by, other cognitions. If that is really so, there seems to be no rational basis for claiming that truth really consists in correspondence with metaphysically independent objects.¹²

There may be more land beyond the Pillars of Hercules (A396), but common sense is ill-equipped to pursue it. Fortunately, there is no need for common sense to leave its firm ground, its own land of truth (A235/B295). For the purposes of our commonsense orientation in the world, we do not need to postulate the ontological priority of the world over the mind, of the object over the subject. Commonsense truths are truths about the world of our experiences and practices; they are human truths. They emerge only insofar as we are situated in the world and interact with it. What we have to consider next is how the way we are makes an impact on the nature of commonsense truths.

III

The human mind has a tendency to charge from one extreme to another. When we recognize that the pictorial relation of 'correspondence' does not sufficiently clarify the nature of truth, we then tend to turn toward explaining truth in terms of 'coherence' or 'usefulness'. When the ideal of one true and complete description of reality is found inappropriate, we are quick to shift toward the relativistic position that 'everything goes' and that what matters is not truth but what passes as such. When the ontological priority of the object is questioned, we incline toward the priority of the subject. Instead of the metaphysical view in which the mind is secondary to a foreign, already determined world, we end up asserting that the only certain thing is the mind, and that—for all we know—the rest of the world may be just an apparition of that mind trapped in a mad scientist's vat.

The trap of extremism could be avoided by carefully examining the way we relate to the world in our ordinary cognitive experience. This is where Kant proves to be an invaluable ally. He helps us to understand that we bring to cognition a colorful package of subjective elements that, at

least partially, accounts for how, while apparently gazing at the same object, two observers can see different things. One of Kant's many claims to greatness resides in the recognition that in the variety of subjective elements there must be something we all share, for otherwise even meaningful disagreement would not be possible. To find out what that common element is, Kant turns his attention to the structural components of cognition. He realized, more clearly than anyone else before him, that our cognition is not simply a passive reaction to the immediate environment, but an intentional and thoughtful effort to orient ourselves in the world:

For truth or illusion are not in the object, insofar as it is intuited, but in the judgment about it, insofar as it is thought. It is therefore correct to say that the senses do not err—not because they always judge correctly but because they do not judge at all. Thus, truth and error, and consequently also illusion as leading to error, are only to be found in the judgment, i.e., only in the relation of the object to our understanding (A294/B350).

The senses do not judge at all, argues Kant. But what does it mean to judge? And what is the final product of judging? Is it a sentence? Or a proposition? Or is a judgment something entirely different? Kant is never explicit on these issues, but we can be fairly sure that he would distinguish between a sentence and a judgment, and even between a proposition and a judgment. If a sentence is taken to be a syntactically correct and meaningful linguistic entity, a proposition can be understood as what the sentence expresses, i.e., as a statement of thought in words.¹³ A judgment is even more than that, insofar as it adds to the proposition in question the mind's decision with respect to its truth or falsity. The proposition may be asserted or denied, or the mind may suspend judgment with respect to it.

Most contemporary analytic philosophers consider the nature of truth in terms of the relation between *linguistic* entities (sentences or propositions) and the world. Kant, by contrast, sees this relation as one between the judging subject and the world. Essential for Kant's approach is his view that, in cognition, the subject is able to integrate and harmonize perceptual and rational elements of his experience (A51/B75). This ability Kant calls 'judging' (*Urteilkraft, judicium*). In the most general terms, judging consists in making a claim and taking a stand with regard to what

we experience. Sound commonsense judging is normally based on much personal experience, on the experience of other people around us, as well as on the experience of past generations. Yet although we all have it, judging is also a “peculiar talent” that cannot be taught but can only be exercised. While in trivial matters and obvious cases the ways in which we exercise this ability need not seem to vary, our life experience is sufficiently rich and complex to show that such differences are present everywhere. Kant takes the opportunity to emphasize that judging is “a natural gift,” “a specific quality of so-called mother wit (*Mutterwitz*), for whose lack no school can compensate” (A133/B172).¹⁴

Kant is in full agreement with the Aristotelian-Scholastic tradition and early modern philosophers, such as Descartes, that truth is essentially the product of judgments. Where he deviates from that tradition is in his specific understanding of the nature of judging. In Kant's view, judging is the ability of subsuming a particular under a universal.¹⁵ By a ‘particular’ Kant means that which is, or could be, intuited as present ‘here’ and ‘now’. This could be a particular individual thing, or it could be a characteristic of an individual thing. By a ‘universal’ Kant has in mind a general (or typical) representation, such as concept or idea, which is applicable to many instances (or tokens) and thus can serve as a rule (A132/B171). Kant's idea is to show that by means of judging we overcome the privacy and idiosyncrasy of the given and turn them into communicable and intersubjective experience.¹⁶ In judging, we make the transition from an isolated ‘here’ to many remote ‘there’s, from an ever-present ‘now’ to the past or even not yet existing ‘then’. Thus, we are not arrested in the immediate environment but—by means of judging—live in an all-encompassing world.

Kant argues that, as human beings, we judge and orient ourselves in the world in two different, although complementary, ways. The more obvious one is in terms of the ‘determinative’ (*bestimmende*) judgments, and the other is by means of ‘reflective’ (*reflektierende*) judgments. In the former case we subsume a given particular under an already given universal. In the latter case an appropriate universal is not given once and forever, but is to be found always anew. Kant takes it that reflective judgments deal with (subjective and objective) purposiveness, and the relevant examples can be found among aesthetic and teleological judgments. Kant also divides determinative judgments into two kinds. What we determine by means of determinative judgments is either ‘what is’ or ‘what ought to be’. In the former case the determination is based on cognitive

concepts and in the latter on moral ones. We will here ignore the question of the truth of moral judgments and focus on the nature of determinative cognitive judgments.

Kant's account of such judgments is as follows. In commonsense cognitive experience we deal with the world insofar as it appears to us. Since we want to determine what is, that which appears provides indispensable material for our cognition. Kant defines that which appears in a perceptual encounter, i.e., an appearance, as "an undetermined object (*unbestimmte Gegenstand*) of an empirical intuition" (A20/B34). Cognition, by contrast, "consists in determinate reference (*bestimmten Beziehung*) of given representations to an object" (B137).¹⁷ That an appearance is an undetermined object means that what is given in perceptual encounters is not a ready made object which we then simply register and cognize. However simple and immediate our cognitive experience may appear in the view of common sense, the story is significantly more complicated than that. In reality, what is given is not yet determined or—better yet—it is insufficiently determined (or underdetermined). What is given leaves choices open and can be seen (determined) in different ways. What is given presents a challenge.

How, then, do we react to this challenge? How do we, starting from underdetermined perceptual objects, form determinative cognitive judgments? To clarify this, consider the following example. Suppose that, after our finished chess game, I state: "These pieces are heavy." What is it that we do when we make a judgment like that? First, we are orienting ourselves, and our audience, towards the things we are talking about. Our judgments narrow down the perceptual field and *direct attention* towards some specific things or occurrences.¹⁸ In the given judgment, we are not talking about the weight of the chess board, nor the table on which the pieces are positioned, nor are we talking about the tea cups that stand next to the chess pieces. Second, we *identify* what our attention is focused on in terms of what Kant calls 'categorical' concepts (A80/B106). In our particular case, the pieces are categorially identified as 'many' (i.e., as 'plurality', in terms of the categories of quantity), as 'real' (among the categories of quality), as 'substances' with properties (among the categories of relation), and as 'existent' (among the categories of modality). That they are 'substances' means, for example, that they are the bearers of properties and relations, rather than properties and relations themselves, that they are permanent objects which are not going to disappear in the next moment, and so on. Third, by means of empirical concepts we *iden-*

tify these substances as belonging to a certain type (kind) of objects, i.e., chess pieces. Any categorial determination still leaves open a range of possible empirical determinations. Thus, had our intentions been different, these 'substances' could have been empirically identified not as chess pieces but as pieces of wood, or as pieces of different shapes and color, or as toys that my children like to play with and throw around. Fourth, by asserting that the pieces are heavy, we narrow down and limit the empirical determination to one important *distinguishing* property of these objects. We assert that these chess pieces have a definite characteristic, and leave everything else out: whether they are new or old, expensive or not, big or small.

This account has crucial implications for our understanding of the nature of the truth of determinative cognitive judgments of common sense. Such judgments are determinative insofar as they purport to establish what is the case (and what is not the case). Since establishing what is the case consists in identifying what is the case, truth consists in identifying what is the case. Hence, a determinative commonsense judgment is true when it identifies the object of inquiry for what it is. Such a judgment is false when, in the attempt to identify what is the case, we misidentify (and mis-take) things for what they are not. Identifying what something is normally involves establishing its limits and differentiating it from other things. Thus, making determinative commonsense judgments means being involved in the processes of identifying what something is and differentiating it from other things or events.¹⁹

What implications does this understanding of commonsense truths have for our previous discussion? As we have seen, the realist version of common sense relies heavily on the idea that truth depends on reference to the external world, and that such reference can exist and be established without any contribution by the cognizing subject. By contrast, the idealist version of common sense points out that there is no such thing as 'pure ostension'. According to this view, reference is impossible without some 'pragmatic' decisions with respect to our intentions and goals, as well as without taking into account our social institutions and practices. What matters for truth is not reference but meaning.

Kant tallies up the strengths and weaknesses of both sides and identifies their underlying assumptions. His central insight is that subjectivity and objectivity, as well as meaning and reference, are not mutually exclusive, but rather complement and permeate each other. Insofar as our commonsense judgments refer to objects, insofar as they are 'objectively

real', they are not some 'subject-less' and 'perspective-less' eternal sentences, but human thoughts and judgments, i.e., thoughts and judgments entertained by human beings situated in the world (A156/B195–6). Furthermore, our thoughts and judgments do not simply refer to external objects because we are causally affected by them. Cognitive reference to objects or events in the world is possible only within the framework of our practices, experiences, concepts, and intentions (A247/B304). Even if this need not be the case with mathematical or logical truths, commonsense truths are distinctively human.

In this point Kant agrees with the idealist side of common sense, and this misleads many commentators into arguing that he holds a version of the coherence theory of truth.²⁰ Yet it is essential to remember that Kant argues that this idealist side goes too far in overemphasizing the significance of meaning and in unjustifiably disregarding the relevance of reference. 'Meaning' is simply broader than 'truth'. Put differently, while it is indispensable that our thoughts are coherently integrated into larger formations of judgments and beliefs, and while it is important to take into account our intentions and goals, the coherentist and pragmatist elements do not suffice for truth. Despite my judgment being internally coherent with the rest of my beliefs, it may still be false. No matter how many people agree on a certain point, they can all be mistaken. Regardless of how useful something may be to believe, it may still be a mistake to do so. Kant is categorical in considering cognition and judging as normative activities, as rational activities that aim at truth. And the truth of our determinative cognitive judgments is the truth of reality.

Although indispensable, the elements contributed by the cognizing subject do not suffice for truth. This becomes especially clear when Kant considers the relevance of the conceptual apparatus in general, and of the categories in particular, for an account of the nature of cognition and truth. By means of concepts, argues Kant, we are capable of making 'objectively valid' judgments (A93–4/B125–6).²¹ Such judgments are communicable and corrigible intersubjective assertions about the world. Yet 'objectively valid' for Kant does not mean 'true'; what he means by it is 'capable of having a determinable truth value'.²² As Lewis White Beck masterfully clarifies Kant's point, "the categories do not differentiate veridical from non-veridical experience; they make the difference between dumbly facing chaos without even knowing it—'less even than a dream'—and telling a connected story, even if it is false."²³ The role of the categories, then, is not to establish the truth of our judgments, but instead to

impose the conditions that must be satisfied if our judgments are to be fitted for truth. They impose the conditions that must be satisfied if our judgments are to have a determinable truth value: "The categories lead to truth, i.e., to the agreement of our concepts with the objects" (A642/B670), but they cannot guarantee truth. Whether a categorially determined and objectively valid judgment is true or false does not depend solely on our concepts but also on the actual empirical conditions. Kant thus agrees with the 'realist' side of common sense that "possible experience is that which can alone give reality to our concepts; in its absence a concept is a mere idea, without truth, that is, without relation to any object" (A489/B517).

Despite his occasionally ambiguous language, Kant's intention clearly is to bring together the elements of both realist and idealist positions into a unified account of truth. He offers various hints and suggestions that point in this direction, but little beyond that. Kant does not provide a full account of their integration and never offers a complete elucidation of the nature of commonsense truths; our task in the next section is to fill in these lacunae.

IV

One of the deepest 'truth traps' in past thinking about the nature of truth has been to assume that truth must depend either on the way the world is, or on the way we are and think about the world. But why would that be so? Since the relation of 'being dependent' admits of degrees, in principle it is possible that truth depends both on the way the world is and on the way we are and think about it. More than with respect to any other kind of truths, this must be the case with commonsense truths. What makes the world in which we live common, and what makes common sense 'common', is the integration and interdependence of the subjective and the objective elements (cf. A51/B75). There could be no clearer case than that of commonsense truths where such integration and interdependence are more indispensable and visible.

Our problem now is to move beyond the initial plausibility of this thesis and further articulate and support the claim that truth presupposes an interaction of our cognitions with their objects. Let us begin by showing that, with respect to this point, all traditional theories of truth are one-sided and inadequate. It is not so much that they completely miss the

nature of commonsense truths; it is rather that they capture only a few of the relevant aspects and disregard all others. For instance, correspondence theories correctly emphasize that truth depends on the way the world is. But they mistakenly separate man from the world, and our thinking and judging from their objects. As a consequence, they try to identify truth in terms of what could be characterized as a dubious ‘pictorial’ or ‘geometrical’ congruence between our cognitions and their objects. Coherence theories, by contrast, correctly emphasize the relevance of our conceptual apparatus and background knowledge. Yet they inflate the relevance of the subjective factors and underestimate the degree to which truth depends on the way the world is. As a result, they sever our ties with reality and make our true judgments appear belong to a consistent but—for all we can say—perhaps fictional story. Pragmatists correctly underline the functional role of truth, i.e., its connectedness to our needs, intentions and goals, and its relevance for our practical orientation in the world. But, as we shall see later, they tend to ignore some of the constraints on the side of the object.

What are these constraints about which we are talking? Roughly speaking, they are the subjective and objective conditions that, on the one hand, create the possibility of objective truth value and, on the other hand, impose some limitations on what is true or false. Kant’s considerations in the *Transcendental Analytic* and the *Transcendental Dialectic* make it is easier to reconstruct and identify the constraints on the side of the subject.

(S1) The first of them concerns the plasticity, or flexibility, of the subject. It is manifested by the degree of fluidity or rigidity of the subject’s goals, intentions, and expectations.²⁴ Rigidly defined expectations and goals could blind us to certain aspects of the presented situation. Fluid and flexible expectations and goals make us open-minded to unexpected things.

(S2) Further, there is the question of the respective simplicity or complexity of the subject. It is not measured by the number of components or parts involved, but by the level of complexity and sophistication of the background knowledge. A layman entering the lab for the first time would not grasp as much as a trained physicist who works there; an amateur chess player would not recognize a certain pattern of pieces as the Sicilian Defense; a person not familiar with chess would not recognize a certain pattern of pieces as a checkmate.

(S3) Finally, there are variables having to do with the availability and structure of the cognitive apparatus. The way that our senses are struc-

tured makes certain dimensions of the observed reality accessible, and some others inaccessible to us. The nature of our intellectual abilities similarly opens some vistas and closes off others. In Kant's terms, the kind of the intellect human beings have—at least as far as the cognition of the world around us is concerned—is in between the *intellectus ectypi* and the *intellectus archetypi*. Our intellect is not intuitive but discursive; we have no intellectual intuition but must rely on a conceptualization of the material provided by the senses.²⁵

What about the constraints on the side of the object? This question appears far more troublesome for an interpreter of Kant. All subjective constraints deal—in one way or another—with the elements of form that are imposed on the material of the senses in the process of cognitive synthesis and judging. The senses provide the needed cognitive material which, when properly formed, leads to cognitive contents and judgments. The material provided by the senses, the underdetermined objects of empirical intuition, have the potential to be determined in various ways. But these underdetermined objects must also impose some limitations on the possibilities for the formation of objectively valid judgments.²⁶ What kind of limitations could Kant have in mind? He does not help us much here, and Kant scholars find themselves frequently wondering what his response could be.²⁷ Here is a reply based on some of the previous considerations.

(O1) One of the constraints concerns the plasticity, or the level of underdetermination, of the observed objects or events. A simple curved line is normally more plastic than an equally simple straight line. The shape of a cloud is more plastic than the shape of a square. All objects and events have their own specific degree of plasticity which functions as a limiting factor in our attempts to perceive and grasp those objects and events.

(O2) Furthermore, there are constraints that deal with the respective simplicity or complexity of objects or events we try to grasp in our cognitions. Objects and events contain more or less components. A geometrical figure involves more components than a straight line; a chess board involves more components than one of its squares. More complex objects and events will normally create more resistance to our attempts to grasp and illuminate them.

(O3) There is also a level of relative accessibility or inaccessibility of observed objects or events. It is easier to grasp one billiard ball hitting another, than the action on a football field; it is easier to grasp a straight-

forward opening move in chess than one that starts a complex combination. Without attempting to postulate here some invisible essences of things, it may be that there are layers of reality that are inaccessible to all of our cognitive advances.²⁸ Without elaborating on this thesis, and without determining whether this may be so because of the nature of objects or the nature of our own limitations, let us nevertheless keep it in mind, for it is important for Kant's discussion of things in themselves.

Of these six constraints, advocates of correspondence theories are mostly concerned with the respective simplicity or complexity of perceived objects (O2), less frequently with the respective accessibility of presented objects (O3), and tend to overlook all others. They rely too heavily on a belief in the existence of an independent reality that is made up of its own component entities, together with inherent properties and structural interconnectedness. Since they take it that reality is so definitely structured, they try to convince us that in cognition we disclose and faithfully copy the objectively determined situation. Partisans of correspondence theories thus make it appear as if truth is a one-way street, a unilateral relation of objects to subjects.

Advocates of coherence theories, by contrast, recognize and emphasize the relevance of the second kind of constraints on the side of the subject (S2), and in some cases also the plasticity of the object (O1) and its relative inaccessibility (O3). They rely on the idea that truth consists in the interconnectedness of our thoughts and cognitions in a unified and internally coherent system. An important shortcoming of coherence theories is that they, like correspondence theories, also make truth look like a one-way relation. The difference is that the direction is reversed in comparison to correspondence theories.

Pragmatism is based on a more dynamic understanding of commonsense cognitions and truths than the other two kinds of theories. The subject and the object are not treated as separated by a gulf that can never be fully overcome. In Dewey's words, "Knowing is one kind of interaction which goes on in the world."²⁹ It is thus not surprising that pragmatists come closer than correspondence and coherence theorists to a recognition of the interactive nature of commonsense truths. As James argued, "The truth of an idea is not a stagnant property inherent in it. Truth *happens* to an idea. It becomes true, is *made* true by events."³⁰

Pragmatists clearly recognize the constraints on the side of the subject, especially (S1) and (S2). Our inquiry, our pursuit of truth, is an active human manipulation of the world in which we live, a manipulation

guided by our needs and interests. Moreover, our inquiry is not an individual project; it is a common enterprise in which cooperation with other human beings is of vital importance. This joint action is based not only on common needs and interests, but also on our previous knowledge and truths. Our various thoughts and insights are closely related to other cognitions and truths; they feed on each other and lean on each other. To quote James again, "Truth lives, in fact, for the most part on the credit system. Our thoughts and beliefs 'pass', so long as nothing challenges them, just as bank-notes pass as long as nobody refuses them."³¹

Pragmatists also recognize the plasticity on the side of the object (O1). Cognition is for them an active intervention, an active manipulation of our world, which allows us to grasp and shape it in multiple ways. In fact, one of the fundamental weaknesses of the pragmatist conception of cognition and truth is that it is based on belief in an almost unlimited plasticity of the environment. Yet the reality with which we interact is not in a state of constant flux, and not everything goes. Reality has limits to its flexibility; it has its own constraints. We can change and use that reality in accordance with our needs and interests, but that manipulation has its limits. When, for instance, a cabinetmaker works on different kinds of wood, his craft is constrained not only by his own, or his customers', needs and interests, but also by the natural properties of wood. To be a good craftsman, one must be able to answer and respond to the properties of the material one is working with. To be a good craftsman, one must make sound judgments.

As genuine products require the interaction of our skills, goals, and needs with the natural constraints of the involved material, commonsense truths require their interaction and integration as well. In cognizing the world and making judgments about it, we are like craftsmen who must integrate the subjective and the objective elements of the situation. Truths emerge as the result of the interaction of these elements, and do not consist in any naively conceived disclosure of a previously structured reality. This is one of the fundamental differences between transcendental realism and Kant's critical idealism, and also between Plato's Divine Craftsman and Kant's human craftsman. In Plato's story, the integration of matter and form is performed by the Divine Craftsman, and the task of the human cognizing subject is to discern and disclose the result of the already performed interaction.³² For Kant, by contrast, the integration of matter and form on the one hand, and the grasp of truth on the other, are no more than two different aspects of the same cognitive and judging activity.

The realization of our active engagement in cognition may prevent us from talking about our judgments as being ‘true of the world’, if that is taken to mean a uniquely true and complete description of the given situation. This does not imply, however, that our judgments are mere constructions. What is important in the craft metaphor and the interactive understanding of truth is that our judgments must be ‘true to the world’.³³ They can be true to the world not by disclosing it and copying it, but by being faithful to it, by taking its own constraints into account. In cognition we are like craftsmen who interact with the world in which we live. The results of such interactions are determinative cognitive judgments; in judging, we clear the paths through the world that surrounds us and stamp them down. Our judgments lead to new practices, and those practices in turn lead to new judgments.

Let us now sum up the results of our inquiry. Although Kant’s conception of commonsense truths is objectivistic in its orientation, it is not captured by any of the traditional theories of truth. It presents Kant’s attempt to move beyond the nominal definition of truth in terms of agreement between cognitions and objects, and brings us closer to a real definition of truth. When Kant considers the nature of commonsense cognitions and truths, his discussion points toward the interactive aspect of judging and truth. He thereby rejects with equal vigor a theocentric conception of cognition and truth on the one hand, and a relativistic one on the other. Kant argues in favor of an anthropocentric model. Our cognitive experience does not consist in a mere reaction to what happens in a detached world. We are not separated from the world but situated in it. Our commonsense truths are neither mere reflections of a predetermined reality, nor are they ways of world-making. Commonsense truths are paths traced through reality, and they emerge as a result of our interactions with that reality.³⁴

Notes

1. I undertake a detailed examination of Kant’s various conceptions of truth in my unpublished manuscript “Truth and Illusion: Reflections on Kant’s Philosophy.” In emphasizing the complexity of the concept of truth, Kant differs significantly from currently popular deflationist and disquotationalist views on truth; for a critical examination of these views from a Kantian point of view, see my “Rethinking the Concept of Truth: A Critique of Deflationism,” in *Truth and Its Nature (If Any)*, ed. J. Peregrin (Dordrecht: Kluwer, 1999), 203–21.

2. By 'real definition' Kant means an account of how something is possible (cf. A240 and A731/B759). Real definitions are accomplished, if at all, only at the end of a philosophical inquiry, not at its beginning. For that reason they may turn out to be quite different from the initially accepted nominal definitions. For a fuller account of Kant's view on nominal and real definitions, see L.W. Beck, "Kant's Theory of Definition," in *Studies in the Philosophy of Kant* (Indianapolis: Bobbs-Merrill, 1965), 61–73.

3. I borrow this metaphor from H. Bergson; see *The Creative Mind*, trans. M.L. Anderson (New York: Philosophical Library, 1946), 217.

4. In this point I follow H. Putnam; see his *Realism and Reason* (Cambridge: Cambridge University Press, 1983), 30.

5. For some of the most prominent versions of the correspondence theory of truth in this century, see L. Wittgenstein, *Tractatus Logico-Philosophicus*, trans. C.K. Ogden (London: Routledge & Kegan Paul, 1922); G.E. Moore, *Some Main Problems in Philosophy* (New York: Macmillan, 1953); and M. Devitt, *Realism and Truth* (Princeton: Princeton University Press, 1984).

6. 'Agreement' could mean: 1. concord, harmony, conformity; 2. unity of opinion or sentiment; 3. resemblance, conformity, similitude; 4. an understanding or arrangement between two or more people, and so on. 'Correspondence', in turn, varies in meaning between the following: 1. relation, fitness, conformity, mutual adaptation of one thing to another; 2. similarity, analogy; or 3. communication by exchange of letters. More persistently than anyone else, Putnam has argued that the word 'correspondence' is empty and noninformative. If it could be articulated in a philosophically fruitful way, that could only be done in terms of truth. Thus Putnam claims that, "Philosophically it is correspondence that depends on truth, not *vice versa*"; *Realism and Reason*, 277. For an important criticism of the correspondence (and indeed all other) theories of truth, see D. Davidson, "The Structure and Content of Truth," *Journal of Philosophy* 87:1990, 279–328.

7. See, for instance, S. Kripke's *Naming and Necessity* (Cambridge: Harvard University Press, 1972). For a criticism of Kripke's causal theory of reference, see E. Pivcevic, *What is Truth?* (Hunts: Ashgate Publishing Limited, 1997), esp. 73–5. Although Kant never developed his own theory of reference, any attempt to reconstruct such a theory based on Kant's text would have to take into account his claim that "Thought is the act of referring given intuitions to an object" (A247/B304).

8. Would not a complete description require the specification of the rules of chess, and perhaps even the history of the game, with an account of how we came to play the game with the presently accepted rules?

9. H. Dreyfus has correctly argued that a human situation, with all the contextual parameters involved, should not be confused with a physical state, or with a set of facts. There is no type-type correlation between human situations on the one hand, and physical states or sets of facts on the other. What we are trying to describe in our commonsense judgments are human situations, not physical states

of affairs, and those situations involve complicated social and institutional relations that cannot be overlooked; *What Computers Still Can't Do* (Cambridge: MIT Press, 1992), 213–20.

10. For a valuable discussion of these points of view, see H. Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983), chapters 1–2.

11. Kant insists that sound common sense is “a great gift from God” when “applied immediately to experience,” but dangerous and damaging in metaphysical waters. In metaphysics “sound common sense has no right to judge at all”; *Prolegomena* (4:259–60). See also A743/B771 and A783/B811.

12. For Kant's view on this problem, see A57–8/B82, A367–97, and various versions of Kant's logic lectures: e.g., 9:50, 16:251, and 24:386–7. For a sympathetic discussion of Kant's view, see Allison, *op. cit.*, 14–34. For further discussion and opposing assessments of the strength of this problem, see B. Allen, *Truth in Philosophy* (Cambridge: Harvard University Press, 1993), 27, 35–6, and D. Davidson, *op. cit.*, 302–5.

13. Not every sentence expresses a proposition; for example, a prayer is a sentence but it does not express a proposition, nor is it true or false.

14. “Deficiency in judging,” claims Kant, “is ordinarily called stupidity, and for a such a failing there is no remedy” (A134/B172). Kant argues that such a deficiency exists both among ignorant and learned people.

15. For further discussion, see *Critique of Judgment*, 5:179–81; *First Introduction*, 20:208–16; and *Anthropology*, 7:197–200. In Schopenhauer's interpretation of Kant, the *Urteilkraft* consists in the ability “to lay down and fix in appropriate concepts for reflection what is known through perception, so that, firstly, what is common to many real objects is thought through *one* concept, and secondly, their points of difference are thought through just as many concepts. . . . From this what is different is known and thought as different, in spite of a partial agreement; and what is identical is known and thought as identical, in spite of a partial difference, all according to the purpose and consideration that actually exists in each case”; *The World as Will and Representation*, trans. E.F.J. Payne (New York: Dover, 1969), vol. 1, 65.

16. Thus the application of concepts turns tokens into types.

17. *Bestimmen* means to ‘determine’, ‘decide’, even to ‘define’; *unbestimmt* means ‘indeterminate’, ‘indefinite’, ‘vague’, and ‘undecided’. This does not show, however, that an *unbestimmte Gegenstand* is chaotic and without discernible differences. It only means that an object of empirical intuition is available to us as a potentiality of a certain kind, a potentiality to which a form has to be imposed in accordance with the internal potentialities and boundaries of that which is given. To determine an object of empirical intuition is to limit it, to establish its limits and thereby differentiate it from other things.

18. For further discussion of this point, see R. Campbell, *Truth and Historicity* (Oxford: Clarendon Press, 1992), 420–1.

19. I take this to be the proper Kantian rendering of Aristotle's famous dictum in *Metaphysics* that "To say that what-is (*to on*) is not, or that what-is-not (*to me on*) is, is false; but to say that what-is is, and that what-is-not is not, is true" (1011b26–7; see also 1017a31–5). This may also be what Leibniz has in mind when he argues that 'A=A' is "the first truth," and then attempts to see if all other truths directly or indirectly reduce to this one, or could be derived from the first truth. Let me also remark that I here side with E. Cassirer (and against Frege) in taking identification to be the first function of language and predication to be just one of several different ways in which identification takes place; for Cassirer's view, see *The Philosophy of Symbolic Forms*, trans. R. Manheim (New Haven: Yale University Press, 1957), vol. 1, 282–4, and vol. 3, 293 ff.

20. E.g., N. Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason"* (New York: Humanities Press, 1962), 36–7; N. Rescher, *The Coherence Theory of Truth* (Oxford: Clarendon Press, 1973), 9; R. Walker, *The Coherence Theory of Truth* (London: Routledge, 1989), 61–82, and 102–123; and S. Neiman, *The Unity of Reason* (New York: Oxford University Press, 1994), 75. One rare exception is T. Nenon, who by contrast argues that Kant accepts a correspondence theory of truth; see his "Limitations of a Coherence Theory of Truth in Kant's Philosophy," *International Studies in Philosophy* 36:1994, 33–50.

21. Kant usually leaves out of his account empirical concepts and discusses only the role that the categories play, but empirical concepts are just as indispensable for his account.

22. This view is defended by G. Prauss, *Erscheinung bei Kant* (Berlin: Walter de Gruyter, 1971), 86–7, and Allision, *op. cit.*, 72–3. Based on A788/B816, K. Reich argues that "objective validity is synonymous with truth"; *The Completeness of Kant's Table of Judgment*, trans. J. Kneller and M. Lasonsky, (Stanford: Stanford University Press, 1992), 44. My opinion is that Reich is wrong. One of the things that speaks against him (which Prauss and Allision do not mention) is Kant's conception of "indeterminate objectively valid" maxims of reason (cf. A647/B675 ff.); they are objectively valid but—with the cognitive capacities we have—their truth value cannot be determined.

23. L.W. Beck, "Did the Sage of Königsberg Have No Dreams?," *Essays on Kant and Hume* (New Haven: Yale University Press, 1978), 54.

24. Kant commonly uses the term 'interest', instead of goals, intentions, and expectations; cf. A462/B490 ff.

25. See *Critique of Judgment*, sections 76–7. See also B68, B72, B138, B165, B306–9, and Kant's letter to Herz from 1772 (11:129–35); *Philosophical Correspondence, 1759–99*, ed. and trans. A. Zweig (Chicago: University of Chicago Press, 1967), 70–6.

26. That Kant believed this to be so is suggested, for instance, by his insistence that there can be no purely universal formal criterion of truth that would disregard the specific properties of the objects in question; see A58–9/B83–4.

27. R. Walker, for instance, remarks: "If however every characteristic of anything that we are ever aware of is the result of our mind's activity, it becomes unclear how there can be any element in perception which we do not ourselves provide in accordance with the rules of the mind's own working" (*op. cit.*, 102–3; see also 119). Cf. P. Parrini, "On Kant's Theory of Knowledge: Truth, Form, Matter," *Kant and Contemporary Epistemology*, ed. P. Parrini (Dordrecht: Kluwer, 1994), 217.

28. For an elaborate defense of this thesis on Kantian grounds, see M. Munitz, *The Question of Reality* (Princeton: Princeton University Press, 1990), 23–79. For a further discussion of the two previous constraints on the side of the object, although not in the context of Kant's philosophy, see N.R. Hanson, *The Patterns of Discovery* (Cambridge: Cambridge University Press, 1959), esp. chapter 1. I have used Hanson's considerations to reconstruct a Kantian theory of perception in my book *Anamorphosis: Kant on Knowledge and Ignorance* (Lanham: University Press of America, 1997), chapter 2.

29. J. Dewey, *The Quest for Certainty* (New York: G.P. Putnam's Sons, 1925), 204–5.

30. W. James, *Pragmatism* (New York: Dover, 1907), 80. Although he is not a pragmatist, Pivcevic develops a similar idea about truth as an event; see *op. cit.*, 174–204.

31. James, *op. cit.*, 80. This effectively means that James is not willing to accept any sharp distinction between what is true and what passes for truth.

32. Plato, *Timaeus*, 28c–31a. For further discussion, see Munitz, *op. cit.*, 23–79.

33. This thesis was defended by R. Campbell, *op. cit.*, 434 ff. In his justified opposition to conceptions of truth which assume that truth consists in *reaction* to the existing world, Campbell goes to the opposite extreme and claims that truth consists in *action*. This, however, is equally mistaken: truth consists neither in reaction nor in action, but in interaction.

34. I am grateful to Robert Howell, Gerold Prauss, Nalin Ranasinghe, and Parick Tinsley for their comments on an earlier version of this paper.

The Anti-Reductionist Kant

Gordon G. Brittan Jr.

1. There is a sense, of course, in which all philosophers are our contemporaries, participants in an ongoing and important conversation. But Kant has special interest for us. He more than anyone else challenges the characteristic twentieth-century task of Anglo-American philosophy, that of providing “analyses,” or, what comes to the same thing, reductive accounts of apparently problematic phenomena or of the concepts describing them. Among these accounts are logicism: the reduction of mathematics to logic; behaviorism: the reduction of mind to bodily behavior and dispositions to behave; phenomenism: the reduction of objects to sequences of sense-impressions; operationalism: the reduction of quantities to measuring operations; and ethical naturalism: the reduction of values to facts.

The “isms” multiply without necessity. Kant rejects them all. My aims in this paper are to clarify the grounds of this rejection, show why they require reductions in science while precluding them in philosophy, and argue for a fundamental distinction between “objects” and “constructs.”

A merit of this collection in honor of Lewis White Beck is that it brings us together not as metaphysicians or moral philosophers, but as Kantians. I will talk mainly about the theoretical side of Kant’s position. But the points made will, at the end, be applied to the practical side as well. At every turn, I will make use of Beck’s insights and underscore his belief that the history of philosophy can be brought to bear on present problems. This, ideally, is reunion in philosophy.

2. Reductionist programs in philosophy are so familiar that little needs to be said to characterize them here. They share a common form, which is typically deductive. Moreover, at least one premise of the deduction asserts a conceptual tie between key terms in the respective sentences. Phenomena are reduced when one sort of sentence, say setting out the Peano Axioms for arithmetic, can be derived by way of a claim about meaning from another sort of sentence, in this case derived by way of definitions of “0” and “is the successor of” from the axioms of set theory. They also

share a common motive, which is typically epistemological. Our talk about what is inferred and otherwise merely probable, the minds of others for example, is to be grounded in what is immediately accessible or incorrigible, in this case our observations of bodily behavior. It is in this way that reductive accounts are held to provide a "foundation" for the ordinary beliefs we hold and the claims we commonly make, a fortress against skeptical attack. Where our claims cannot be well-founded, as in religion, by deriving them from others more basic or secure, they should be given up.

3. Kant rejects both the motive and the form of reductionist programs as just characterized. To begin with, there is no special reason to be skeptical about our ordinary sorts of claims about the world. The argument from illusion does not entail its intended conclusion, and the fact that other people have minds is not to be doubted. The claims made in mathematics and science are particularly well-founded, and even as regards morality there is no systematic consideration to suggest that our routine judgments might be in error. Foundationalism is thus not necessary. Neither is it possible. That is, it is not possible to translate or otherwise reduce any of these sorts of claims to what is more accessible or more certain, since nothing is more accessible or more certain, neither sense-impressions nor principles that reason purportedly discovers by reflecting on itself.

It is sometimes held that Kant attempts to derive the laws of Newtonian mechanics from his own *Principles of Pure Understanding*, the latter providing a firm foundation for the former. But this interpretation mistakes his strategy. He wants to validate the *Principles*, not science, and on one main line of argument must assume the epistemological security of the laws of Newtonian mechanics to do so. His question is never "are these laws possible?" that is, *do* they apply to the objects of our experience, but always "*how* are they possible?" that is, what are the conditions of their application? It is in *this* connection that the apparatus of the *Critical Philosophy* is introduced.

Moreover, Kant rejects the form of typical reductionist programs. To put his point very briefly, derivations of this kind illegitimately introduce the mathematical method into philosophy, and the key terms involved cannot be provided with explicit and satisfactory definitions of the kind required for the reduction to go through.

4. But the real grounds of Kant's anti-reductionism lie much deeper.¹ In the process of reduction, the objectivity of experience, the axis around

which his position turns, is invariably subverted. For in the process of reduction, objects are replaced by *constructs*,² and without *objects* there can be no objectivity. It will take some time to illustrate and defend this two-part claim. But first let me outline the argument for it. In what follows I will simply assume the linchpin of Kant's master argument, the Transcendental Deduction, that application of the concept of the self as subject of experience requires that at least some of our judgments about that experience be objective.

For Kant, judgments are objective in that they are made true or false by the properties of the objects to which they refer. It is in this sense that without objects there can be no objectivity.³ But, he continues, such objects are not given in experience; they are represented. The problem is to say when our representations truly represent or, I prefer to say, refer to an object. Among the conditions that must be satisfied is this: the truth or falsity of a judgment which "objectively refers" must be independent of the *manner* in which the reference is made. If the truth or falsity of a judgment depends on the object to which it refers, then it should not matter *how* one refers to it. Or again, objects, introduced originally in the context of judgments about them, must be independent of any particular *conception* of them, which is just another way of saying that in order to be "objects" in the appropriate sense the *same* objects must be conceivable or representable in different ways. To speak of objects is thus to presuppose, among other things, the existence and satisfaction of certain identity conditions. It is at this point, particularly with regard to the objects of outer experience, that the forms of space and time, and the Principles that make their further determination possible, are invoked.⁴ For it is largely in terms of their spatial-temporal locations that we identify and individuate such objects. But the "constructs" which emerge as the result of our reductive activities are not objects. This is to say that they do not satisfy appropriate and intuitive identity conditions.

This, then, is Kant's argument in outline: no objectivity without objects; no objects without identity; reductive "constructs" lack identity; therefore, reductive "constructs" lack objectivity. But, as assumed, we cannot do without objectivity. Therefore, reductive "constructs," and the accounts which issue in them, must be rejected.

5. Kant's rejection of logicism, the presumed derivation of mathematics from logic and definitions, provides an initial example of his anti-reductionism. First, there is no genuine epistemological motive to provide

mathematics with a firmer foundation. Mathematics is at least as secure as the logic from which it is ostensibly derived, a conclusion only reinforced by the post-Cantor vagaries of the set theories with which that logic must be identified if the reduction is to go through. Second, insofar as logic is construed as “analytic,” in some sense true in or of all possible worlds, it does not suffice for the derivation of all mathematical propositions, many of which, owing to their existential character, are true in or of only a restricted number of possible worlds. Third, and most important, the various set-theoretical constructs that have been proposed as analyses of the concept of number differ between themselves in a number of ways. Although all do duty for the concept of number in the sense that they satisfy the Peano Axioms for arithmetic, there is no way in which they can be identified. The concept of number is thus left indeterminate in the course of reduction, which is to say that arithmetical judgments have lost their objectivity.⁵

6. Perhaps an even clearer example of the point I want to make has to do with operationalism, the reduction of quantities to measuring operations, an account most often associated in the United States with the physicist P.W. Bridgman. This example is particularly important both because the concept of quantity is at the heart of Kant's theory of corporeal nature and because his own account of the concept superficially resembles Bridgman's.

Bridgman's position is familiar. But allow me to sketch it briefly. Quantities are not properties of objects, inherent in things. So much was made clear, at least for Bridgman, by Einstein's theory of relativity. Rather, quantities result from, and hence can be reduced to, the various measuring operations by which they are determined:

The concept of length is therefore fixed when the operations by which length is measured are fixed: that is, the concept of length involves as much as and nothing more than the set of operations by which length is determined. In general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations.⁶

Moreover, if we adopt this point of view,

we need run no danger of having to revise our attitude toward nature. For if experience is always described in terms of experience, there

must always be correspondence between experience and our description of it, and we need never be embarrassed, as we were in attempting to find in nature the prototype of Newton's absolute time.⁷

So long as quantities are identified with the ways in which we measure them, our judgments about them need never be revised.

Two aspects of the operationalist position should be noted before we proceed to the Kantian criticism of it. One is Bridgman's assumption that the only alternative to naive realism concerning quantities is their reduction to measuring operations. But naive realism, he contends, is incompatible with the deliverances of twentieth-century physics. Therefore, reduction of quantities to measuring operations is the only tenable view. The other aspect of the operationalist position is that it equates objectivity with intersubjectivity.⁸ If different researchers performing the same specified measuring operations generate approximately the same numbers, then to that extent these operations, and the quantitative judgments which reduce to them, are objective.

7. From Kant's perspective, the operationalist account of quantities is problematic in a number of different ways. But the one on which I wish to concentrate is this. It follows from Bridgman's characterization that different measurement procedures, even when they issue in the same numerical result, must be taken as defining different quantitative concepts. Thus, when we measure length in the usual ways, optically and tactly, we are in fact operating with separate concepts which, in order to avoid all ambiguity, should be terminologically separated. It is, for Bridgman, a merely contingent fact that different measurement procedures should issue in the same result, a fact that more precise procedures might at any moment overturn.⁹

Of course, the multiplication of quantities on this account would complicate science enormously. But more to the point, so far as Kant is concerned, is that in tying quantities to specific procedures, they lose their objectivity. For quantities on the operationalist account are not independent of the way in which they are conceived. Or, what comes to roughly the same point, there is no way on this account to say that different procedures measure or determine the *same quantity*. But unless and until conditions for the identity of quantities, length for example, or mass, can be given independently of the procedures that measure them, then our judgments about quantities are not, for the reasons already given, objective.

It is not surprising that Bridgman, in the typically American straightforward and innocent way, concedes the point. After agreeing that identity is "one of the most fundamental of all concepts with which we describe the external world" and "thinking would be almost inconceivable without such a concept," he goes on to maintain that it is no more than an empirical, and once again contingent fact, that we are able to identify and individuate objects, a fact this time jeopardized by developments in quantum mechanics.¹⁰

But as Brian Ellis points out,¹¹ this position very much resembles Berkelean idealism. If one procedure, using our eyes, is logically independent of another, using our feet, to measure "lengths" (put in quotation marks to highlight the ambiguity of this term), then there is no way in which to determine that there is *a* length which a particular object has, hence no way in which to determine that there is a *particular* object which has *it*. Separation of what are ordinarily taken to be the same quantities, in the process of reducing them to diverse sets of measuring operations, has the effect of dissolving the object that is ordinarily taken to be their locus. But to remind you once more, this Berkelean dissolution of the object (and its later reconstitution in the mind of God) is, Kant argues, incompatible with the requirements of the unity of the self as subject of experience. Without an it, an object, there can be no I, a subject. Operationalism undermines the object, and eventually the subject of experience, and must therefore be rejected.

8. It should be clear from the foregoing that the objectivity of quantities depends on the possibility of their being identified. Such identification takes place when various quantities are reduced, not to measuring operations or, more generally, to something which is not itself a quantity, but in the course of scientific progress, to more fundamental quantities. This is, I think, one burden of the Appendix to the Transcendental Dialectic in the first *Critique*, and one aspect of his insistence there (and in the *Critique of Judgment*) on the central role played by the concept of the systematic unity of nature.

It might otherwise be thought misleading to call Kant an anti-reductionist since the Appendix brims with reductionist imperatives (see, e.g., A649/B677). There is no difficulty, however, as soon as we distinguish between philosophical reductions, guided by an epistemological motive, and scientific reductions, which have a unifying, hence explanatory, role. Philosophical reductions undermine the objectivity of experience. Scien-

tific reductions help to secure it. For what scientific reductions in fact both demonstrate and depend upon is the identity of certain quantities, gravitational and inertial mass for instance, on which the objectivity of experience depends. Given the contingency of the identification and eventually the reduction of quantities with and to one another, scientific reducibility must remain a regulative ideal. Its successful application cannot be guaranteed from the outset. *A fortiori*, there is no guarantee that the quantities we now think of as basic, among them, perhaps, the five fundamental forces, will remain so. What is necessary for us to make good on the claim that the same quantity can be referred to or measured in different ways, however, is that *some* reduction of quantities can in principle be carried out, in the context of the elaboration of successful scientific theories, which is what the objectivity of quantities, and thus of our scientific picture of the world, comes down to.¹²

I realize that this is a rather glib reading of a very difficult section of the *Critique*. What makes the section so difficult is that in it Kant is ambivalent about the status of the concept of systematic unity, which I here understand in terms of the reducibility of quantities (although Kant's examples indicate it also has to do with the reducibility of mental states and activities and a variety of other qualitative properties).¹³ At times, he indicates that it is a subjective maxim (A680/B708), and yet at other times he insists that it must be presupposed as "objectively valid and necessary" (A660/B688). More to my immediate point, he never provides an argument for the conclusion that the systematic unity of nature must be presupposed as objectively valid and necessary. One line suggested in the text is that the unification of experience, one source of which is the unity of nature, is presupposed by the *necessity* of the laws of nature. But another, and I now think more powerful line of argument, results from reflection on Kant's general philosophical position. Unless nature were systematically unified, that is, unless it were possible to refer to the same quantities (and qualities) in different ways, and thereby to presuppose their identity, then such reference would no longer be objective, with all that that entails.

9. At the risk of belaboring the point, let me summarize it. The operationalist reduction of quantities to measuring operations would not simply complicate science. There is more to the concept of systematicity than order and neatness. It would also reduce quantities to "constructs" which, lacking the possibility of further identification, would not be "objects" in the appropriate sense. Therefore claims about them would not,

in the same sense, be “objective.” Since the claims about these constructs would, as I have already indicated, be intersubjective, it follows that intersubjectivity is a possibly necessary, but not also sufficient condition of objectivity. Identification is also required. But such identification is at the heart of scientific reduction and, Kant clearly suggests, scientific progress. The objectivity of (outer) experience thus depends, at least in part, on the possibility of such progress being made. It is for this reason, mainly, that it is imperative to pursue it.

10. Kant rejects Bridgman’s conclusion, the reduction of quantities to measuring operations. Like all such reductions it is tantamount to empirical idealism, and thus incompatible with the requirements of the Transcendental Deduction. But he accepts one of Bridgman’s premises, that having to do with the untenability of naive, or—as Kant would term it in this context—“transcendental,” realism with respect to quantities. The option is to develop a more sophisticated or “empirical” realism.

Recall that naive realism with respect to quantities is the view that quantities are inherent properties of objects. If there were but one (physical) object in the world, and no one around to measure any of its properties, it would still be the case that it had a particular length, mass, and so on. Bridgman thinks that reflection on the theory of relativity shows the untenability of naive realism with respect to quantities. Kant has a much more general epistemological argument. This argument is so familiar that I will omit many of its premises (see B66). It is to the effect that we can have *knowledge* only of the relational properties of objects, not as they are in themselves, but only as they can be compared, eventually measured, against one another. All knowable properties of objects are relational properties and, Kant adds, only knowable properties of objects are significant. It is coherent to think about the nonrelational or intrinsic properties of objects, as they are in themselves, but there is no way in which we can give such thoughts truth-values.¹⁴

This epistemological argument has two apparent ontological corollaries. One is that a knowable world must consist of at least two objects, whose properties can then be compared. In the simplest quantitative case, we would have a rock and a ruler. The other corollary is that a knowable world must also contain a being like ourselves who would select one of the objects as a standard of congruence and comparison, in terms of which the others could be measured. For Kant, then, a world without human quantifiers, or something very like them, would be a world without quan-

ties. Which is to say that quantities, although not “constructs,” are in an important sense *constructed*.

To this point, Kant’s position roughly resembles Bridgman’s, not only in its rejection of naive realism, but also in its insistence that measurement depends in certain ways on human intervention. But there is a crucial difference between them. It does not follow for Kant, as it does for Bridgman, that quantities can, indeed must, be reduced to human measuring operations. In this respect, his account of quantities, unlike Bridgman’s, remains in an important sense *realistic*.

The difference between them turns on two distinctions. One has to do with the existence, the other with the identity of quantities. I will put the distinctions in contemporary terms, although the ideas involved are clearly present in Kant’s account of the rule-governed character of mathematical construction¹⁵ and in the section entitled Axioms of Intuition in the *Critique of Pure Reason*.

Quantities or, as he also calls them, extensive magnitudes arise on Kant’s account just in case two different kinds of conditions are met. First, the properties of objects in question must be capable of being ordered in a particular, serial way. Length, for example, is a quantity only if the relation “longer than” is both asymmetric and transitive. Whether or not properties are capable of being so ordered is independent of our ordering them. In this sense, the existence of quantities is entailed by the existence of certain relationships between objects that hold apart from the presence or absence of human quantifiers. In this same sense, quantities are real.¹⁶ Second, in order to be or have quantities, the properties in question must be capable of being added, that is, there must be for them an empirical operation formally similar to addition in arithmetic. Thus, lengths can be added. The total length of two rods put end to end is the sum of their respective lengths.

If the first set of conditions introduces a (quasi-)serial order, the second introduces a metric. It is the second set that, in turn, requires human intervention. For it is Kant’s position, on the basis of the epistemological argument already sketched, that quantities generally (certainly space and time in particular) do not have an intrinsic metric. The metric, rather, is in certain ways “contributed by us.” In fact, we make two contributions in connection with the quantification of properties. One has to do with the choice of units, yards or meters in the case of length, and is clearly arbitrary within a constant factor. The other has to do with the relationship between a property and the standard, a rigid rod once again in the case of

length, used to measure it. The choice of this standard is, for Kant as for Bridgman, up to us, but it is not for all of that, according to Kant, arbitrary. To put a very complex claim in a cursory way, the specification and choice of the standard is not independent of the character of the physical theories we admit, and the character of admissible physical theories is tied to the ways in which we conceptualize experience, in terms of a certain concept of causality for instance. But the ways in which we conceptualize experience are not arbitrary; the unity of consciousness depends upon experience being conceptualized in some, but not other, ways.

The issues here are much more technical and complex than I have made them out to be. The basic point is that quantities do not reduce to measuring operations because, on the one hand, order relationships exist or fail to exist as the case may be, independent of the operations we do or fail to carry out and, on the other hand, the choice of a metric, while ours in certain respects, is not in the most important respect arbitrary or conventional. Bridgman suggests that we can choose any metric we like, by way of the specification of a set of measuring operations, the only constraints being purely pragmatic. Kant disagrees.

The second distinction that distances Kant and Bridgman is that between the order and metrical conditions, the conditions which when satisfied allow measurement to take place, and the quantities to be measured. For Bridgman, as we have seen, each set of specified measuring operations *defines* a quantity. It follows that for any two such sets of operations, there are two distinct quantities. But on Kant's account, a set of order and metrical relationships does *not* define a quantity. Rather, it describes an order and a metric. It is this order and metric, rather than the order and metrical relationships, which is to be identified with a given quantity. Crucially, the same order and metric, the same quantity, can be described in logically independent ways. The criterion of "sameness," it follows, is simply this: if logically independent sets of order and metrical relationships generate the same order and metric under the same conditions, then they are relationships for the same quantity.¹⁷

There are obvious problems in the application of this criterion, particularly with respect to what constitutes "same conditions." Moreover, it suggests that there is a quantity which our various attempts to determine merely approximate. Everything depends on exactly how this suggestion is to be understood. If we understand it in a naively realistic way, that there are quantities "in the world" so to speak, independent of any human intervention, then it is, from Kant's point of view, without significance.

There is no meaningful way in which we can compare our measurements with the quantity in question, and determine the degree to which they approximate it. Rather, we must take the notion of a “real quantity” here as a regulative ideal, the limit of the successive application of our conditions on measurement. This limit is in part a function of, and so constrained by, the considerations that Kant lists in the Appendix to the Transcendental Dialectic: scope, refinement, theoretical coherence. But again, it is the aim of scientific reduction to come closer to this limit, without in principle ever being able to reach it. It is in this idealized sense that science gives us a picture of “reality.”¹⁸

This, then, is the core of Kant’s “empirical” realism. It is not naive insofar as Kant construes all quantities as relational in character and it is “constructive” to the extent that a metric must in certain ways be “contributed by us.” But it is realism nevertheless in that the order conditions can be described without describing any measuring operations or scales of measurement of our devising, in that the metric which we “contribute” is not simply arbitrary but is constrained by certain necessary conditions on possible experience, and also in that the quantities which the order and metrical conditions describe can, on occasion, be identified, although the general possibility of such identification can never be more than a regulative ideal. This last, in fact, is just what is at stake in the formulation and development of the sciences worthy of the name, the description and eventual reduction of quantities to still more fundamental quantities. This is, among several other things, what gives these sciences their “objective” character. Operationalism, in reducing quantities to measuring operations, at the same time forecloses an objective rendering of them, the view that the operations we human beings carry out measure quantities that, in the sense I have tried to spell out, exist independently of us. The quest for certainty that operationalism and other such reductive projects involves inevitably, and perhaps paradoxically, leads to a type of subjectivity. Possibly Descartes taught Kant this lesson.

11. I now want to shift my focus to Kant’s moral philosophy, in the hope that the discussion of the irreducibility of quantities will throw some light on the objectivity of values. There are, admittedly, a variety of ways in which quantities and values differ. But there are also some important analogies in Kant’s respective treatment of them.

John Rawls is in many respects the greatest contemporary interpreter of Kant’s moral philosophy and by all odds the most influential. One

criticizes him only very reluctantly, and not at all for questing after certainty. Still, I think that his own “constructivist” position misses Kant’s point about objectivity, and in missing that point falls prey to certain features of the case against reductionist projects generally. For what results from Rawls’ account, if I read it correctly, are more nearly akin to what might be called “value-constructs” than to “value-objects,” the real things.

We can begin with Rawls’ list of the five elements essential to an adequate conception of objectivity:

The first essential is that a conception of objectivity must establish a public framework of thought sufficient for the concept of judgment to apply and for conclusions to be reached on the basis of reason and evidence after discussion and due reflection. . . . The second is this: It is definitive of a judgment (moral or otherwise) that it aims at being reasonable, or true, as the case may be. . . . As a third essential, a conception of objectivity must specify an order of reasons as given by its principles and criteria, and it must assign these reasons to agents . . . as reasons they are to weigh and be guided by in certain circumstances . . . as a corollary, we have a fourth essential: namely, a conception of objectivity must distinguish the objective point of view . . . from the point of view of any particular agent . . . or of any particular group of agents . . . at any particular time. . . . The fifth essential is that a conception of objectivity has an account of agreement in judgment among reasonable agents.¹⁹

It should be clear that, with one notable exception, these five elements frame a conception of objectivity as intersubjectivity. Indeed, they capture very well Bridgman’s conception of the objectivity of our judgments about quantities: carefully specified and well understood procedures, publicly formulated and publicly applied, that lead to approximately the same results when put into practice by rational and unbiased individuals. There is the same superb caution, the same willingness to revise our judgments on the basis of further experience and reflection.

The exception, of course, concerns the second item on the list, that it is definitive of judgments that they “aim at being reasonable, or true, as the case may be.” For what is true, as opposed to what is reasonable, has little to do with widespread agreement, however “correct” the procedures leading to it.

12. Rawls’ discussion is framed in terms of three model positions: rational intuitionism, Kantian constructivism, and political constructivism.

According to him, the first position contrasts with the second two in its insistence on truth as against reasonableness as the aim of all judgments. Since he is wary at best of the notion of truth in moral philosophy, he favors the constructivist variants over their intuitionist alternative. Hence despite the conciliatory things he says about intuitionism, his own conception of objectivity is tantamount to reasonable inter-subjectivity.

But surely this is a mistaken reading of Kant. In the first place, Kant is explicit that moral judgments are true or false. The most striking passage for me in this connection is that at A476/B504 in which he asserts that “We must be able, in every possible case, in accordance with a rule, to know what is *right* and what is *wrong* . . .,” from which it follows by way of some easy premises that every moral judgment is knowably true or false. In the second place, and more important, this reading is incompatible with Kant’s general account of objectivity, that judgments are objective in that they are made true or false by the properties of the objects to which they refer.

Of course, there are important differences between Kant’s moral and theoretical positions, and moral objects, however they are rightly to be characterized, differ from the objects of outer experience in particular. I will explore some of these differences in a moment. But it nevertheless remains a chief plank of Kant’s philosophical framework that there can be no objectivity without some sort of object, no objective judgments that are not also true or false. This is as true of his practical as it is of his theoretical philosophy.

Rawls apparently comes to his reading of Kant by way of the following argument. A judgment is either reasonable or true, as the case may be. But if a judgment is true, then it is true of “an independent order of values.” But Kant does not think that any judgment is true of an independent order of values. Therefore, Kant must hold that moral judgments are reasonable at best, a view reinforced by a consideration of the ways in which we, as rational agents, come to agree to the terms of our common social contract.

The problem is with the second premise of Rawls’ argument. Kant does not think that any moral judgment is true of “an independent order of values,” where this phrase is taken as implying the possible existence of values in a world devoid of rational agents. Values require valuers, just as quantities require quantifiers, one great parallel between the practical and theoretical halves of Kant’s position. But it does not follow that moral judgments are not true or that values are not real. It is just that we have to

understand “true” and “real,” and for that matter “independent order,” in a different, undeniably complex way. Rawls, like Bridgman before him, seems to think that the only option to naive realism, which is what rational intuitionism in moral philosophy comes to, is a kind of (in Rawls’ case very attenuated) philosophical reductionism, of values to procedures. Kant rejects naive realism everywhere, including its appearance in moral philosophy. But his position continues to be realistic in important respects, even when it comes to moral matters.

13. I do not have anything like a developed interpretation of Kant’s moral realism, and certainly can add very little to Beck’s account of Chapter II of the *Analytic in the Critique of Practical Reason*.²⁰ A developed interpretation would include an account of the ways in which moral objects, while in some sense “constructed” or, to use Kant’s term, “produced” by us, nevertheless have a kind of reality independent of the ways in which we construe them (*Practical Reason*, 5:66). It is enough at this point to claim that what Kant calls the objects of practical experience must be enough like the objects of theoretical experience to cohere with his general position on objectivity and truth. But following up one aspect of the parallel suggested between quantities and, as a blanket term, values (good and evil) leads us in the direction of such an interpretation.

In fact, the parallel between quantities and values, and between the kinds of objects to which we ascribe them, is far from perfect. The objects of theoretical or (what comes to the same thing for my purposes here) outer experience are located in space and time. They have certain causal properties, among which is the ability to incite our sense organs; their behavior is subject to lawlike regularities; and they can, with regard to fundamental properties, be measured. All of these facts are sources of the objectivity of the judgments we make about them. They are in some general sense constitutive of their empirical reality. Presumably in these respects and perhaps others they are unlike the objects of practical experience which, to begin with, are not really “experienced.”²¹

What qualifies them all as *objects*, and is the principal source of their “independence” of the specific ways in which we conceive of them, is that they involve a rule-governed synthesis.²² In the theoretical case, these rules are the principles of the understanding; in the practical case they are fundamental moral principles, disclosed to us in rational reflection, the paradigm of which is what Kant calls the “moral law.” In the *Critique of Practical Reason* (5:65), he brings them all together as “laws of causal-

ity,” on the one hand of nature, on the other of freedom. They are, to continue the analogy, rules for the synthesis of intuitions and of desires, respectively, and in this sense confer the unity characteristic of objects in general.

Surely in all of this Kant is stretching the meaning of “object” to the point where the assumed parallel between the “laws” of theoretical understanding and of practical reason begins to beg all of the most important questions. Yet two initial points might be made in its behalf. One is that Kant provides a ground for distinguishing between what is objective and subjective in my moral experience, between “X is good” and “I desire X,” over and above the fact that the first judgment might happen to reflect a common moral consensus. It is that “X is good,” if objectively valid, is a judgment that any rational person as such would make. Kant assumes, of course, that rationality is itself unconditioned by psychological, social, and historical factors, but this assumption (which can indeed be justified) is as important for his theoretical as for his practical philosophy and does not beg any special questions with regard to the latter.

The other point to be made in behalf of Kant’s emphasis on the objectivity of moral judgment is that the distinction between “X is good” and “I desire X” applies as much to particular historical communities as it does to individual persons. That is, the concept of the good (embodied in various moral principles) restrains the behavior of particular historical communities as it does my own, and provides a basis (which is otherwise impossible) for evaluating them. It is also, according to Kant, that which most importantly distinguishes us from other animals.

Now it follows from his characterization of them, I think, that all objects satisfy certain conditions of individuation and identification. They are relatively straightforward in the case of the objects of theoretical or outer experience (sameness of spatial-temporal position, itself precisely determined only by way of the Principles of the Understanding) and of their measurable properties (sameness of order and metric), much less so in the case of the objects of practical “experience.”²³ What we have to say in the second case is that the moral principles that determine the objects of practical experience at the same time provide their identifying and individuating conditions. An action as object of the will is “the same” if it instantiates these various principles together. Indeed, it is tempting to say that for Kant the objectivity of morality is demonstrated by and depends upon the fact that different formulations of the Categorical Imperative yield the same results, that is, lead to our identification of the same actions as good.²⁴

But against the background of our earlier discussion of the imperative nature of scientific reduction, there is an interesting corollary to my claim that moral objects, too, must be capable of individuation and identification if they are to play their proper role in anchoring judgments about them.

14. Before spelling out this corollary, we need to remind ourselves of the distinction between philosophical and scientific reduction. Philosophical reduction is foundational in character. No such reduction is, according to Kant, either possible or necessary, a claim which he develops in detail on a case by case basis. In particular, quantities cannot, and need not, be reduced to measuring operations and values cannot, and need not, be reduced to facts. But as we have seen, his point is much deeper. The reduction of quantities to measuring operations would undermine the objectivity of judgments about them, which in turn would preclude the unity of the conscious self. In *somewhat* the same way, the reduction of values to facts (or something else allegedly more epistemologically secure than values) would undermine the objectivity of judgments about *them*, which in turn would preclude the autonomy of the rational agent. This last point is emphasized by Dostoyevsky and others of like mind in their attacks on Mill in behalf of personal responsibility. But it is clearly implied by Kant's insistence on the freedom of the rational agent in the choice of ends.

So much for philosophical reductionism. Now to the role of scientific reduction. It is sometimes suggested, to repeat an earlier point, that it is in the interests of the economy of thought, a matter of organization, clarification, and simplification. Kant himself sometimes indicates that the greater order that scientific reduction affords exhausts reason's ends. But as we saw, such reduction, involving the identification (and on occasion the separation, as in the distinction forced on us between relativistic and rest mass) of quantities at the same time makes good on the claim that judgments about them are objective. It is in this rather novel way that Kant argues for the familiar conclusion that objectivity requires and results from unification in the sciences.

Something like this, I believe, is also the case with respect to the reduction of values. Kant is not an ethical reductionist; judgments about values cannot be based on or derived from something different or more secure.²⁵ But his position does require that some values are more basic than others, to the extent that they can be used to ground or at the very least rank-order the others.

It is sometimes suggested that conflicts between moral principles pose problems for agents who must act on them. Indeed, this is often taken as a crucial objection to principle-based theories of action and evaluation like Kant's. But the much deeper difficulty for him is that a conflict of principles is an indication of the fact that they cannot be reduced to and, in a way analogous to the case of quantities, identified in terms of a more basic principle. To put it in another and simpler way, a conflict of moral principles indicates a lack of objectivity in our moral judgments. Once again this is an old conclusion, but Kant's argument for it follows by way of a very general premise concerning the nature of objects and objectivity. For our moral judgments, in particular, to be objective they must be about objects that are capable of being identified, but such identification is demonstrated when something like a reduction or "derivation" of at least some of these judgments is carried out. The reduction does not *prove* that a given moral judgment is true. But it does demonstrate its objectivity. This is the task of all philosophy, practical as well as theoretical. The details of the demonstration vary with the type of judgment at stake. But in every case, I claim, something very roughly *like* a "scientific" reduction is involved. Moreover, just as scientific progress is a function of this sort of reduction being carried out, so too there is a type of moral progress, under the aegis of a similar regulative ideal, which depends, if only in part, upon the reduction of some moral principles to others.²⁶

15. Following Kant, I have stressed the importance of the conceptual tie between objects and objectivity. It might be wondered whether this tie really obtains in the case of moral philosophy, or whether the talk about the "objects" of practical experience is a misguided effort to make the overall position more symmetrical, more architectonic than it really is. Put another way, it might be wondered whether there is room between rational intuitionism and political constructivism for a more sophisticated and robust alternative, in the same way that Kant's empirical realism is an alternative to naive realism and philosophical reductionism. Certainly Kant's attempts to clarify the notion of an object of practical experience are tortuous, and I have not yet succeeded in making their identity conditions precise. But at least one consideration seems to me to motivate more effort in this direction.

Rawls provides us with a theory of justice that is largely procedural in character. Again, in this respect it resembles Bridgman's account of quantities. This is a plausible move on Rawls' part since justice, after all, has

to do with the fair distribution of goods that have already been characterized as such. A procedural account of these goods, in turn, however much it coheres with the way in which, increasingly, we make individual and social decisions in democratic societies, is not completely satisfactory. We also need a substantive account of goods or more generally values which, unlike justice, cannot be reduced to the formulation and following of certain agreement-generating procedures. But this is to say we need an account of values as objects and not as the mere end-products of these procedures, objects about which our various moral claims can be judged as true or false. For we must always leave open the possibility that, however great the consensus among reasonable and like-minded individuals, we might still be wrong about what is right. But this is among other things to say that there is a regulative or normative ideal towards which we strive and which can, on occasion, correct our strivings. This is a conclusion that I owe as much to Beck as to Kant, and that the former by his example as by his teaching demonstrated.²⁷

Notes

1. In *Kant's Theory of Science* (Princeton: Princeton University Press, 1978), I pursued the anti-reductionist theme along several different lines. What follows is the result of a subsequent twenty years of thinking about it.

2. Ralf Meerbote has pointed out to me that this expression is misleading. We need to distinguish between constructions, such as we find in Kant's philosophy of mathematics and in a roughly analogous way in his moral philosophy, and constructs *per se*, the result of the reductive activities outlined above. Constructs *per se* are at stake in my account of Kant's rejection of them.

3. "Understanding," Kant writes at *Critique of Pure Reason*, A199/B244 (Kemp Smith translation, New York: St. Martin's Press, 1965), "is required for all experience and for its possibility [read objectivity]. Its primary contribution [in making objectivity possible] does not consist in making the representation of objects distinct, but in making the representation of an object [and thus reference] possible at all."

4. The forms of space and time, as they are characterized in the Aesthetic, do not serve to individuate and identify the objects of outer experience. It is only when space and time have been fully determined by the machinery of the Analytic that these objects can be fully individuated and identified, and reference to them made possible. It is in this connection, in the course of the B-Deduction, that Kant parenthetically equates objects with determinate spaces (B138). Simi-

larly, the passage at A199/B244 (already quoted) goes on to make clear both that the “possibility of an object” depends on being able to locate it in a determinate temporal order and that such location is itself dependent upon prior application of the principle of causality to our experience.

5. Two points of clarification. First, Kant is not at all clear about the existence and status of mathematical *objects*. See Charles Parsons, “Kant’s Philosophy of Arithmetic,” in *Kant’s Philosophy of Mathematics*, ed. C. Posy (Dordrecht: Kluwer, 1993), 73–5. But whether and in what precise sense they are objects for him, Kant insists that numbers must be determinate in that different characterizations of them pick out or refer to “the same number.” Second, Kant holds that numbers are in a sense “constructed”; they result from our counting activities, the generation of a series from a basic unit. But this properly arithmetical construction, on which the properties of numbers derive from their place in the series and hence do not vary as a function of the various ways in which they are characterized, must be distinguished from the set-theoretical constructions to which it allegedly reduces and on which the properties of numbers are not directly and uniquely determined.

6. *The Logic of Modern Physics* (New York: Macmillan, 1948), 5.

7. *Ibid.*, 6–7.

8. See Carl G. Hempel, *Aspects of Scientific Explanation* (New York: Free Press, 1965), 123 ff.

9. See *The Logic of Modern Physics*, 10.

10. *Ibid.*, chapter 3, “The Concept of Identity.”

11. In *Basic Concepts of Measurement* (Cambridge: Cambridge University Press, 1966), 36 ff.

12. It is not at all clear what *form* scientific reductions are to take, if indeed they are all to take the same form. But it is surely a mistake to think that, for Kant, they must have a deductive form, the propositions of classical thermodynamics (for one example) following from the basic principles of statistical mechanics. The reduction, at the very least, is not by way of sentences but concepts.

13. The point of these psychological and otherwise scientific reductions remains, however, the same: the identification of objects and properties on which objectivity depends.

14. Newton claimed in the *Principia* that, while “the common people conceive these quantities [i.e., time, space, place, and motion] under no other notions but from the relation they bear to sensible objects,” it was nonetheless possible to conceive of them as they were in themselves, absolutely. Kant denies this claim. It is only relative to “sensible measures” that we *can* attribute a metric to them, although quantities cannot be reduced to such measures.

15. See A715/B743 ff.: “[M]athematics can achieve nothing by concepts alone, but hastens at once to intuition, in which it considers the concept *in concreto*, though not empirically, but only in an intuition which it presents a priori, that is, which it has constructed, and *in which whatever follows from the universal condi-*

tions of the construction must be universally valid of the object of the concept thus constructed" (Kemp Smith translation; my italics). I understand the last part of this passage as distinguishing sharply between our constructive activities or operations and the conditions that determine the correctness of these operations.

16. On my reading, Kant puts this by saying that the ordering or topological relationships are "intuitive," properties of space and time as they are given, prior to the metrical relationships that require *our* further conceptualization. Thus, topological relationships are at stake in the Aesthetic, metrical relationships in the Analytic.

17. Following Ellis, *Basic Concepts of Measurement*, 34. See also his discussion on 48 ff., of the difficulties in making this criterion precise.

18. Thus as against Bridgman, who wants to abandon absolute space and time as mere fictions, Kant wants to claim that there are fundamental senses in which they are both ideal and real, and that a purely reductive account of space and time cannot be given. For a deeply illuminating discussion of the way in which Kant works out the details of this position in connection with absolute space, see Michael Friedman, *Kant and the Exact Sciences* (Cambridge: Harvard University Press, 1992).

19. *Political Liberalism* (New York: Columbia University Press, 1993), 110 ff.

20. See his magisterial *Commentary on Kant's "Critique of Practical Reason"* (Chicago: University of Chicago Press, 1960). Carl Posy's paper, "Kant's Moral Realism," has also had an important influence on my thinking.

21. This is to say that there is nothing corresponding to intuition in the case of the objects of practical experience, although Kant does suggest that sensations are to theoretical understanding what objects of desire are to rational choice. See Beck's *Commentary*, 131. And just as there is nothing quite corresponding to sensibility in the case of moral experience, so too there is nothing quite corresponding to understanding and the way in which it constitutes objects. Reason's regulative role is broadened so as to provide the principles on the basis of which the objects of practical experience are determined. I owe these reflections to Ralf Meerbote.

22. This claim appears to contradict what Kant says at A477/B506 that "in the case of right and wrong, the object is not to be met with outside the concept." However this sentence is to be understood, it cannot contradict my claim, since the same thing is said of "transcendental philosophy" where the independence of the object from the specific ways in which it is conceived is crucial to Kant's argument. In fact, the point here seems to be that which is taken up again in Section II of the Analytic in the *Critique of Practical Reason*, that practical knowledge does not require intuition. As Beck notes in his *Commentary*, 141, here "we have the one instance in the whole of knowledge where thought can create its object directly, and not by being a cause of a series of events which may terminate in an object."

23. Indeed, practical reason is parasitic on theoretical understanding in just this sense, that objects must first be constituted by the former before they can be brought, by practical reason, under the concepts of good and evil. See the *Critique of Practical Reason*, 5:65.

24. I owe this point to James Allard.

25. In just the same way that Kant is often pictured as deriving the basic laws of Newtonian mechanics from the principles of the understanding, so too is he pictured as deriving basic moral principles from the essentially formal, and thus nonmoral, principle expressed in the Categorical Imperative. Both pictures are mistaken.

26. I owe this suggestion to James Allard.

27. James Allard, Prasanta Bandyopadhyay, Predrag Cicovacki, and Ralf Meerbote have made a number of useful comments on an earlier version, first read at a conference on Kant at the University of St. Andrews, for which I thank them.

Analyticity and the Semantics of Predicates

Carsten Held

I. Introduction

Semantics investigates linguistic entities in their most fundamental dimension; it studies the relations of signs or strings of signs, their meanings, and their eventual references in the world. While linguistic semantics investigates concrete languages, in all their empirical details, it yet presupposes an intuitive understanding of the different kinds of signs and their functions: assertions, commands, and questions; names, predicates, relations; quantifiers, demonstratives, indexicals; and so on. As we all command at least one language, we have a pre-theoretical mastery of the linguistic instruments studied by theoretical analysis.

Philosophical semantics cannot rest content with such intuitions. Its proper task may be defined as the theoretical clarification of different kinds of signs and their functions: how signs in general have meanings and, eventually, references; how they interlock to form meaningful complexes and how their functions differ; what gives them the status of names, predicates, relations; what structurally guarantees that we understand certain strings of signs as assertions, questions, commands; and so on.

Indeed, *the* fundamental problem of semantics may be to explain how subsentential semantic entities combine to form the basic self-sustained semantic unit, an assertion, and to describe what these subsentential entities must be like. In a sense, this was the main occupation of Frege in his attempt to lay foundations for philosophical semantics. In recent decades, philosophers have at the same time found Frege's compositional approach seriously wanting and occupied themselves more globally with meaning—all too often with skeptical results.¹ I believe that Frege's program of clarifying the status of subsentential semantic bits in regard to their function to form full-fledged assertions is a promising one and that it deserves to be refurbished and continued.

The purpose of the present reflections is to show how the analytic-synthetic distinction, as set out by Kant, can be exploited to give answers

to some of the Fregean questions in philosophical semantics. Of course, I do not mean we should interpret Kant as an early philosopher of language; if we should list those topics of present-day philosophical concern ignored by Kant, language certainly would be included. However, pioneers of modern philosophy of language, such as Frege, Carnap, and Quine, have occupied themselves with the analytic-synthetic distinction, and invariably have addressed Kant's influential approach. What these pioneers shared was the tacit presupposition that the theoretical instruments used by Kant for an analysis of inner thought processes can be used to analyze linguistic utterances as well. This presupposition seems a good one, and congenial to Kant's original transcendental project. What we produce in the mental processes, understood by Kant as the exercise of understanding and reason, are thoughts about objects; what we linguistically communicate is, *inter alia*, these thoughts. Hence, it is a good (if not trivial) premise to assume that language and thought are structurally alike, since the former is produced to transmit the latter. This presupposition is what makes implanting Kant's epistemological categories into the philosophy of language a promising project.

This project seems all the more promising, as it is the neglect of the epistemological or cognitive origin of conceptions such as the analytic-synthetic distinction which has led some (mostly empiricist) philosophers to expunge them altogether from the philosophical study of language. I will try to reverse this development, by showing that especially Kant's conception of analyticity, properly understood, can make a fruitful contribution to philosophical semantics. Specifically, I will show that Kant's conception of analyticity illuminates his own view of the functioning of concepts. I will interpret concepts as the meanings of predicates, and thus will use this view to present a more general conception of the semantic function of predicates.

II. Kant's Conception of Analyticity

Kant's various accounts of analyticity are always a part of his project of establishing the distinction between analytic and synthetic judgments. The distinction, although prefigured in the work of earlier philosophers, originates with Kant.² We can, thus, broadly say that analyticity as it appears in modern philosophy of language, analyticity as discussed by Frege and Carnap, or as attacked by White and Quine, is *Kantian* analyticity.

Typical accounts of the conception are presented in the introduction of the *Critique of Pure Reason*, in the *Prolegomena*, and in the late *Preisschrift*:

In all judgments where the relation of a subject and a predicate is thought . . . that relation is possible in two ways. Either the predicate B belongs to the subject A as something that is (covertly) contained in this concept A; or B lies entirely outside the concept of A, although it is connected with it. In the former case I call the judgment analytic, in the latter synthetic (A6/B10).

Analytic judgments say in the predicate nothing but that which was already really thought [*wirklich gedacht*] in the concept of the subject, albeit not so clearly and with the same apprehension" [*Bewußsein*] (*Prolegomena*, 4:266).

Judgments are analytic, if their predicate only clearly (*explicite*) represents what was already thought, albeit obscurely (*implicite*), in the concept of the subject (*Preisschrift*, 20:250).³

These statements occur in prominent passages in texts published (or meant to be published) by Kant to introduce the conception of analyticity to his audience. Kant, quite unambiguously, sets forth *one* conception of analyticity which is usually called the concept-containment conception and is roughly the following: The predicate is implicitly contained in the subject, and this containment is made explicit by the predication. Other explications—e.g., those presenting the principle of contradiction as a criterion, or those reflecting upon the way we come to know an analytic truth—occur in other places, but should not puzzle us;⁴ they are usually given not in places where Kant *introduces* the concept, but in the course of specific trains of thought where he counts analyticity as an already established notion; their position in the argument always clearly indicates that Kant considers them as derivative characterizations of analyticity. Our task is to first understand the primary explication of concept-containment, and then comprehend what follows from it.

This is not to say that Kant's introduction of the primary notion in the passages quoted is crystal-clear. The explication quoted from the *Critique*, e.g., is immediately followed by the statement that analyticity is a peculiar "identity," something that is at first blush unintelligible, even given the rudimentary logical facilities of Kant's time. But, again, Kant

clearly indicates that this is only meant to further elucidate the original exposition; thus, we should ignore the additional explanation and attend to the original exposition.⁵

Looking closer into the passages expounding the concept-containment conception, we find that this conception has three features which deserve special attention: (1) It is spelled out entirely in terms of one form of judgment, the subject-predicate form. (2) There is an important time component on the side of the thinker, namely that something thought *prior* to the judgment in question is made explicit in it. (3) The essential feature is that something implicitly contained in the “subject” is explicitly predicated of that “subject”; the nature of the relation is conveyed by two pairs of metaphor: inside—outside, and obscure [*dunkel*]*—*clear.

(1) At least from Frege's *Foundations of Arithmetic* onward, Kant has been criticized for this implausible limitation. The analytic-synthetic distinction is meant to describe a very general feature of human judgment, namely that it is either purely explicative or genuinely informative. Now, without argument, Kant seems to limit analyticity to a very special kind of judgment, those that have the subject-predicate form. Such limitation would certainly require additional argument on Kant's side, but it is missing and it is not even clear what he limits himself to.

Given Kant's taxonomy of judgments (categorical—hypothetical—disjunctive) in his famous table of judgments, subject-predicate judgments are categorical judgments. We now have the choice to interpret the analytic-synthetic distinction to “apply unrestrictedly . . . to categorical propositions,”⁶ or, as Frege suggested, only to *generalized* categorical propositions.⁷ As already quoted, in the opening sentence of the relevant paragraph in the *Critique* Kant explicitly says: “In *all* judgments in which the relation of a subject to the predicate is thought,” that relation is of a kind as to make the judgment either analytic or synthetic, he thus refers explicitly to all categorical judgments. On the other hand, Kant's paradigmatic example, “All bodies are extended,” for which he even gives a logical analysis in his logic lectures, is a generalized categorical judgment.

So Kant himself does not unambiguously indicate that he wishes to be understood in the restricted Fregean sense. Nevertheless, we should follow Frege's suggestion, for the following reason. Consider Kant's own explanation in the *Critique* that in an analytic proposition “I must not go outside” my subject concept and “must not transcend” it. Kant repeatedly emphasizes this feature in the following paragraph, to show that in an analytic judgment I need not take recourse to experience in order to judge.

The suggestion is that I need not investigate the object (to which, initially, the subject concept applies) and see whether the predicate concept applies. This is the meaning of the sentence at B11 which says: "For it would be inconsistent to base an analytic judgment on experience, since *I must not go outside my concept in order to frame the judgment . . .*" Consider now a singular categorical statement. Examples of singular statements are scarce in Kant's work, but two workable examples are: "The rose is red" and "This man is black"⁸—they constitute what Kant would call a categorical judgment and what we nowadays call a singular predication. In such examples (which are synthetic, of course) the subject concepts ("rose," "man") are parts of names or singular terms that have the semantic function of specifying individual objects. It is clear that "going outside" a concept in a way illegitimate for an analytic statement is going into experience; picking out an individual object, thus, seems to be the paradigmatic case of going outside my own concept and into experience. Now, if these statements, by virtue of their form, were admissible candidates for the analytic-synthetic distinction, Kant's idea of forbidding, in an analytic judgment, any transcending of the subject concept would become unintelligible. But how is it possible to exclude these singular categorical statements? The fact that the examples quoted constitute, by virtue of content, synthetic judgments, clearly does not yield an argument. We could make up examples which, though formally similar to Kant's own examples, must count as analytic, such as "This rose is extended," or (to vary Quine's famous example) "This bachelor is unmarried." Excluding empirical concepts altogether seems implausible, since Kant himself admits analytic statements containing empirical concepts (cf. *Prolegomena*, 4:267).

The only way to resolve the difficulty is by way of a formal restriction, namely Frege's proposal. Generalized categorical propositions are, again following Frege, best analyzed as generalized indicative conditionals, and these generalized conditionals by their very form (viz., the presence of a quantifier) indicate that no reference to an individual empirical object is made in them. The interpretation also tallies with Kant's own example of an analytic judgment involving empirical concepts, namely: "Gold is a yellow metal" (*Prolegomena*, 4:267). "Gold" is a mass term, not a name, and the statement is best rendered as follows: For anything, if it is gold, it is a yellow metal.

I say that analytic statements in Kant's sense are best understood as not containing a singular term in the subject position, but I do not say that

synthetic statements essentially do so. A synthetic statement may indeed be formally identical to an analytic one, since both may be without singular terms. Kant's examples, "All bodies are extended" and "All bodies are heavy," aptly illustrate this fact. Thus, what Kant maintains by means of these examples (at A8 as well as at B11 f.) is different from what I say here. Kant argues that for the second judgment, I start with the concept of body, and then need to consult experience, then finally return to my general reasoning to fully frame a judgment. It is this consultation of experience, after possessing the concept, that constitutes the transcending of the concept. What I have argued is that picking out some individual empirical object by a name or singular term constructed from a concept is another way of transcending the concept which is similarly opposed to Kant's purpose. This is the main argument to exclude singular predications from considerations of analyticity, in accord with Frege's proposal.

A second feature that underlines the present interpretation is the fact that Kant sometimes says, misleadingly, that the analytic judgment is about, or has as its subject matter, the subject concept itself.⁹ Some have suggested, for linguistically framed judgments, i.e., statements, that if they are analytic, they are about words.¹⁰ This is obviously false. Nothing in an analytic statement indicates the use of meta-language, and nothing indicates that in an analytic judgment the subject concept *refers* to itself. But Kant's slip is explainable. Given the implicit question of what an analytic statement is about, Kant shies away from mentioning any individual object (since it has no individual object as a subject matter), and turns, instead, to the concept.¹¹

(2) As stated above, Kant's own picture of how a judgment which is possibly analytic is formed bears an essential time-component: A thinker is said to possess a concept *before* judging; *then* he either analyses or transcends it in order to add new empirical information; *finally* he frames the judgment in question. This is not accidental. Kant repeatedly formulates the relevant issues in this manner, saying that the predicate in an analytic statement explicates what was "*already* thought," though perhaps obscurely, in the subject concept (see A8/B11). As Kant also puts it, in an analytic judgment "the concept *which I already possess*" is analyzed (A8).

The temporal succession signals, first of all, a constitutive dependence: An analytic judgment depends constitutively on the possession of a concept (the subject concept) by the thinker, in the sense that such possession is a necessary condition of the existence of the judgment. This seems

to express nothing but the platitude that judgments, being constituted by concepts, depend on these constituents. Thus, temporal succession is a mere metaphor for a constitutive dependence.

This, however, is not what Kant intended. The temporal succession is to be taken here as literally true. Consider, for a contrast, the famous opening passage of the Transcendental Logic in the first *Critique* where Kant says that “through intuition . . . an object is given to us,” and “through the understanding” *this very object* “is thought” (A50/B74). Obviously, Kant defends a constitutive dependence of understanding on intuition (insofar as being given an object is necessary in order to think it). Similarly, a full-fledged judgment constitutively depends on the possession of (at least) one intuition and (at least) one concept. But in both cases it would be mistaken to read the dependence as a succession in time. In the present case, however, a real succession in time must be assumed, at least in the sense that the forming of an analytic judgment presupposes possessing a concept, produced through a *prior* judgment. Given the intuitively plausible condition that a thinker cannot simultaneously produce two judgments, it follows that the prior possession Kant talks about is literally the *earlier* possession of a concept.

Hence, the temporal component in Kant’s exposition signifies that the analytic judgment analyses a concept which the thinker possessed from an earlier judgment. This implies a permanence of the concept beyond the actual existence in the mind of the judgment of which it is a part. But to view a concept as a permanent entity in this sense would create a severe tension between one aspect of Kant’s epistemology and another more closely connected with his philosophy of mind. That latter aspect of the Kantian doctrine includes as the central tenet that there is nothing permanent in the mind.¹² *A fortiori*, a thinker cannot possess a concept (or a judgment, for that matter) permanently. In some epistemological reflections Kant underlines that the concepts a thinker uses in a judgment, *qua* form, always originate with him. They are his creations just as the produced judgment is.¹³ Thus, neither concepts nor judgments enjoy permanence in a thinker’s mind; they come and go.

One way to defuse the tension is to propose a theory where a thinker is, at will, able to produce a concept matching an earlier one in content. Though the new concept is numerically different from the earlier one, it equals it in content. Kant’s figure of speech that the analytic judgment explicates a concept “which I already possess” must be understood thus: The thinker, for the purpose of an analytic judgment, forms a new con-

cept equal in content to another formed earlier in the context of another judgment. Thus, if a thinker is able to deliberately produce concepts equal or different in content to others produced earlier, it is a metaphor, strictly speaking, to say that he permanently possesses a concept or modifies a previously acquired concept. With this proviso, we can say that, according to Kant, an analytic judgment analyzes a concept used by the thinker in an earlier act of cognition.

(3) Kant fleshes out his containment conception of analyticity with the help of two metaphors, inside—outside, obscure—clear, which we need to investigate in turn. Kant, especially at A6/B10, uses the metaphor of inside and outside, extensively. In an analytic judgment the predicate concept is “*contained*” in the subject concept; the thinker “breaks up” the subject concept “into those constituent concepts that have all along been thought *in it*” (A7/B11); “in framing the judgment, I must not go *outside* my concept” (B11), since “I have already *in* the concept all the conditions required for my judgment. I have only to *extract* from it . . . the required predicate” (B12). In a synthetic judgment the exact opposite holds: the predicate concept “*lies outside*” the subject concept (A6/B10), and “I do not *include*” the former in the latter (A8/B12), i.e., one “is *not contained* in the other” (B12).

First of all, this string of metaphors is an application of the form-content distinction. Concepts have a content, hence, they must have form and content.¹⁴ Another feature of concepts—one being contained in the other—is easily understood on this basis. Of course, in any judgment of the form in question the “predicate concept” must “lie outside” the “subject concept.” Regardless of whether the judgment is a mere thought or is linguistically framed, its internal structure must guarantee a recognizable difference of the concepts or predicates, involved as constituent parts. In this sense, as a constituent part of the thought or statement, one concept is never “contained” or “included” in the other. But this, again, is a purely formal matter. The two concepts are different entities, but one can have (and in the case of analyticity, does have) a content equal to the whole or a part of the other. Accordingly, the possibility of extracting a concept from another is construed thus: The thinker forms a new concept the content of which equals the whole or a part of the content of the first concept. What matters is that the thinker must have access to his own conceptual content, but, as I have argued above, without explicitly referring to that concept or its content in the final judgment. Kant depicts this ability of a thinker as a scrutiny of the concept—but, as I have already pointed out, he

is wrong in doing so insofar as an analytic judgment has neither the subject concept nor its content as the subject matter (although it may be logically equivalent to another judgment of a different form that *has* the concept as a subject matter). He is correct, however, insofar as the thinker must have access to the conceptual content.

What does the second pair of metaphors, obscure—clear, add to this picture? The content of a concept obviously must be such as to be possibly complex, and that content must be of a kind as to be not always fully apparent to the thinker. Both features are remarkable. That a concept can have a complex content, i.e., can have different characteristics [*Merkmale*], is always presupposed by Kant when he reflects on concepts. Familiarity with our own thought activities as well as linguistic experience suggest that we follow him in this respect; we know that there are complex concepts and predicates. Nevertheless, this feature is again nontrivial considering Kant's transcendental project to synthesize, in philosophical reflection, the units of empirical and nonempirical thought from their constituent parts. This project must, first, reconstruct the primitive units of experience where an object is originally cognized by way of the original formation of a concept. As we will presently see, the formation of a complex concept presupposes full cognition of a plurality of objects. Thus, the original formation of a concept in one of these cognitions cannot result in a complex concept. Kant sometimes seems to allude to this when he says that, in the case of a visual experience of "a dog," the schematized concept gives us "the figure" of the animal, not a collection of characteristics like "four-footed" (A141/B180).¹⁵ And he expressly says that a concept relates to an object "mediately by means of a feature" [*Merkmal*] (A320/B376), using the singular explicitly.¹⁶ Even if the concept contains a characteristic by means of which it could apply to "several things" (A320/B376), we cannot see how it could contain *several* such characteristics, since, as Kant indicates, no more than one characteristic is necessary for a thinker in order to conceptually relate to an object. Now that possibility of a concept being complex is essential, of course, for an understanding of analyticity, at least for an understanding of those analytic judgments that are truly explicative and not mere tautologies.¹⁷ So, any full construal of Kantian analyticity requires a theory of the cognitive acquisition of complex concepts.

Second, the obscurity—clarity metaphor means that a thinker can possess a complex concept that is not fully transparent to him in a given thought process. This is surprising, since we should expect a complex

concept to be nothing but a collection or string of characteristics that, as with all concepts, are the thinker's own products and have been collected by him. It is *prima facie* unintelligible how, if the constituent parts of a concept have been produced by a thinker and have been synthesized by him, any of these parts can be hidden from that thinker. Of course, Kant's intuitive presupposition that this can be the case is correct. I can explicate a concept to myself that I used all along with all its characteristics, but without any longer taking notice of these characteristics. This has, however, an interesting consequence. Thinkers must form abbreviations of collections or strings of characteristics when organizing their own conceptual contents, just as a linguistic community forms abbreviations of collections of predicates such as the word 'bachelor' for 'unmarried male person'. This feature yields, thus, a welcome argument for the existence of a 'language of thought', at least as far as concepts are concerned. (This consequence is welcome, since the strategy of this paper is to argue away any substantial difference between thought structure and linguistic structure.)

We now have a good idea of what must be provided for a full understanding of analyticity: a theory of cognitive acquisition of complex general concepts. Indeed, Kant discusses a theory of the acquisition of general concepts in his logic lectures, but in quite a sketchy way that lacks much originality. He, by and large, subscribes to the traditional abstraction theory inherited from Aristotle. The preceding reflections on analyticity, however, will provide a basis for interpreting Kant's sketchy remarks.¹⁸

Kant presents a theory of concept formation where a thinker compares different objects, abstracts from their differences and "attends only to what they have in common" (24:566; cf. 9:94). He says explicitly that "these operations of the understanding are the essential and general conditions of the formation of any concept" (9:94). So, he claims to be presenting a fundamental theory of the origin of concepts. As such, this theory has been criticized as running counter to the transcendental project on the following ground. Kant's project, as indicated above, wishes to reconstruct empirical cognition *ab ovo*, by first reconstructing the most primitive cognitive encounter a thinker has with objects. Even such a primitive cognition must already involve intuitive and conceptual elements. If a theory of concept formation, which presupposes the awareness of different objects, is presented, initial contact is already presupposed. Thus the initial concept formation involved in a primitive object cognition is also presupposed.¹⁹

Correct as this criticism may be, it is irrelevant for our present purpose, for even if “abstraction” is not fundamental for the production of a concept the idea doubtlessly catches part of what happens in concept production—if not in an elementary act of cognition, then at least in ‘higher’, more complex thought processes. Indeed, the situation should be interpreted thus: On the one hand, the traditional abstraction theory, insofar as it implicitly presupposes the possession of concepts, is useless as a fundamental theory of the mental origin of concepts. Such a fundamental theory—as Kant himself notices only in late reflections—would not presuppose that concepts are given to us by the objects, but rather would make intelligible that “the concepts make that through them [*an ihnen*] we have objects of cognition.”²⁰ On the other hand, construed so fundamentally, concepts would have to be ‘individual representations’—as Kant takes intuitions to be ‘individual representations’.²¹ The nearly trivial fact that a concept (*qua* content) *can* apply to more than one object remains unexplained along these lines. Even if we had a fundamental theory of concept origin, that theory would be one of concepts as *individual* concepts, and we would still be in need of an additional theory of the origin of concepts as *general* concepts.²² This additional theory of how a general representation comes into existence given individual ones is what Kant has in mind whenever he reflects about the origin of concepts (cf. 8:93–4); and for this purpose, the traditional abstraction theory is a more plausible candidate.

Kant offers a theory of “comparison, reflection, and abstraction,” through the following simple example:

I see, e.g., a spruce, a willow, and a lime. By comparing, first of all, these objects to each other, I notice that they differ from one another in respect to the trunk, branches, leaves, and so on; now, I reflect only on that which they have in common, the trunk, branches, leaves themselves and abstract from their magnitude, shape, and so on; in this way I get a concept of a tree.²³

Again, three peculiar features should be pointed out. First, Kant is presenting a brief theory of the acquisition of a *complex* general concept. If we follow Kant, the concept ‘tree’, after the procedure described, has at least three characteristics: ‘having a trunk’, ‘having branches’, and ‘having leaves’. This is a superfluous complication. Second, Kant’s proposal is insufficient as a theory of concept origin, even if it is interpreted as an account of the origin of general concepts. The reason is that the first step,

i.e., comparison of different objects, is unintelligible without the thinker already possessing general concepts. Judging that two objects differ in a certain respect is, presumably, to judge that one does fall under a certain concept while the other does not. Hence, a concept, in order to be possibly applicable to more than one object, must be a general concept already; and this theory is, thus, inherently circular. Third—and this is the good news—the final activity of abstraction, despite the mentioned flaws, can be interpreted as a suggested theory for the acquisition of a general concept. Abstraction means that the thinker mentally manipulates a given conceptual content (of an individual concept) and consciously impoverishes that content. And this is what a thinker must do in order to arrive at a concept which, *qua* having a certain content, can apply to more than one object.²⁴

I do not wish to flesh out this idea into a full theory of general concepts, but only to note two consequences that matter in the present context. First, when for whatever reason a thinker manipulates a conceptual content, the resulting concept is used in the very same way in which, according to Kant, any concept is used; that is, in order to “determine” something.²⁵ Indeed, only a general concept is fit for “subsumption,” and “subsumption” is both the paradigmatic activity of the mind in forming judgments and the paradigmatic explication of what Kant means when he characterizes the activities of the understanding as acts of “determination.”

Second, Kant's example illustrates that a thinker's possession of a complex general concept reflects its cognitive history. The thinker possessing such a concept has consciously framed the concept by composing it from simple characteristics, and he is able to use it in acts of subsuming a thing under a complex general concept by checking properties that correspond to the concept's characteristics. (Schematically: ‘This thing has a trunk, branches, and leaves. So, it is a tree.’)

Given the possibility of subsumption, the thinker is able to recognize an object by his criteria; he is able to judge whether a given object falls under a given concept. The thinker now could explicitly give what Kant might call a “real definition” of the concept, a collection of characteristics necessary and sufficient to subsume an object under it (see A241n, 9:143–4). Kant has, of course, serious reservations about calling the analysis of any empirical concept a “definition” (see 9:141–2). This is because he requires, reasonably enough, that any definition of a concept to be complete and precise, but “the limits of the [empirical] concept never are assured.” Kant immediately goes on to ask: “And indeed what useful pur-

pose could be served by defining an empirical concept . . . ?” (A728/B756). However, in this passage Kant’s argument makes explicit use of *several persons* who have different characterizations of the concept “gold.” In order for several persons to perceive the fuzzy limits of a concept, there obviously must be communication between them, and the concept must be linguistically framed, e.g., as the word “gold.” For one sole thinker, a set of characteristics collected in a concept of “gold” has initially a purpose different than communication. It serves the thinker as a collection of what he regards as necessary and sufficient characteristics to subsume an object under the concept. The thinker is able to give what I propose to call a “cognitive definition” of his own private empirical concept.

An explicit definition of ‘cognitive definition’ of a concept might be proposed in terms of the concept’s cognitive role for a thinker. A concept *A* cognitively defines another concept *B*, if *A* is a nontrivial collection of characteristics that jointly could perform the cognitive role *B* performs. (Note for future reference the double asymmetry in this definition: *A* necessarily is a nontrivial collection of characteristics, while *B* is not necessarily one. A certain cognitive role *is actually* performed by *B*, but *could* instead be performed by *A*.) The explicit form of such a definition is such that explicit mention is made of *A* and *B*; the concept *B* performs a cognitive role which could be performed by concept *A*.

Some final clarifications are in order. Does a thinker who possesses a complex general concept thereby possess a cognitive definition? And is the latter identical to an analytic judgment? To both questions the answer is No. First, a thinker who possesses such a concept and uses it for subsuming something under it can be said to possess a rule which he follows. In fact, it is this activity of subsumption which leads Kant to call the understanding “the faculty of rules.”²⁶ The cognitive definition, explicitly formulated, is what makes this rule explicit to the thinker himself; yet a thinker can follow a rule without being explicitly aware of it. Second, the explicit definition or rule has an analytic judgment corresponding to it, but they are not identical. The definition or rule explicitly refers to the concepts *A* and *B*, something which the analytic judgment, as I maintained, does not.

Even if we could not reconstruct a fundamental theory of concept production in Kantian terms, the analysis provides us with a fairly complete picture of analyticity. We are able, first, to give a concise definition. An analytic judgment is one which, if every concept contained in it is replaced by its cognitive *definiendum*, transforms into a logical truth. Sec-

ond, we see more clearly the role of the temporal component in analytic judgments suggested by Kant. We also see that the cognitive history of acquisition of a concept explicated in a cognitive definition or, indirectly, in an analytic judgment, provides an understanding of *why* a concrete analytic judgment is necessarily true. A thinker must have used a characteristic contained in concept *A* in order to subsume an object under *A*, and the analytic judgment, after the fact, unfolds this cognitive procedure. I have concentrated here on empirical cases where such analyticity is relative to an individual thinker. Yet an obvious direction of generalization comes to mind as well. An analytic judgment may reflect that *every* thinker, in order to subsume an object under concept *A*, must turn to the characteristics contained in *A*. Kant primarily takes interest in these latter cases and locates them in mathematics and philosophy.

Finally, we note that analyticity, apart from trivial cases, is never a matter of pure form. We need to have access to conceptual *contents* in order to judge whether a judgment is analytically true or not. Kant acknowledges this himself in one of the above-mentioned explications of analyticity: "But whatever be their origin or logical form, there is nevertheless a distinction in judgments as to their *content*" (4:266).²⁷

III. Analyticity and the Semantics of Predicates

In the introductory section I have suggested that a reconstruction of Kantian analyticity could shed light on key issues in present-day semantics, especially with respect to the meaning of predicates. This conjecture was grounded in the fact that analyticity is generally construed—by its adherents as well as its critics—as the truth of a proposition solely in virtue of the meaning of its nonlogical constituents (which on our analysis turned out to be nothing but predicates). Thus, a reconstruction of analyticity as originally conceived by Kant harbors the hope of clarifying what predicate meaning is.

We recall that in modern analytic philosophy, from Carnap onwards, what sparked the debate about analyticity was an interest in predicate meaning. Quine's famous skepticism about analyticity ultimately is one about the clear definability of linguistic meaning. The thrust of Quine's argument, though, is that analyticity is to be defined in terms of the cognitive synonymy of predicates: A statement is analytic if, on the replacement of predicates by others which are cognitively synonymous, it turns

into a logical truth. Quine suggests that we cannot come up with a better explication for cognitive synonymy than interchangeability *salva veritate*, which is a demonstrably insufficient proposal.²⁸

It is quite remarkable that Quine labels the required conception of synonymy a *cognitive* one. Indeed, the Kantian account of analyticity is a cognitive conception, and it can help to dispel many of Quine's doubts. The proposal of the cognitive definition is more substantial than a mere affirmation of cognitive synonymy.

It is not clear initially what 'cognitive synonymy' and 'cognitive definition' should mean in the case of linguistic expressions. I propose the following clarifications: An expression as used by a certain thinker *qua* speaker expresses a certain concept performing a certain cognitive role in the thinker's thinking. Two expressions *A* and *B* are cognitively synonymous, if the cognitive role performed by the concept expressed by *A* can be performed by the concept expressed by *B*. An expression *A* is a cognitive definition of another *B*, if *A* expresses a nontrivial collection of concepts which jointly could perform the cognitive role of the concept expressed by *B*.

Thus defined, the cognitive definition of one expression by another has two notable features: Asymmetry and speaker/thinker relevance. (1) Cognitive definition, like Kant's original concept-containment, but unlike synonymy and interchangeability, is an asymmetric relation between expressions. This is signposted, formally, by the difference of *definiens* and *definiendum*, and—in the content of the definition—by 'a nontrivial collection' and the difference between 'actually *A*' and 'possibly *B*'. This preserves an important intuition about analytic statements, namely that they are *explicative*. We regard statements as analytic where the second ('predicate') concept explicates the first ('subject') concept, but not *vice versa*. Our analysis, so far, has left open the question of whether a cognitive definition can, or must be, or sometimes is, complete, or full-fledged, or final (thus, is a definition in Kant's strict sense). Assume that such a full cognitive definition is possible for a concept *B*. Then a full cognitive definition of *B* in terms of *A* entails synonymy and interchangeability *salva veritate*, but not *vice versa*. The reason is that even in this ideal case, there is an inbuilt asymmetry between expressions *A* and *B*, *B* being the *definiendum* (in the definition) or *explicandum* expression (in the corresponding analytic statement), *A* being the *definiens* (in the definition) or *explicans* expression (in the analytic statement). This asymmetry in the semantics of explication or definition carries over to our understand-

ing of what an analytic statement does, namely explicate something; but in attempts to capture analyticity by symmetric relations of expressions like synonymy and interchangeability this feature gets lost. This is a good indicator that 'explication by cognitive definition' captures the fundamental conception, while cognitive synonymy and interchangeability are derivative notions.

Note, furthermore, that in empirical cases cognitive definitions are hardly ever complete, but nevertheless give rise to statements that we think are analytic. In fact, we often regard statements as analytic that only partly unpack the subject concept, such as "Every mother is female" or (to use Quine's skeptical example) "Everything green is extended." The characteristic explicated in the expression in 'predicate position' is properly "contained" or is a proper subset of the characteristics defining the expression in 'subject position'. (Note that the set-theoretic formulation makes explicit Kant's original intuition of 'concept containment'.) Here both expressions are neither synonymous, nor interchangeable, nor does the latter fully define the former. The latter is, however, part of the definition of the former. So the idea of definition naturally extends to these imperfect cases, while symmetric proposals have nothing to say about them.

(2) What the definition of 'cognitive definition' expresses is the cognitive relevance of the characteristics contained in a concept for a speaker *qua* thinker who uses that concept. The above conception of cognitive definitions in judgments was that a thinker possesses concepts which are his own private abbreviations of collections of concepts. Let a concept *A* abbreviate a collection *A'*; the thinker will then regard concepts *B*, *C*, *D*, . . . as individually necessary and jointly sufficient for subsumption of an object under *A*, because *B*, *C*, *D*, . . . are the parts of the collection *A'*. The thinker *qua* speaker will assent to a sentence purporting to express an analytic truth, if he regards the concepts to stand in this relation: that subsumption of an object under one is necessary for subsumption under the other. Analyticity of judgments is, at least insofar as empirical matters are concerned, relative to individual thinkers. A thinker will in principle be able to judge whether a property is necessary for judging another or not. For speakers the matter is more complicated. An expression may be used by a speaker to express an observational concept that is general, but simple. Assume, for example, that a speaker uses 'dog' to point out something which has the typical dog-shape. The same expression can be used, however, to express a subsumption concept which contains 'four-

footed' as a characteristic.²⁹ This difference in use will not show up on the linguistic surface and will not be perceptible to the audience. This is the reason, I believe, why generally acceptable examples of analyticity involving empirical concepts come from cases where the complex property cannot be attributed on the basis of direct observation, e.g.: "Mothers are female," or "Bachelors are unmarried." By contrast, "Dogs are four-footed" is doubtful as a candidate for an analytic statement.³⁰

In sum, the Kantian conception of analyticity carries over straightforwardly from judgments to declarative sentences as linguistically framed judgments. What we learn from it, in the first place, is a clear conception of how predicates which are syntactically simple (e.g., 'a mother') can be semantically complex (e.g., 'a female parent'). This complexity is one of content, and a predicate has a complex content, if it expresses a complex concept. A complex concept, in turn, is one which the speaker *qua* thinker forms for some cognitive purpose. A complex concept—like any other concept—is a thinker's cognitive achievement. Thus, if we ask what complex concept an expression represents, whether it stands for a complex concept at all, the prime authority is the speaker *qua* thinker. Hence, speaker-meaning is decidedly prior to expression-meaning. A speaker may fail to successfully communicate what he means, if his own meaning deviates too far from conventions, or if he cannot, by auxiliary means, convey the deviations to the audience, but these complications do not change the primacy of speaker-meaning for the use of predicates.

More interesting, however, is what we learn from our brief study of Kantian analyticity about predicates *qua* semantic forms. We have seen that making sense of analyticity as a semantic category requires to view speakers as thinkers and view their linguistic products as expressions of their cognitive achievements. An analytic statement unfolds what a speaker *qua* thinker has put into a concept when producing it for some arbitrary cognitive purpose. This line of thought has important consequences for an understanding of the semantics of predicates *qua* predicate forms. Since predicates are expressions of concepts they express the latter's cognitive function. Put succinctly, predicates do for a statement what concepts do for a judgment.

Initially, the relation of predicates and concepts deserves further clarification. So far, it was unquestioningly presupposed that predicates are expressions of concepts. Does that mean that predicates and concepts are alike in content? My claim is that concepts *are* just the meanings of predicates, that predicates are concepts in linguistic form.

I have argued above that in analytic judgments concepts appear in their normal function, in the very same function in which they appear in, say, predications, despite the fact that analytic judgments have no subject matter, in the sense of essentially containing referential expressions.³¹ In fact, they appear in the very function in which they appear in the cognitive process of subsumption that the analytic judgment unfolds. That function, according to Kant, is nothing but determination. In fact, we can, with little interpretive strain, say that for Kant concepts are *determinators*.³²

This conception, I propose, should be carried over directly to predicates. Predicates are expressions of concepts *qua* determinators or, more straightforwardly, are expressed determinators. Thus, their function is to expressly determinate, just as the function of concepts is to determine. The sole argument I presently have to offer for this interpretational move is that it provides important semantic insights. Note, first, how it helps to understand the principal difference of names and predicates. Names never function to express determinators or expressly determine, but instead expressly indicate determinate objects; predicates never function to expressly indicate determinate objects, but instead expressly determine. The reason is that names and predicates express mental entities with different cognitive function: indicating something determinate and determining. Note some consequences of this apparently simple result: (1) Predicates do not name properties or refer to properties.³³ (2) The intuitive semantic difference between names and predicates cashes out as one of different semantic functions: referring and determining. (3) Names come out as the results of determination, since they indicate *determinate* objects. This has the consequence that a predication, as in Kant's example "This man is black," is to be analyzed as a further determination of some determinate entity, thus of some previously determined entity.³⁴ Accordingly, names, whatever they are, are the outcome of acts of determinations. Also, it seems that there must be predicates that determine without determining some pre-determined object. (4) Finally, the difference between an analytic statement and an expressed cognitive definition emerges more clearly. The definition contains terms *referring* to determinators, thus names for them; the analytic statement contains the determinators themselves, which in the statement actually perform their determining functions.

The conception of the meaning of predicates which results is a mentalist one: A predicate is a word token of a certain type in actual predicative use by an individual speaker. The meaning of that predicate is the concept the speaker *qua* thinker wants it to express. Such mentalism is, first of all,

congenial to the Kantian approach. As Kant in his epistemology famously reverses the primacy of empirical objects with respect to their being cognized, the present essay advocates, from a Kantian platform, that the primacy of a language as an abstract system of meanings of types of words with respect to concrete tokens of linguistic entities should be reversed. The present brief analysis of predicate meaning built on Kantian grounds even helps to defend such mentalist semantics more concretely. Mentalist approaches to semantics have been criticized on the grounds that the internal structure of propositions, especially the references of predicates to properties or natural kinds in them, cannot be put correctly by the mentalist.³⁵ In this respect, at least, mentalist semantics can immediately be defended. Brief analysis reveals that predicates never do refer to properties or kinds at all, thus the mentalist cannot be held to account for how, according to his approach, they can.

Let us at the end summarize (in reverse order) our two answers to the question concerning the meaning of a predicate. Understood formally, the question aims at a characterization of predicate-meaning as opposed to, say, singular-term-meaning or sentence-meaning. And it must be properly regarded as one for the different function of these different semantic particles. My proposal with respect to what predicate-meaning is formally has been to say that it is a concept *qua* determinator.

The question about the meaning of a predicate can also be posed in view of a concrete predicate and be understood as directed toward its content. The answer should, by now, be apparent to the reader. One could explicate the meaning of a predicate—and in cases of complex predicates this can be quite informative—by offering one's own cognitive definition of it (which in many cases is just the lexical definition). However, instead of mentioning the *definiens* and *definienda* concepts explicitly one might, equivalently, put the same matter in an analytic sentence.³⁶

Notes

1. I am thinking here of both the collapse of the analysis of meaning into truth-conditions in Davidson-style truth-conditional semantics and of a general meaning skepticism like Quine's. For an overview of both see, e.g., Stephan Schiffer, *Remnants of Meaning* (Cambridge: MIT Press, 1987).

2. See, e.g., the reference to Locke in Jerrold Katz, *Cogitations* (New York: Oxford University Press, 1986), chapter 4; the reference to Leibniz in Newton

Garver, "Analyticity and Grammar," in *Kant Studies Today*, ed. L.W. Beck (La Salle, Ill.: Open Court, 1969), 245–73; and the references to Locke, Leibniz, and Hume in George Bealer, *Quality and Concept* (Oxford: Clarendon Press, 1982), 211; for the view that Kant gave the analytic-synthetic distinction its unique frame and title, see Lewis White Beck, *Studies in the Philosophy of Kant* (Indianapolis: Bobbs-Merrill 1965), 75.

3. All translations from Kant's works as well as emphases in them are mine.

4. Such puzzlement is indicated, e.g., by Garver (*op. cit.*).

5. The sentence immediately following the quoted passage in the *Critique* exposes the identity characterization, but the derivative character is indicated by "therefore" [*also*]. See also Kant's unqualified characterization of analyticity in terms of identity in the *Jäsche-Logik* (9:111), and his attempt to repair the slip in the *Preisschrift* (20:322).

6. As, e.g., Garver (*op. cit.*, 248) does, apparently in accordance with the then current literature.

7. *Grundlagen der Arithmetik* (Breslau:Koebner, 1884), §88.

8. The first example is implicitly given at B70n, the second explicitly in a letter to J.S. Beck, July 3, 1792 (11:347); see also: "The air is elastic" in the *Prolegomena* (4:301), and "The tower is red" and "The stone is warm" in the *Jäsche-Logik* (9:113).

9. See explicitly *Reflexion* 3127 (16:671).

10. E.g., Garver (*op. cit.*, 267).

11. A tentative answer to the question of what an analytic statement is about would be the following. The question of what a piece of information is about may be understood as asking for the referent of the information, but not every piece of information need have a reference. So, the question of what the statement "Bill Clinton is on vacation" is about has the answer "Bill Clinton," but the question of what an analytic statement such as "Every mother is female" is about in this sense has no answer, like the question of what a statement such as "There are aardvaarks" is about has no answer, in this very sense. On the other hand, the question what a piece of information is about may be understood as asking about what the piece of information reports. In this case, the first statement is about Bill Clinton being on vacation, the second about every mother's being female, the third about aardvarks existing.

12. See, e.g., A381, B412–13. A similar point is made by Predrag Cicovacki, *Anamorphosis: Kant on Knowledge and Ignorance* (Lanham: University Press of America, 1997), 94–5, though in the course of a quite different argument (see below, note 26).

13. In the *Critique* there is, as far as I see, only one passage where Kant explicitly says that all concepts are created by the thinker, namely A19/B33: "[F]rom the understanding arise [*entspringen*] concepts." Indirectly the same point can be defended by noticing that Kant regards the categories as the thinker's product (cf. e.g., A86/B118, B144) and empirical notions as categories "*in concreto*" (cf. e.g.,

A138/B177), but both points are open to interpretation. Important confirmations outside the *Critique* are found in Kant's logic lectures. In the *Jäsche-Logik* he says that concepts, as regards their content, can be "given" to or "made" by the thinker, but as regards form, they are always "made": "The form of a concept as a discursive idea is always made" [*Die Form eines Begriffs als einer discursiven Vorstellung ist jederzeit gemacht.*] (9:93; cf. 9:94, and 16:547). In the *Logik Pölitz* he defines the understanding as "the ability to make concepts" [*das Vermögen Begriffe zu machen*] (24:568).

14. This follows, given some premises which are open to dispute: (1), that 'content' and 'matter' are synonymous (see Henry Allison, *Kant's Transcendental Idealism* [New Haven: Yale University Press, 1983], 66–7; and Robert Pippin, *Kant's Theory of Form* [New Haven: Yale University Press, 1982], 120; but see Cicovacki, *op. cit.*, 97–8, for the opposite view); (2) that 'matter' and 'form', as concepts of reflection, always refer to the same entity (this is advocated by Kant in the key passage A266/B322); (3) that 'concept' is in itself a formal characterization of a mental entity.

15. The full sentence runs: "The concept 'dog' signifies a rule according to which my imagination can delineate the figure of a [certain] four-footed animal in a general manner." [Conjecture following Mellin and Erdmann; see the German *Critique* edition by Raymund Schmidt (Hamburg: Meiner, 1956).] Kant does not say explicitly what I interpret him to be intending: that the concept contains nothing as a *Merkmal* but a certain "figure," i.e., the characteristic 'doghood' complementing the typical visual shape.

16. Similarly, a concept and *one single* characteristic are discussed at B133 n.

17. For this distinction see the *Jäsche-Logik* (9:111).

18. The following three paragraphs summarize and improve parts of my paper "Analytizität," presented at the 9th International Kant Congress, 2000, in Berlin.

19. This criticism was first voiced by Kant's pupil J.S. Beck. It is discussed and repeated by Gerold Prauss, "Zu Kants Theorie kategorialen Gegenstandsbezugs," in *Kategorie und Kategorialität. Festschrift für Klaus Hartmann zum 65. Geburtstag*, ed. D. Koch, K. Bort (Würzburg: Königshausen & Neumann, 1990), 99–106; and by Edmund Heller, "Kant und J.S. Beck über Anschauung und Begriff," *Philosophisches Jahrbuch* 100:1993, 72–95.

20. *Reflexion* 5636; cf. *Reflexion* 5643 (quoted from Heller, *op. cit.*, 83).

21. This is the conclusion reached by Heller, following Beck's criticism of Kant (*op. cit.*, 83 n. 18). Kant indirectly suggests something similar when he reasons, in a logic reflection, how a thinker can, given a single intuition without the possibility of comparison, reach a concept by means of abstraction (*Reflexion* 2878; 16:556). Note also that it is a consequence of Kant's approach that the concept have exactly the *same* content as the corresponding intuition (cf. "homogeneous" [*gleichartig*] at A137/B176), not a *poorer* one (as the view—presented below—on concept formation in terms of abstraction suggests).

22. Recall the above proviso that every concept *qua* being a mental entity and thus non-persistent is an individual entity. What, then, does the difference of individual and general concept consist in? I suggest that a concept is an individual concept if it is, *qua* content, only fit for one object. I assume that our concepts of the faces of individual persons or of shapes of countries on maps are of this kind. Below, in the main text, I suggest that concepts are to be interpreted as “determinators.” In what sense, then, do individual concepts determine? I assume that they only determine, i.e., delimit, a shape against its background. This may be viewed as an interpretation of Kant’s idea that the understanding, by means of the concept, determines the corresponding intuition. (For references, see note 25 below.)

23. “*Ich sehe z. B. eine Fichte, eine Weide und eine Linde. Indem ich diese Gegenstände zuvörderst unter einander vergleiche, bemerke ich, daß sie von einander verschieden sind in Ansehung des Stammes, der Äste, der Blätter u. dgl. m.; nun reflectire ich aber hiernächst nur auf das, was sie unter sich gemein haben, den Stamm, die Äste, die Blätter selbst und abstrahiere von der Größe, der Figur derselben u.s.w.; so bekomme ich einen Begriff vom Baume*” (8:94–5; cf. 24:566–7).

24. Kant should have seen that abstraction, as the necessary procedure for the origin of a general concept, must precede any act of comparison. Strikingly, in the immediate neighborhood of the passage quoted, he discusses the formation of the general concept ‘red’, following the perception of a scarlet cloth, by mere abstraction (without mentioning comparison and reflection), but does not draw the obvious conclusion that abstraction is the necessary and sufficient activity for the formation of a general concept (cf. 9:95; 24:567).

25. See for the activity of the understanding, through a concept, as determining [*bestimmen*] something, e.g., Bxvii, B142, B168 f., A267/B322 f., B406, and A656/B684. I owe this list to Gerold Prauss, *Einführung in die Erkenntnistheorie* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1980), 68. Prauss points out that the notion is insufficiently suited for Kant’s epistemological purpose, since Kant takes sometimes an intuition, sometimes an object as that which is determined through the concept. Contrary to the solution proposed by Prauss, I understand Kant as reflecting on elementary cognition in the former cases (B142, B406), but on what he calls “subsumption” (of an object under a *general* concept, see, e.g., A137/B176) in the latter cases (B169, A656/B684).

26. See, e.g., A126 f., A132/B171, A158/B197, A299/B356, and A302/B359. Considering that following a nontrivial rule (or cognitive definition) of a complex general concept (in judgments by subsumption) presupposes the formation and application of individual concepts, plus the formation (by abstraction) of general concepts, plus the synthesis of several general characteristics into one complex concept, Kant’s own characterization of the understanding in general as the “fac-

ulty of rules” seems far off the mark. For a defense of Kant’s view on this point, see Cicovacki (*op. cit.*, 93 ff.).

27. Note that, on the surface, this is inconsistent with his other prominent characterizations of logic, e.g.: “General logic . . . abstracts from all *content of cognition* and looks to some external source, whatever that may be, for representations in order to transform them into concepts *which is an analytic process*” (A76/B102; emphasis added). Here analyticity seems to entail abstraction from content and thus is associated to formal logic, while in the quoted passage from the *Prolegomena*, analyticity involves regard to content, thus is divorced from formal logic. The tension disappears, however, if we say in modern terms that analyticity is a matter of content; it is, thus, not dealing merely with pure logical form, but involves semantic considerations as well. This, then, is just what Kant wants to say in the *Prolegomena* passage. In the *Critique* passage, however, he wishes to distinguish general logic (here including semantics), that has no content stemming from sensibility; and transcendental logic, that just has such content.

28. See his *From a Logical Point of View* (Cambridge: Harvard University Press, 1953), 20–46, and *Word and Object* (Cambridge: MIT Press 1960), 61–7.

29. It is perhaps not superfluous to remark in passing that the Kantian reservations about the permanence of concepts apply to linguistic expressions as well. As the repeated use of the same concept is nothing but a repeated production of instances of the same type of concept, so the repeated use of the same expression is nothing but a repeated production of instances of the same type of expression.

30. For a similar point, see Searle’s criticism of Quine’s “borderline” examples in *Speech Acts* (Cambridge: Cambridge University Press 1969), 8.

31. See also above, note 11.

32. See also above, note 25.

33. The idea that predicates are or can be interpreted as names seems to originate with Carnap, who argues that a predicate *can* be construed as naming either its extension or its intension; see his *Meaning and Necessity* (Chicago: University of Chicago Press, 1947), 106. Carnap says that a predicate can alternatively be construed as just *having* extension and intension, if a name is construed in the very same way; see 23, 40. Quine emphasizes that predicates do not name their extensions; see his *Philosophy of Logic* (Cambridge: Harvard University Press 1970), 67. Bealer stresses that predicates name neither properties nor functions (see Bealer, *op. cit.*, 86–7 n. 2), and rightly criticizes second-order logic as glossing over the difference of a predicate and its nominalization (which is a name of a property). The claim that predicates do not name at all seems to be implied by both, but I found it explicitly only in George Boolos, *Logic, Logic, and Logic* (Cambridge: Harvard University Press 1998), 38, where he interprets the mentioned passage from Quine.

34. With a detailed analysis of the black man example (see 11:347), but in a different context, this has been urged, by Prauss, “Zu Kants Theorie kategorialen Gegenstandsbezugs.”

35. This is the main argument in Schiffer (*op. cit.*, chapter 3; see esp. sections 3.2–3.5).

36. Special thanks for critical comments on this essay are due to Hans-Ulrich Baumgarten, Predrag Cicovacki, and Katharina Held.

Kant, the ‘I Think’, and Self-Awareness

Robert Howell

Self-awareness, and Kant’s treatment of it in the *Critique of Pure Reason*, are fascinating, difficult topics that have received much attention in the last three decades. It is impossible in a single essay to do justice to all the issues that arise. However, despite the flood of enlightening discussion, I think that crucial points remain ignored or underappreciated. Below, I try to bring out these points and some of their main implications. I hope that my efforts will serve as a tribute, however modest, to Lewis White Beck, a model of clarity and humanity and a scholar wonderfully receptive to work on Kant from many different standpoints.¹ I begin with a brief review of Kant’s account of knowledge and of self-awareness. Then I compare Kant’s view of such awareness with the modern idea of direct reference by means of “I”; it is in terms of the framework provided by that idea that, I will argue, the insights in Kant’s view are best appreciated. I conclude by considering two basic questions which this framework requires him, and us, to confront. Kant’s remarks on self-awareness suggest three answers to these questions, two of which are of great interest today.

I. Kant on Self-Awareness

Fundamental to the first *Critique* is the idea that beings like us know objects as single, individuated things that are of general types or kinds (“*this* thing before me is a *tree*”). In knowledge, we characteristically have a *de re* awareness of the object as having a general property that specifies the type or kind of which that object is an instance. Such awareness we of course achieve through intuition and concept. An intuition is a singular, immediate, and sensible representation that presents to the mind, in a *de re* fashion, the single, individuated object that we know. A concept is a general, mediate representation that, through activities of thought, comes to present to the mind (or comes itself to be) a general property.

Through its presentation of this property, the concept represents to the mind, in a *de dicto* fashion, the objects, whichever in particular they may be, that have this property. When we attain knowledge, the manifold in which the intuition is given is synthesized through thought and imagination, and we perceive the unified object presented by the intuition as having the property presented by the concept.²

Our intuitions fall into two groups. Outer-sense intuitions allow us to know objects as in space. Through inner-sense intuitions, we are made aware of various of the properties of our mind—including our thoughts, feelings, sensations, and outer-sense intuitions—as being in time. However, inner sense by itself gives us no awareness of such items as belonging to ourselves (or, for that matter, as belonging to any entity at all). To achieve that sort of awareness, we must use our understanding to think those items as so belonging, as being ours. We think them as being ours through the use of what Kant calls the representation *I think*. Through the *I think*, and only through it, we gain a first-person awareness of ourselves, a genuine self-awareness.

Kant's views of the *I think* and of the self-awareness that it yields are the main focus of this essay. However, in order to understand the depth and continuing plausibility of his views, it is important to recognize that they are a composite of positions. To what I will call Kant's core account of the *I think* and self-awareness, he attaches a number of other points that are philosophically fascinating but also highly questionable. The result is that the deep insights and originality of the core account are overlaid by famous Kantian phrases and doctrines that lead readers to miss the continuing relevance of the core or else to see in it only an early stage in the discovery of the truth. Yet if my conclusions are correct, it is not an exaggeration to say that Kant's core account is itself, in its most fundamental features, as close to the truth about first-person awareness as anyone has yet reached.

Kant's transcendental idealism is the main overlay to the core account and a main source of confusion about its interpretation. As everyone knows, Kant argues in the first *Critique* that we do not know objects as they are in themselves but only as they appear to us in spatiotemporal forms through our intuitions. It follows, in particular, that we never can attain a knowledge of our mind as it is in itself; we can know it (including our own thoughts, feelings, and so on) only as it appears to us in inner sense.

Yet, as I note below, Kant also wants to hold that, through the *I think* and the self-awareness that it yields, we designate the "subject of thoughts,"

a subject that—although nothing about its nature or properties is revealed through the *I think*—is not a mere appearance but “the being itself” (B429).³ This point means merely that through the *I think*, taken by itself, we gain no awareness or knowledge of anything about the nature of the self, existing in itself, with which the *I think* puts us into contact. The point allows that we might in some other way come to know a great deal about the self as it so exists. However, Kant’s transcendental idealism now operates to bar any knowledge of that self. The self that we somehow contact through the *I think* becomes the mysterious and unknowable “transcendental subject of thoughts = X” (A346/B404), “this I or he or it that thinks” (A346/B404), “the being which thinks in us” (A401). Attention is deflected away from Kant’s deep insight into the kind of representation afforded us by first-person self-awareness to the idea of a transcendental subject, and Kant is subsequently congratulated (or excoriated) for introducing that alluring notion. However, and as I think much patient research has shown, Kant fails to make a convincing case for his idealism.⁴ It is thus desirable to separate Kant’s specific, core claims about the *I think* from the further views he reaches once his idealism is accepted. (Moreover, even those who accept the idealism may find it enlightening to see what the core claims by themselves allow Kant to say about self-awareness.)

A subsidiary overlay to the core account is provided by various inferences that Kant draws from his view of the *I think*—for example, his inference (one form of which I note below) that the self cannot be known as an object because it is that through which objects are known. Some intelligent, alert readers—including some exceptionally good Kant scholars—have accepted these inferences, and there is no doubt that they are philosophically stimulating and deserve careful reflection. However, I find them very doubtful, at least in their popular forms. For that reason, I will mention them below only in passing.

If we set aside transcendental idealism and the inferences just mentioned, Kant’s core claims about the *I think* and self-awareness run as follows.⁵ (i) The *I think* is a spontaneously occurring act of thought that can—although it need not actually—accompany all our representations.⁶ (ii) The *I think*—or the “bare representation” *I* that it contains and that represents myself—is a simple representation (B135, B138, B68, A345–6/B404, A355, A443/B471, A784–5/B812–13, and many other places). As such, it is a “completely empty representation,” not even a concept but only “a bare consciousness which accompanies all concepts” (A346/B404).

As it contains no manifold of subrepresentations, (iii) the *I think* represents nothing about the properties or nature of the I of which we are aware through it, neither as that I appears to us in inner sense nor as it may be in itself:

In the synthetic original unity of apperception [through the *I think*], I am conscious of myself, not as I appear to myself nor as I am in myself, but only that I am (B157).

If [through the *I think* taken by itself] I would be conscious of myself simply as thinking, then since I am not considering how my own self may be given in [inner] intuition, the self may be mere appearance to me, the I that thinks, but is no mere appearance insofar as I think; in the consciousness of myself in mere thought I am the *being itself*, although nothing in myself is thereby given for thought (B429).⁷

Nevertheless, and although it thus represents no properties of the I of which we are aware through that representation, (iv) the *I think* does indeed represent (or “designate”) an entity (the self or I):

[I]n attaching *I* to our thoughts we designate [*bezeichnen*] the subject of inherence only transcendently, without noting in it any quality whatsoever—in fact, without knowing anything of it either by direct acquaintance or otherwise (A355).

However, because the *I think* involves no intuition, (v) that entity is not marked out for us in our experience as one particular object that we can distinguish cognitively from other particular objects with which we have to deal:

Thought, taken by itself [in the *I think*], . . . does not exhibit the subject of consciousness as appearance (B428–9).

[C]onsciousness [through the *I think*] in itself is not a representation distinguishing [*unterscheidet*] a particular [*besonderes*] object, but a form of representation in general [through which I can say that] I am thereby thinking something (A346/B404).

[T]he *I* is indeed in all thoughts, but there is not in this representation the least trace of intuition distinguishing [*unterschiede*] it from other objects of intuition (A350).

The fact that the entity represented or designated by the *I think* is not marked out in our experience as a particular object is supported also by point (ii) above, insofar as it will not be possible to distinguish this entity by means of properties ascribed to it by the *I think*.

On the basis of such facts, some commentators have supposed that the *I think* does not refer to or designate a subject of consciousness at all, but the quotations given under (iii) show that this interpretation is mistaken.⁸ One can also see the mistake by noting Kant's well-known remarks about my uniting a manifold of intuition through *I think*-accompaniment and thereby becoming "conscious of the self as identical in respect of the manifold of representations that are given to me in an intuition" (B135). As these remarks bring out, Kant's talk of the *I think* or *I* as designating an entity, *I*, is quite full-blooded. He is in effect committed (as the normal natural-language readings of "I" and "ich" require) to there being a single, self-identical entity that is referred to by the multiple occurrences of "I" in a claim like "I am conscious that I have intuition element *i* and I have intuition-element *j* and I have intuition-element *k*." It is simply that for him (and given points (ii), (iii), and (v)), (vi) through the *I think* we gain no insight into the nature or identity conditions of this entity as it is in itself in independence of its grasp through the *I think*. As far as our grasp of this entity through the *I think* goes, its identity is purely "formal" or, one might say, "functional." What, in itself, plays the role of the *I* on one occasion might be completely different from what plays that same role on another occasion.⁹ But the *I* itself (taken as the entity designated by the *I think* or *I*) remains the same formal entity through these (possibly) changing manifestations or role-players of it.

A further group of points about the *I think* is of great importance to Kant's treatment of self-awareness. He holds that the representation *I think* is an *original* apperception, in that

it is that self-consciousness which, while generating the representation *I think* (a representation which must be capable of accompanying all other representations), and which [self-consciousness] in all consciousness is one and the same, cannot itself be accompanied by any further representation (B132; note also A106–7).¹⁰

The interpretation of this passage is tricky. However, it should be read alongside texts like A544/B572 ("an *original* act, such as can *by itself* [my emphasis] bring about what did not exist before"; see also B72) while

taking into account Kant's stress, in the B132 sentence preceding the above quotation, on the spontaneity of the *I think*. When the passage is read in this way, it can be seen to be a main part of a crucial B-Deduction §16 claim about the occurrence (and the nature) of the *I think* as yielding self-awareness and accompanying the elements of the manifold of intuition (and as yielding the unity of the manifold that Kant argues to follow from that accompaniment). That claim is that (vii) the occurrence (and nature) of the *I think*, as yielding these things, is not to be explained in terms of, derived from, or reduced to, the occurrence of any further acts, representations, or other phenomena in our mind. No further explanation of the occurrence of the *I think* or analysis of its nature can be given; its existence is ultimate.¹¹

This stress on the originality and unanalyzability of the *I think* can be supported by two further, and related, Kantian points. At A346/B404 and elsewhere Kant holds in effect that (to state the first of these points) (viii) any understanding that we have of the I, the subject of thoughts—aside from our simple understanding of the I as that that has various specific thoughts as its predicates—can be got only through the representation *I*. Any attempt to understand the I without the use of that representation will either fail or (by covertly reintroducing the I into the explanans through the covert use of the representation *I*) will become circular.¹²

The second of the points can be developed by first considering, for a bit, comments that Kant makes at A402 and at B422. At A402 he argues that in order to know any object, I must presuppose the thinking I—the determining self that I am aware of when I accompany the representations through which I know by the *I* or *I think*.¹³ Again, at B422 he says somewhat similarly that the subject of the categories—the I of which I am aware via the *I think*—cannot through use of the categories come to conceive of itself as an object that is known (or conceived of adequately) through the categories. The reason, he says, is that in order for the I to use the categories to know (or to conceive) the I, the I will have already to presuppose its *I think*-arrived-at self-awareness. But it is that *I think* self-awareness that, as belonging to the putative object of knowledge here (the I that is the subject of the categories), the I is trying first to know through the I's use of the categories.¹⁴ Nor does the I have any special intuition that takes the thinking I of the *I think* as its own object and so lets the I know that thinking I and its nature like any other object of knowledge.¹⁵

Where the A402 reasoning just noted argues in terms of presupposing the thinking I, the B422 reasoning argues in terms of presupposing that

I's *I think*-yielded self-awareness. Hence (aside from the fact that B422, but not A402, mentions the categories) A402 and B422 do not make exactly the same claim.¹⁶ Moreover, neither the A402 nor the B422 reasoning is absolutely pellucid, and one can debate their exact interpretations at length. On one, perhaps superficial, reading these pieces of reasoning are indeed analogous to, and no better than, the bad argument that we can never discover an adequate theory of the origins and nature of language because we will have to use language to state any such theory. (On such a reading, the A402 and B422 reasoning are examples of the shaky inferences that, above, I excluded from Kant's core account of the *I think*.)¹⁷ It seems to me, however, that in his A402 and B422 reasoning Kant can be read as defending (or at least as pointing towards) the following more specific—and more plausible—claim, which is the second of the two further points that I mentioned above. (ix) The nature of the self-awareness that we arrive at via the *I think* (and the nature of the I that produces that self-awareness) cannot be explained in terms of, or reduced to, any objects or representations that can themselves be known through or accompanied by the *I think*; any attempt to achieve such an explanation or reduction will either fail or else will reintroduce the very *I think*-style self-awareness (or the very I) that we are trying to explain.¹⁸ Or so, for purposes of this paper, I will interpret the A402 and B422 claims.

Like all his philosophical views, Kant's account of self-awareness via the *I think* is rich and subtle, and we could pursue its details considerably further than I have above. Here, however, I want to note only one more explicit pair of points that he makes. First, while inner sense does not, by itself, present our self to us as being our own self, Kant nevertheless means to insist that (x) the self that thinks the *I think* is, in fact, the same subject, the same entity, as the self that appears in inner sense. Thus he does not take there to be two distinct subjects or selves, the empirical self or I that is an appearance in inner sense and the transcendental self or I that thinks and knows that empirical self. Rather, in inner sense "the mind intuits *itself* or *its* inner state" (A23/B37, my emphases); one and the same soul both is an object of experience (appears in inner sense) and has an existence in itself (Bxxvii). "*I* [am] an object to *myself*, an object of intuition and of inner perceptions" (B155–6, my emphasis); when I claim that "*I* exist thinking," I determine "the subject (which is then at the same time object [of inner sense])" (B429).

Second, Kant does not hold just that the I that thinks is the same entity as the self that appears to that I through inner sense. He also seems to

hold that (xi) when I think the thought *I think*, in that thought *itself* (and in independence of the operations of inner sense) I have a kind of reflexive awareness of that very thought as having my self as its subject.¹⁹ Thus we have already noted Kant's B429 claim that if [via the *I think*] "I want to become *conscious of myself only as thinking* [my emphasis], then . . . the self [as I grasp it] . . . is no mere appearance insofar as I think; in the consciousness of myself in mere thought I am *the being itself*."²⁰

In remarking points (x) and (xi), we are moving on difficult ground. For one thing, and as I have argued elsewhere, Kant holds two distinct views of the objects of knowledge. He holds an appearing theory, which takes there to be a single, selfsame group of entities (including our own self or mind) that both appear to us—through our outer- and inner-sense intuitions—as knowable spatial and temporal objects and also have an unknowable, nonspatiotemporal existence in themselves. It is such a theory in terms of which I have been interpreting Kant so far in this paper. However, Kant also holds an appearance theory, which takes the spatial and temporal objects of our knowledge (including our own self or mind as it is known through inner sense) to be appearances in intuition (roughly, synthesized intuitions themselves) that are distinct from the entities, having existence in themselves, that affect our mind and yield those appearances to us. Moreover, the interactions between these theories are quite complex and can be seen to affect Kant's view of self-awareness in complex ways.²¹

For another thing, the difficulties of the topic of self-awareness, and further complications in Kant's own theories, lead him in a number of places to speak as though, contrary to point (x), he recognizes at least two different selves, the thinking self and the known, inner-sense-appearing self that is thought. Thus in the extremely difficult B155 text (quoted in part above) he worries about "how the I that thinks can be distinct from the I that intuit itself (for I can represent still other modes of intuition as at least possible) and yet, as being the same subject, can be identical with the latter"; and in the B429 text just quoted, he speaks of the I that thinks as being "no mere appearance insofar as I think"—as though the I that thinks were distinct from the intuited I that we know through inner sense. Independently of such texts, the language that Kant uses elsewhere to describe the thinking self may raise questions as to how he can hold that the self that thinks is identical to the self that appears to me in inner sense.²²

In a thorough study of Kant's account of self-awareness it would be important to work through all these points, taking account of both ap-

pearing and appearance theories and trying to explain in detail Kant's occasional (apparent) language of two selves. However, the appearance theory can be seen to lead to intolerable difficulties for Kant, and the two-self language does not, I think, reflect any of the insights into self-awareness that his account embodies.²³ As my goal here is to focus on the points in that account that are most central to it and philosophically most enlightening, I will therefore ignore both the appearance-theory view of the self and the interpretive difficulties raised by the two-self language.

II. Direct Reference and the 'I Think'

We have now seen that, for Kant, we achieve self-awareness through the act or representation *I think* (or *I*). This representation can (but need not actually) accompany all our representations. It is a simple representation that represents nothing about the properties or nature of the I of which we are aware through it. Although it designates that entity, it does so in a way that does not mark that entity out as one particular object that is distinguishable cognitively for us from other particular objects. In particular, through the *I think* we have no insight into the nature of the I as it is in itself in independence of its grasp through the *I think*. Our *I think* awareness is original and cannot be explained in terms of or reduced to any group of further mental entities or relations among such entities; our understanding of the I can be got only through such irreducible *I think* awareness. Indeed, Kant holds, more specifically, that the nature of the *I think* awareness and of the I itself cannot be explained by means of, or reduced to, any objects or representations that are themselves known through or accompanied by the *I think*. The I that itself produces the *I think* awareness is, in fact, identical to the self that is known through that awareness as that self appears in inner sense. Moreover, in the thought that is my *I think* awareness, I have a reflexive awareness of that very thought as having myself as its subject.

We have now seen the basic points about self-awareness that Kant makes in his core account of the *I think*. Of course at this stage in our history any theory of such awareness (and so Kant's) is bound to be incomplete or somewhat speculative. Many basic facts about the nature and neurological underpinnings of consciousness (let alone about the nature of first-person consciousness) are simply not known; and, lacking knowledge of those facts, we must leave gaps in our theory in various places. Moreover,

if Kantian views about the irreducibility and unanalyzability of the *I think* turn out to be correct, the goal of at least certain sorts of reductive theories of such awareness will have to be abandoned. However, even with their inevitable gaps—and obvious obscurities—Kant's points provide the beginnings of as plausible a theory of self-awareness as anyone has so far proposed. Even after 200 years, they remain a deep, original attempt to comprehend the subject, an attempt from which one can still learn much.

This fact has been stressed in various ways by recent philosophers who have seen the relevance of Kant's views to their own accounts of self-reference and self-awareness, including Strawson, Castañeda, and Chisholm.²⁴ However, these philosophers do not develop all Kant's views in detail, and (as the Kantian ideas presented in the next section will show), I think that the interest of those views goes beyond the claims even of such writers. To my mind, the best way of appreciating the depth and insights of Kant's views is to note their resemblance to current claims about direct reference, "I," and self-awareness. I believe, in fact, that Kant's views are most adequately understood by recognizing that his account of the *I think* contains deep insights into the same sort of phenomena that contemporary direct-reference and related theorists see as present in linguistic reference by means of "I."²⁵ I will try to illustrate the power of that understanding below.

About direct reference, "I," and self-awareness three matters especially should be stressed.²⁶

(I) *Direct, indexical self-reference via the pronoun "I" (or "ich") and similar linguistic devices.* A definite description like "the man wearing a black hat and speaking slowly" achieves reference because its referent has (and is the unique entity that has) the properties that the general terms in that description express, of wearing a black hat and speaking slowly.²⁷ The referential relation here runs from the term to the referent by way of those properties; it is thus a mediate, indirect relation. Moreover, it is not that the properties function simply as means of determining the referent, that referent itself then occurring in the propositional content of what we say by using that term. Rather, the properties expressed by the term themselves make up part of that propositional content. For example, the propositional content of what I say by the statement "the man wearing a black hat and speaking slowly is alert" is made up of properties in standing in various logical relations. What I say is, in effect, that the unique man having the properties of wearing a black hat and speaking slowly also has the prop-

erty of being alert. What I say is not that the entity George Smith has that property, even if George Smith in fact is the unique man in question.

However, and as has become familiar since the work of Kripke, Kaplan, and others, it is possible in other cases for the referential relation to be direct and not mediated in the above way by properties. Thus, if Kripke is right in his Millian views, the proper name "Bill Clinton" refers to the current U. S. President not because of any properties that that name expresses or connotes and that Clinton has. Rather, that name so refers because Clinton's parents dubbed him with the name and its reference was then passed on from speaker to speaker. The reference that we make by using that name is direct. The facts of the dubbing and the chain of transmission determine the reference; but they are not part of the meaning of the name itself, and they do not occur in any proposition expressed by a sentence using the name. All that occurs in such a proposition is the entity Clinton. The statement "Bill Clinton is U. S. President" does not say that the name "Bill Clinton" was applied on such-and-such occasion and then passed on to you from other speakers. Rather, it simply makes a claim about the particular individual, Clinton.

Again, and to come to the main case of interest here, the reference of the pronoun "I" (or "*ich*") is governed by the following rule:²⁸

(R) In any statement in which it occurs, "I" refers to the statement's speaker.

The reference of "I" to the speaker is once more direct, although it turns on contextual factors in a way that the reference of proper names does not. While (R)—and the fact that I produce the relevant token of "I"—determines that I am the referent of that token, (R) and that fact about my production do not themselves occur within the propositional content of what I say using that token. For example, in my statement "I am happy that I am up and about," I attribute no properties to myself—including the property of being the speaker—merely by using "I" to refer to myself. Rather, the entity that I am is established as the referent of my use of "I" simply because I am the maker of that statement. That entity, I—and not the property of being the speaker—then is the sole element that my self-reference by means of "I" introduces into the proposition about my happiness. Although my being the speaker determines that my use of "I" refers to me, only I, and not the fact of my being the speaker, am part of what I say in making the statement.²⁹

(II) *"I" reference without observation, identification, or the possibility of misidentification.* Together with plausible epistemological views, rule (R) implies a group of points stressed in various forms by writers like Strawson and Shoemaker.³⁰ (R) guarantees that in any context in which a person produces "I" in the course of making a statement, the very fact of that production guarantees that "I" refers to that person. Hence in referring to himself via "I," a person does not, and so need not, observe himself to determine that he possesses properties associated with "I" (as he would observe someone to determine that that person possesses the properties required for successful reference by means of "the man wearing a black hat and speaking slowly"). Similarly, a person does not (and need not) identify herself in some way (for example, as this self or I) when she refers to herself by using "I." Nor, given (R), is it possible for her, using "I," to misidentify the entity to which she means to be referring. Nor can her usual self-reference via "I" (unlike her attempted reference via a definite description or a putative proper name) ever fail.

(III) *Irreducibility, unanalyzability, and ineliminability of "I."* Because reference via "I" is governed by (R), the distinctive referential use of "I" in a given statement cannot be reduced to, analyzed in terms of, or eliminated in favor of, the use of any other term that is not governed by (R)—or that is not governed by a rule whose use implies the satisfaction of (R) or of some rule analogous to (R).³¹ In particular, then, and as is well known, such uses of "I" cannot be analyzed in terms of (or eliminated in favor of) the usual sort of definite descriptions like "the person speaking (or thinking)." After all, the term "the person speaking" evaluated at a given context refers to the person, if any, in that context who has the property of speaking. But even if I now have that property, in some other context at which that term is evaluated (a context in which I still exist) I might not. In that other context the speaker might be someone else. ("It is possible that you and not I are the person speaking" is not inconsistent.) Similarly, "I" is obviously not analyzable as some proper name, say "Robert Howell"; my production of that name does not automatically guarantee that, by means of it, I refer to myself.

Facts like (I), (II), and (III) have become familiar in the recent literature on "I," although I do not think that their bearing on Kant's account of self-awareness has yet been fully appreciated by all Kant scholars.³² That bearing is in any case straightforward to rehearse. First, given (I) and rule (R), "I" in its referential use is "simple" in the sense that, as just seen, no properties are attributed to myself through that use (compare Kant's points

(ii) and (iii)), although through that use I designate an entity, myself (compare point (iv)). Yet, because, by (I), my use of "I" involves no attribution of properties to myself, the *mere* fact that I successfully refer to myself by means of "I" does not mark me out as one particular thing that I can distinguish cognitively from other particular things with which I have to deal (point (v)). Second, (II) itself of course leads to a similar conclusion; and it is clear also from (I) and (R), as well as from (II), that the mere fact that I successfully refer to myself by means of "I" does not give me any insight into the nature or identity conditions of myself as I exist in independence of making that reference to myself (point (vi)). Third, given our referential practice with "I," as noted in (I) and (R), the entity that I refer to by "I" is, in fact, the same entity as the entity of whose thoughts, sensations, and emotions I am made aware by whatever internal monitoring system I use (whether it is literally an inner *sense* or not) to know such things. Given (I) and (R), the two occurrences of "I" in the claim "I think (say, believe, know) that I am now thinking of New York, seeing a table, and feeling restless" refer to the same entity (point (x)).

Fourth, although this fact is not an immediate consequence of points (I) to (III), anyone who is in a position to make claims about the world (including her own inner experiences), such as "the book is on the table," is also in a position, except perhaps in certain special cases, to self-ascribe those claims by using "I" in the directly referential manner characterized by (I) to (III). She can always say "I think that the book is on the table," and so on. Thus that characterization allows for the analogue of Kant's usual view (as in (i) above) that it is possible, but not required, for the *I think* to accompany all my representations.³³

Fifth, I want to put off detailed comments on the relation of (III) to Kant's irreducibility claims in (vii), (viii), and (ix), since those claims raise complex questions that are best postponed until below. However, consider briefly the observation in (III) that, roughly, the referential use of "I" cannot be reduced to, analyzed in terms of, or eliminated in favor of, the use of any other term that is not governed by (R) or by some rule analogous to (R). It is at least clear immediately that, given this observation, the referential use of "I" is itself "original" in a linguistically Kantian sense. Without loss of its distinctive nature, that use cannot be reduced to, or analyzed in terms of, the use of any other referring term whose use does not imply the occurrence of the distinctively (R)-style reference or at least the occurrence of a reference achieved via a rule analogous to (R). This point is not, even at the linguistic level, as strong as Kant's own

originality claim about the *I think*. (After all, it seems possible that one might achieve the equivalent of “I” reference in some (R)-style way that does not involve “I” even if it turned out that one did not, thereby, preserve everything that is distinctive about our use of “I.”)³⁴ However, this point shows that the claims in (III) lead to results analogous to at least some of Kant’s views on the *I think* and its originality (points (vii), (viii), and (ix)).

III. Two Basic Questions and Three Kantian Answers

The crucial observation (I) makes it clear that self-reference via “I” sets up a directly referential relation to oneself that imports into the content of what one is saying only the entity that one is, without introducing into that content any properties of that entity or any of the means whereby that referential relation is set up. (When I state “I am a reader of Kant,” my use of “I” refers to myself, the maker of the statement; but the fact that I am the statement’s maker is not part of what I say in the statement itself.) We have just seen that Kant takes the self of which I am made aware through the *I think* to be introduced into the content of my thought in a similar way. Through my production of that representation—the *I think* is an act of thought, recall—I am brought into an awareness that designates myself, the entity that I am. About that entity I can then go on to make specific claims (“*I think that it will rain*”). But beyond that mere entity (“the being itself” of B429), through the *I* or *I think* taken by itself nothing about myself is given to or related to that state of awareness.³⁵

This comparison of Kant’s account of self-awareness with direct-reference ideas is illuminating as far as it goes, and it shows that Kant’s core account of the *I think* offers a coherent, persuasive approach to such awareness. With the advent of direct-reference and related ideas, we have begun to develop the sorts of tools needed to appreciate that approach properly. Nevertheless, Kant’s account leaves open two related, basic questions that need immediate attention.

First, point (I), like (II) and (III), is about self-reference by means of “I.” Kant, however, is talking not about language but about what he takes to be the mental representation *I think* or *I* through which we achieve self-awareness. While self-reference by means of “I” standardly involves self-awareness, we clearly can achieve such awareness (or, indeed, self-reference) without using “I”—or indeed without using any linguistic term

or device specifically designed to express that awareness. Here Kant agrees.³⁶ Moreover, accepting (I) to (III) does not immediately force us to accept Kant's view that there is a special mental *I*-act or representation that yields self-awareness. (I return to this point below.) Yet even if, following Kant, we do accept such a representation, it is clear that linguistic direct reference via "I" is one thing; the relation of the Kantian *I* to the self is another thing. It is a fact that the representation *I* introduces the self into the contents of our thoughts in a direct, propertyless way. However, that fact by itself does not imply that the *I* represents the self in any way analogous in detail to the way in which "I" refers.

In particular, that fact leaves open two related issues. (a) What exactly does the designation relation between the *I* and the self amount to, and how is that relation established? (b) How does the establishing of that relation come to introduce the self, the thing designated, into the content of the *I*-thought or awareness, if only in a propertyless way? (After all, the *I* can surely come to have many relations to other entities without those entities entering into the contents of my thoughts by means of *I*.) These issues could perhaps be resolved if the thinker somehow applied a mental analogue of (R) to her production of *I* in order to use *I* to think of herself. But the mere fact that the *I* introduces the self into one's thoughts in the way just noted does not imply any such behavior by the thinker. We thus have a group of issues that I will lump together into a first basic question: How, exactly, does the Kantian *I* or *I think* function, designatively, to represent the self and to bring it to our thought-awareness? To this question Kant's own comments on the *I think*, while offering the beginnings of several different views, provide no definite answer.

Second, and before turning to those views, I should stress that although point (I) seems true as far as the use of "I" goes, it does not get at the distinctively *first-person* aspect of the self-awareness that the use of "I" normally involves. The mere fact that a being—say, Smith—utters a term like "I," successfully intending his use of that term to follow rule (R), does not imply that that being is aware, in the first person, that he himself is the referent of that term. Watching himself in a mirror, Smith might come to realize, *of* the very being who he in fact is, that that being has just intentionally uttered "I" and thereby referred to itself. Smith might indeed come to realize that that being is expressing some claim about itself, for example the claim that it is hungry. Yet Smith need not be aware in a first-person fashion that he himself is so referring and claiming. Indeed, it is possible that, without using mirrors, Smith could come to be aware

of the very entity that he is, that that being is referring to itself with "I" and expressing a claim about itself. Yet Smith might still not be aware of himself in a first-person way. (In a dissociated state of shock an accident victim is aware of her body, hears it scream out "I'm hurt," but does not realize that she herself is the one screaming.)³⁷ When our reference by means of "I" expresses genuine first-person self-awareness, more is therefore involved than the mere satisfaction of (I).

What more? Some earlier writers have argued that everything here relevant to the use of "I" can be accounted for by appeal merely to points like (I) and, roughly, facts about the relations of the speaker to information available in the environment about bodily orientation, location, approaching objects, and so on. No appeal to anything like a special *I* representation is needed.³⁸ Needless to say, Kant's irreducibility points (vii), (viii), and (ix) show that he will strongly disagree. He will hold that "I" reference expresses genuine first-person self-awareness only when that reference expresses the awareness of self achieved through the special mental representation *I* or *I think*. I cannot defend all the details of Kant's disagreement here, but I believe that he is fundamentally right. Some sort of self-representation—or some special type of inner self-grasp—is necessary for full-fledged first-person self-awareness. An animal or robot might be specially attuned to its location and to objects in its environment, and it might carry on a constant, non-first-person monitoring of its bodily and mental states (as a computer can monitor its own computational states). This attunement and monitoring might enable that animal, without thinking of itself in a genuinely first-person way, to behave so as to take special care of itself, of the entity that it in fact is. But such behavior still will not be fully first-person.³⁹

However, even if Kant is right about the need for a special representation of self, fundamental issues remain about his own attempt to account for first-person self-awareness through the *I* or *I think*. In particular, and as we have seen, that representation is supposed to be simple, a mere "designation" of the self. But—and here is the second basic question I mentioned above—how can a mere mental blip, a momentary blink of the mind, not just designate the self but also do so in a way that (unlike the Smith cases just discussed) gives the self a genuinely first-person self-awareness? After all, even if we can answer our first basic question by showing how the *I*'s simple designation of the self can lead to the self's becoming an object of my thought (in a propertyless way), we may still have not shown that this thought is of a first-person sort.⁴⁰

Kant's irreducibility claims suggest that he may take this second question to be unanswerable. It is simply a primitive fact about beings like us that we have a thought-representation that yields such awareness, a representation whose mode of presentation of its object is first-person. However, there is a clear difference between a mental representation that directly designates the being that one actually is, without giving any first-person access to that being, and the *I* representation, which provides both such designation and such access.⁴¹ Even if much about the *I* cannot be further explained, it is desirable to try to make this difference understandable.

We thus have two questions for Kant's *I think* approach to self-awareness: (A) How, exactly, does the *I think* function, designatively, to represent the self and bring it to our thought-awareness? (B) How does the *I think* or *I*, a simple representation and a mere designation of the self, nevertheless yield its possessor genuine first-person awareness? Kant provides no clear answers to these questions. Indeed, his contemporaries realized early on that his remarks on the *I think* are in many ways incomplete—and to some extent internally inconsistent.⁴² This fact is not surprising, for current versions of questions like the last two remain open, and Kant's initial goal in discussing the *I think* was simply to provide a starting point for the Transcendental Deduction and for his criticisms of rational psychology. Nevertheless, study of points (i) to (xi) above, combined with various of his passing comments, suggests the germs of at least three distinct, and mutually incompatible, answers to these questions.

The first of these answers goes beyond Kant's core views on the *I think*, as represented by claims like (i) to (xi), to involve his transcendental idealism and the rest of his critical philosophy. Although this answer is the most developed of the approaches that he suggests to the *I think*, in my opinion it is philosophically unsatisfactory.⁴³ The approach is presented especially in the Paralogisms. Kant there, in effect, notes that acts of thought cannot, by themselves, grasp single, individuated objects as such. Hence the *I think*, as such an act, can at best be a thought—indeed, a concept—that in a *de dicto* fashion specifies the unique being (whichever it turns out to be) that thinks; the *I think* must in fact be what Kant at B426 describes as “the completely undetermined concept of a thinking being in general.”⁴⁴ This thought, or concept, does not by itself relate us to any specific thinking being (in effect, it simply thinks that there is a unique such thing without specifying which). However, when it is ap-

plied to the manifold of inner sense, it serves to unify that manifold and yields us the specific, inner-sense-mediated awareness of the particular subject of thoughts that is appearing to us through inner sense, namely our self. Thus (A), the question of how the *I think* functions to represent the self, is answered by appeal to the application of the *I think* to inner sense. (B), the question of how the *I think* yields its possessor genuine first-person awareness, is not explicitly addressed.

Smoothly though this first approach to *I think* self-awareness fits into Kant's overall transcendental idealism, it is inadequate as an account of first-person awareness and as an answer to (A) and (B). I mention three problems. (a) The answer to (A) is unsatisfactory. As we have seen, according to Kant the representation *I*, like the word "I," gets us, by itself, into the designation relation with a specific being, our self, although it does not thereby enable us to distinguish that being from other beings. But a *de dicto* concept of a subject of thoughts does not get us into any relation to our self (let alone a designation relation) until it is applied to the manifold of inner sense.⁴⁵ (b) No answer at all is given to (B). While the *I* is supposed to express genuine, first-person awareness, the mere concept of a subject of thought clearly does not. (c) Moreover, the first approach also conflicts with one of Kant's deepest insights about the *I*. For him, and as I have been stressing, the *I* represents the self in a direct fashion; it introduces no properties of the self into the contents of the thoughts that it enables us to think concerning that referent. However, the concept of a subject of thought introduces into the contents of such thoughts the property of being a thinker, a possessor of thoughts.⁴⁶ Given (a), (b), and (c), the first answer to questions (A) and (B) is therefore inadequate.

The second and third answers to these questions are at best pointed to in Kant's core remarks on the *I think*. I will not try, here, to develop these answers in full detail. According to the second answer, and as suggested by Kant's irreducibility claims (vii), (viii), and (ix), the *I think* is to be understood simply as an act of thought that designates the self and thereby gives us an awareness of the self in a direct, propertyless way that is also genuinely first-person. The second answer agrees that we can thus describe—as in the last sentence—the structure of the *I think*'s operations and its first-person character. However, this answer holds that the way in which the *I think* does these things and the reason that it has that first-person character cannot be further explained by, or reduced to, the operation of any other representations or mental entities. According to the stron-

gest way of stating this second answer (or refusal to answer), no informative, genuinely explanatory responses to (A) or (B) are possible at all. The *I think* is *sui generis*.⁴⁷

In connection with this second answer, I should note that the claims (vii) to (ix) make a variety of points about the irreducibility or unanalyzability of the *I think* and of the I itself. Defenders of any detailed form of the second answer will need to decide exactly which of these points they wish to endorse. After all, even if no informative answers to (A) and (B) can be given, as (vii) implies, it does not follow that we can have no understanding of the nature of the I itself that produces the *I think* except through the *I think* itself (as (viii) claims). In addition, and although all attempts that I know of to do without first-person awareness (insofar as it involves an *I* representation or a functionally equivalent device) have been unsuccessful, some type of explanation of that awareness may still be possible even if that awareness is in some important sense ineliminable. One can, after all, explain the use of indexicals like "I" by appeal to the rules, like (R), that link uses of those indexicals to objects (for example, the speaker) that occur in the context of use. Such an explanation is not undermined by (but indeed can be used to support) the fact that the characteristic use of indexicals cannot be eliminated in favor of the use of nonindexical terms. Similarly, even if the second answer turns out to be correct, it may still be possible to provide some nonreductive, noneliminative explanation of how first-person self-awareness (and Kant's representation *I*) comes into existence and achieves its characteristic effects. Whether such an explanation is possible seems likely to turn at least in part on empirical facts about mind and brain. I hope that it is; the second answer may be correct, but the existence of an in-principle completely inexplicable *I think* phenomenon is not easy to swallow.

The third answer to (A) and (B) is suggested by Kant's emphasis on the simplicity and lack of content of the representation *I*—"the simple, and in itself completely empty representation *I*" (A345–6/B404); the "in-content-entirely-empty expression 'I' [*den an Inhalt gänzlich leeren Ausdruck Ich*] . . . which I can apply to every thinking subject" (A355). Noting Kant's talk of the "expression" *I* here, let us couple such descriptions with his *Metaphysical Foundations of Natural Science* (4:542) remark that "the *I*, . . . itself a mere thought, designates as a mere prefix [*Vorwort*] a thing of indeterminate signification" and with his *Anthropology*, §1 (7:127), comment that even if someone, say a child, "cannot yet say 'I' [*das Ich*] he still has it in mind."⁴⁸ We may observe, also, that Kant

talks on occasion of our "perceiving" the presence of the representation *I* in our thoughts.⁴⁹ These various descriptions and turns of phrase suggest that Kant often thinks of the representation *I* as though it were actually a word in the mind—in fact, as though the act of thought that that representation constitutes were functionally equivalent to a referential use of the first-person pronoun or of a similar linguistic device. And that suggestion, in turn, leads one to see the representation *I* as functioning, for Kant, as a mental indexical.

If Kant did think of the *I* in this way, he hardly did so emphatically or in detail, but it is nevertheless a tantalizing idea well worth exploring. The third answer supposes that human beings possess internalized versions of the linguistic (or other) external ways that they have of directly referring to themselves in the first person; they use thought-productions of these internalized designations to refer to themselves and to think of themselves in a first-person manner.⁵⁰ Thus, following the mental analogue of rule (R), I, the thinker, think *I*. That is, I, the thinker, produce the act of thought that, for Kant, is the representation *I*; and, by that very fact of producing the *I* while following the analogue of (R), I refer to myself, in a directly referential, designative way. Because, normally, my thinking *I* while following the analogue of (R) is something that I am aware of, I understand this *I* to refer to its producer. Now, in the case of the linguistic "I," your knowing (R) and hearing me say "I am hungry" specifies me, the producer of "I," as the direct referent of "I" without importing any of my properties (including the property of being the speaker) into the proposition that you take my words to express.⁵¹ Similarly, when I am aware of this *I*-thinking and understand it to follow the mental analogue of (R), the producer of this *I*-thinking, namely myself, is specified for me as the referent of *I*. This producer is so specified without importing any of this producer's properties (including the property of being the producer) into the proposition that I take that *I*-thinking to express. So, by thinking *I*, I think of myself in such a way that I myself get into the content of my thought in a propertyless manner. (Also, to the extent that my use of the mental indexical *I* is primitive and ineliminable, my reference to and thought of myself is primitive in a way satisfying to Kant.)

Does the third answer, as thus developed, give a satisfactory answer to (A) and (B)? It answers (A), but internalized versions of the Smith cases above show that it fails to answer (B). Unless I already realize that the internal producer of the *I*-thinking act that I am now aware of is I myself, my mere internal recognition that that act refers to its producer will not

amount to my recognition that I myself am that producer. Is there any way around this problem? Here we are well beyond anything that Kant's characterizations of the *I* as a mental indexical immediately imply, but I would like to suggest in conclusion that a defender of the third answer ought now to focus on the reflexivity of first-person thought, a puzzling but I believe central phenomenon that, in (xi) above, we noted that Kant accepts.

The reflexivity of the *I* or *I think* act of thought means, roughly, that that act of thought is not just the thinker's act of awareness of the content that the thinker thinks through that act. (I think the content that: it is raining, etc.) Instead, it is an act of thought-awareness, by the thinker, that not only grasps that content but also takes that very act itself to have the thinker as its subject. How this occurs, and how it might be explained, are deeply perplexing matters that I cannot investigate here.⁵² But, assuming that it does occur, a defender of the third answer can now argue as follows. The thinker's inner awareness of the *I*-thinking act is not, so to speak, separate from and external to that act. Rather, that internal awareness amounts to the *I*-act's thought-awareness, of *itself*, that it refers to its producer. After all, the reflexivity of the *I*-act implies that that internal awareness of that act belongs to the producer of that awareness in his very production of that act. Moreover, the producer has internalized a mental analogue of rule (R). The producer now implicitly applies that analogue of (R) to his production of the *I*-act at the same time as, *in* that very *I*-act, he is made aware *of* that very *I*-act itself. In consequence, the producer is made aware of that *I*-act as being an act referring to its producer.⁵³ The *I*-act is thus at one and the same time—and at one and the same time is experienced through itself as being—an expression of agency that points back to its agent. Given these facts, the defender of the third answer should maintain that such an *I*-act constitutes genuine first-person self-awareness. (B) as well as (A) is thus answered.

Moreover, the defender should hold, (B) is answered (as in the earlier version of the third answer) in a manner that respects the peculiar way in which Kant takes the representation *I* to designate the self but not to distinguish it from other entities. In saying that the *I* is simple, has no manifold, and represents nothing about the self that it designates, Kant is agreeing with part of Hume's observation that there is no special idea (or, in Kant's terms, no special sensory intuition) of the self as such. Yet, Kant can observe, we do indeed have inner sense. When, in producing the *I*-act, the producer is made aware (in the way set out above) that that very *I*-

act refers to its producer, inner sense will ordinarily give the producer an intuition of that very act as emanating from the empirical self that has various other thoughts and feelings. In conjunction with the operation of inner sense, the producer's awareness that the *I*-act refers to its producer will thus ordinarily identify that producer to him- or herself as the particular empirical being that that producer is. What, however, of the fact that the production of the *I* act by itself, and without inner sense, designates, but does not distinguish, the producer? Since the representation *I* is not an intuition and contains no manifold, this fact cannot mean that the *I*-act gives the producer the *de re* thought-awareness of a self that the producer grasps in some mysterious fashion as a definite but wholly propertyless entity, a bare particular. Rather, the production of the *I* act, governed as it is by a mental analogue of (R) and reflexivity, guarantees that that act refers to (designates) the specific producer of that act and nothing else. Thus, a designation relation between the *I*-act and a specific entity (namely, the act's producer) is established. However, through the production of the *I*-act all that the producer is made reflexively aware of is that the *I*-act itself is occurring and that that *I*-act (because it follows the mental analogue of (R)) has, as its designation, its producer. In this reflexive thought-awareness of the *I*-act as designating its producer, the producer is thus not further identified *for* the producer. But through this reflexive awareness the producer is introduced—in the direct, propertyless way that we have noted—into the specific content of the thought that is thought *through* the *I*-act.⁵⁴

To my mind the third answer thus allows us to explain some puzzling, subtle Kantian claims. However, it will surely now be objected that question (B) has still not been answered. In the sort of self-referential act just described, the subject and producer of the act does indeed experience that very act as being produced by, and as having as its designation, its subject. But, the objector will say, such an experience is not the producer's first-person experience of *herself* as producing and being designated by that act. It seems to be such an experience only to the extent that the third answer covertly builds into that act the very first-person character that that answer is trying to explain.

To this objection I think that the defender of the third answer should reply that we are no longer talking of a Smith case. In such a case, and from outside the utterance of "I" or the production of *I*, someone—who, unbeknownst to herself, may well be the utterer or the producer—is aware of that utterance or production and realizes (by applying rule (R)) that it

refers to someone; but she does not necessarily realize, in the first person, that it refers to herself. Here, however, we have, instead, a case of an act centered on (experienced as referring back to) a certain subject, an act that *itself*, in its very occurrence and production, takes *itself* to be produced by and to refer back to that subject. And, the defender should say, first-person thought and experience can amount to no more than this type of thought.⁵⁵ Nothing more than a label is lost if we have not expressed this thought using a first-person expression. Rather, we have succeeded in analyzing first-person thought by appeal to the notions of a mental indexical *I* and of the reflexivity of the *I think*. Objectors may still disagree; but it does not seem that the third answer, defended in this way, is obviously and flagrantly mistaken.⁵⁶

I have now carried the discussion of possible Kantian answers to (A) and (B) well beyond anything explicit in Kant's texts, although I believe that the major points that I have appealed to in sketching even the third answer are present in those texts "like a germ in which the parts are still undeveloped" (A834/B862). (I hope also that, to complete the quotation, at least many of the major parts of the third answer are more than "barely recognizable" in the texts "even under microscopic observation.") If the organization and development in terms of (R) (and, earlier, of direct reference) have been mine, the starting points are Kant's. The second and third answers are incomplete (I take the first to be clearly inadequate); they also disagree strongly about various claimed irreducibilities; and their statements could be sharpened in numerous ways. Moreover, a great deal obviously remains to be learned, both philosophically and scientifically, about consciousness and self-awareness, and there is no reason to think that Kant has anticipated every future discovery. However, together with the direct-reference insights that I have argued are also to be found in his treatment of first-person self-awareness, all three answers show how deeply he probed this phenomenon. Moreover, I take the second and third answers especially to present live options for future research. Either one, I think, may turn out to embody or to point towards an important truth.

Notes

1. Beck took an interest in my earlier accounts of intuition and of the Transcendental Deduction, even though their starting point was very different from

his own. I know that he showed many other fledgling Kant scholars the same mixture of encouragement and insistence on reaching Kant's deepest insights that he did me. He was also a source of wise advice—and very good humor—about life, both academic and not. The humane ideal that he embodied is not one easy to attain. I want in addition to pay tribute to Virgil Aldrich, the aesthetician, who died recently at the age of 93, still actively pursuing philosophy. Aldrich was a luminous person, my first philosophy teacher, and the first to suggest I study Kant.

2. For this interpretation, and for further details of the manifold of intuition and its synthesis, see my book *Kant's Transcendental Deduction* (Dordrecht: Kluwer, 1992), especially chapters 1–3; henceforth *KTD*.

3. I cite Kemp Smith's translation of the *Critique of Pure Reason* (New York: St. Martin's Press, 1965), sometimes with alterations.

4. My own doubts are sketched in *KTD*, chapter 1. Many other writers, including P.F. Strawson, J. Bennett, and P. Guyer, have criticized Kant's grounds for transcendental idealism.

5. There are various ways of organizing and presenting these claims. I adopt the groupings below (which involve some overlap and looseness) mostly for ease of exposition. I have arrived at the core by considering the basic claims that Kant makes in the first *Critique* and in related texts about the *I think* taken on its own (and in its relations to inner sense). The core is not a collection of points about the *I think* that are united simply by the fact that they all seem both plausible and insightful.

6. See B131 ff. and A117 note. For further text references and a discussion of various intricacies in Kant's descriptions of the *I think* (or, as he sometimes says, of the *I am* or *I*), see *KTD*, chapter 7, and also my paper "Apperception and the 1787 Transcendental Deduction," *Synthese* 47:1981, 385–448. In order to avoid confusion, I use "*I think*" in italics to designate the act or representation (and "*I*" to designate its first-person part). Quotes—as in "my saying 'I think'" or "Kant's German expression '*Ich denke*'"—mention a linguistic expression that is used to express the representation. Reference, without quotes or italics, to the *I* and its properties refers to the self (whatever its nature may be) to which I refer in saying "*I think*."

7. See also B423 note: Through the awareness yielded by the *I think*, "something real . . . is given, given indeed to thought in general, and so not as appearance, nor as thing in itself (*noumenon*), but as something that actually [*in der Tat*] exists and which in the proposition, *I think*, is designated [*bezeichnet*] as such." B423 note (in part), B157, and B429 are part of Kant's answer to H.A. Pistorius. Pistorius objected that, given Kant's restriction of all the categories—and so the category of existence—to phenomena, the self, insofar as it exists, becomes a mere phenomenon, and so "there will be nothing but illusion [*Schein*], for nothing remains to which anything can appear"; quoted from Pistorius' review of

Kant's *Prolegomena*; originally published in *Allgemeine Deutsche Bibliothek* 59:1784, 322–57; quotation from page 345, as translated in N. Kemp Smith, *A Commentary to Kant's "Critique of Pure Reason,"* 2nd ed. (New York: Humanities Press, 1962), 323. See the further remarks in Kemp Smith and my *KTD*, especially chapters 2 and 7. (Those who deny that, for Kant, *I* or "I" refers to a genuine entity not only do not have a good account of texts like B157 ff., B122–3 note, and B429, but also seem to ignore Kant's need to answer Pistorius.)

8. In my "Apperception" and *KTD*, chapter 7, I argue that the *I think* refers in this way; that view also is present in the Kant interpretations proposed by philosophers like P.F. Strawson, *The Bounds of Sense* (London: Methuen, 1966); H.N. Castañeda, "Indicators and Quasi-Indicators," *American Philosophical Quarterly* 4:1967, 85–100; and R.M. Chisholm, *The First Person* (Minneapolis: University of Minnesota Press, 1981). It is well argued by A. Brook, *Kant and the Mind* (Cambridge: Cambridge University Press, 1994), chapter 4 (see especially 88); he notes accounts by H. Allison, J. Rosenberg, Patricia Kitcher, and C.T. Powell that deny such reference. See also the discussion by R. Aquila, "Self as Matter and Form: Some Reflections on Kant's View of the Soul," in *Figuring the Self*, ed. D.E. Klemm and G. Zöllner (Albany: State University of New York Press, 1997), 31–54, which gives further guides to the literature.

9. This is the best interpretation of the famous elastic-ball example in A363–4 note and of the contrast Kant develops in the main A363 text between the "logical identity of the I" and the fact that that logical identity does not imply that the *I* represents a self-identical person or any other type of entity with more than a formal or logical identity. (Such an identity could be satisfied by a changing committee of dependent particulars acting always in concert.) The basic claims of the first three Paralogisms should be interpreted along the lines of this sort of logical-identity interpretation. See my *KTD*, 379. There is an extended discussion in Brook, *op. cit.*, chapters 7–8.

10. On a proposed emendation—see R. Schmidt's edition of the first *Critique* (Hamburg: Felix Meiner, 1956), 141b—"cannot be accompanied by" here should be "cannot be derived from." The interpretation below can be given in terms of this emendation but does not require it.

11. Note A107 and A106 on transcendental apperception as the "pure, original, unchangeable consciousness" that is the ground—and ultimate source—"of the unity of consciousness in the synthesis of the manifold of all our intuitions." I interpret these claims (and the allied reasoning in B-Deduction, §§16–17) in *KTD*, chapter 8, 229–30. The implication of the literal language of B132 that the *I think* is a pure apperception that generates the representation *I think* itself is paradoxical (unless it is simply a hasty way of stressing the spontaneity and underivedness of the *I think*). It seems best to understand Kant as saying simply that our capacity for or faculty of apperception is original in the sense that that faculty generates the underived and not-further-explicable *I think*. (Compare A117 note: "the pos-

sibility of the logical form of all knowledge is necessarily conditioned by relation to this apperception *as a faculty* [*Vermögen*].")

12. A346/B404: the transcendental subject of thoughts "is known only through the thoughts which are its predicates, and of it, apart from them, we cannot have any concept whatsoever, but can only revolve in a perpetual circle, since any judgment upon it has always already made use of its representation." (Compare A347/B402, A353–4, and A366.) To speak strictly, while this A346/B404 passage argues that any adequate understanding of the I must involve the use of the representation *I*, it does not, unlike B132, say that that representation could not itself be analyzed as a combination of or relationship between further representations.

13. So I interpret (and flesh out) his A402 argument that "I cannot know as an object that which I must presuppose in order to know any object, and . . . the determining self (the thought) is distinguished from the self that is to be determined (the thinking subject) in the same way that knowledge is to be distinguished from its object."

14. This interprets (and again fleshes out) B422: "the subject of the categories cannot by thinking the categories acquire a concept of itself as an object of the categories. For in order to think them, its pure self-consciousness, which is what was to be explained, must itself be presupposed."

15. See B422 before the sentence just referred to. Kant makes a similar, no-special-I-intuition point at A399–400 in the text preceding A402. See also B157–8 note.

16. However, A402 speaks of "the determining self (the thought)" and thus seems perhaps to identify the *I self*, in our awareness of it via the *I*, with that *awareness* itself. So perhaps A402 and B422 make very much the same point. (Compare also A350: "in it [consciousness] . . . as the transcendental subject, all our perceptions must be found.") However, if Kant is to respect the usual working of "*ich*" or "*I*"—and, also, if he is to have a hope of answering Pistorius' objection that, for him, the *I* to which objects appear can be only a mere illusion (see note 7)—then he must distinguish the *I* that is the referent of "*I*" and of the representation *I* from the self-awareness experience that occurs in the *I think*. I will therefore ignore the idea that the *I* might itself be identical to the *I think* experience.

17. So also may be the somewhat similar A346/B404 reasoning quoted in note 12 in connection with point (viii). From the fact that I have to use my *I think* in the course of arriving at claims that explain the *I think*, it doesn't follow that my *I think* occurs in those claims in such a way as to make the explanation circular. (Compare the fallacious argument about language above.)

18. As thus interpreted as amounting to (ix), the A402 and B422 claims relate closely to claims (viii) and (vii) above. ((viii) indeed entails the part of (ix) concerned with the *I*, and (vii) entails the part of (ix) concerned with the *I think*.) It may well be that Kant is also arguing in these claims in a way analogous to the

bad argument about language that I noted above. (Observe, for instance, A402: "I cannot know as an object that which I must presuppose in order to know any object." Taken literally, this seems very doubtful. Can I not, as knower, know my own existence or space and time or the principles of logic?) However, Kant often packs numerous distinct claims, some deep and insightful, into single, succinct bits of reasoning; and his arguing in this bad way (if he is) still allows that he also is aiming at point (ix) above. There is no space to pursue this issue further here.

19. Such a reflexive awareness of that thought will clearly be different from the awareness that I have, through *I think* accompaniment of my inner-sense intuitions, of my self as having temporally ordered thoughts, sensations, and other mental items. (Kant indicates at the end of B422–3 note that this reflexive awareness via the *I think* would not take place without some empirical—some outer- or inner-sense—representation to supply the "material" for "the *actus I think*." But earlier in that note he observes that the *I think*, and so such reflexive awareness, "precedes," and hence is independent in logic of, "the experience which is required to determine the object of perception." See also my *KTD*, chapter 7, especially 196–7.) Moreover, this reflexive awareness will of course grasp my self in the purely designative and nonproperty way that we have already noted.

20. See also B423 note, B157, and B157–8 note.

21. See S.F. Barker, "Appearing and Appearances in Kant," in *Kant Studies Today*, ed. L.W. Beck (LaSalle, Ill.: Open Court, 1969), 274–89; Howell, "A Problem for Kant," in *Essays in Honour of Jaakko Hintikka*, ed. E. Saarinen, et al. (Dordrecht: Reidel, 1979), 331–49; and Howell, *KTD*, chapters 1 and 2. (The considerations there about outer-sense objects can be applied to the self appearing in inner sense.) The issues raised by these theories are extremely complex—more so than most writers in the literature have realized.

22. Thus note the A401 discussion of the fact "that the being that thinks in us is under the impression that it knows itself through the pure categories" and the whole issue, famously unresolved by Kant, of the relation of our individual empirical selves to one another and to the (single or multiple?) I that knows them. Outside the first *Critique*, note *Progress* (20:270–1); and observe that in *Anthropology* §4 footnote (7:134), Kant says firmly that apperception and inner sense give us two ways of representing the same subject, not a doubled I. See also Brook, *op. cit.*, 92–3.

23. Given the B429 language ("no mere appearance insofar as I think"), I suspect that the appearance theory is responsible for a good deal of Kant's two-selves talk. (His use of that theory does not give the whole explanation, however; B155 suggests the additional influence of familiar merging-and-branching issues of the sort discussed in my *KTD*, chapter 2, 56–7, and notes 37, 41–2; see also chapter 1, note 29.) The appearance theory in any case leads to intolerable difficulties for Kant when combined with his idealism. See *KTD*, chapter 1, and "A Problem for Kant." As I also argue in those places, Kant's appearing theory itself

leads to related difficulties; but since—among other reasons—our normal view of self-awareness is that I, who think my current thoughts, *am* the same entity as the temporal self who has those thoughts, the appearing theory is the best way to interpret Kant's account of self-awareness. Moreover, we are ignoring transcendental idealism in this paper.

24. Strawson, *Individuals* (London: Methuen, 1959), and *The Bounds of Sense*; Castañeda, *op. cit.*; and Chisholm, *op. cit.*

25. My point is not that Kant accepts some explicit version of a modern direct-reference theory. Rather, his account of the *I think*, which proceeds at the level of representational thought, has hold of the same sort of phenomena that contemporary direct-reference and related theorists see as present in linguistic reference via "I." Given its own intellectual virtues, the direct-reference framework thus offers the best current way of coming to grips with the distinctive (and I think correct) points in Kant's account. Really to appreciate what is going on in an author's text, we need to understand the problems that he or she was confronting. To reach that understanding, we need to consider the most successful current theories of those problems. Of course these theories may be superseded (some better approach to the issues than direct-reference views may be developed). But, until they are, we should apply the current theories, with sensitivity, in order to understand the import of our author's ideas.

26. (I) is based especially on points in J. Perry, *The Problem of the Essential Indexical* (New York and London: Oxford University Press, 1993); and D. Kaplan's essays "Demonstratives," and "Afterthoughts" in *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein (New York: Oxford University Press, 1989) 481–563, 565–614. For the idea of direct-reference, see 492–7, 523–4, and 568–9 of Kaplan's essays. D. Lewis' "Attitudes *De Dicto* and *De Se*," *Philosophical Review* 88:1979, 513–43, on *de se* reference also should be noted. Related points are found in Strawson's *Individuals*; Castañeda, *op. cit.*; and Chisholm, *op. cit.* Points akin to (II) are defended by many philosophers, especially by Strawson in both *The Bounds of Sense* and *Individuals*, and by S. Shoemaker, "Self-Reference and Self-Awareness," *Journal of Philosophy* 65:1968, 555–67, and his *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996). The claims in (III) derive from reflection on ideas expressed by all these philosophers; Castañeda was among the first to discuss them in depth.

27. I focus on the so-called attributive use of definite descriptions and ignore the frequent role of context in establishing their referents.

28. I adapt the formulation in Perry, *op. cit.*, 95. A number of other philosophers, including especially Kaplan in "Demonstratives," 520, have noted this rule and its philosophical interest. By the statement's speaker, in (R), I understand its maker; a spokesperson who reads out my first-person statement is not its speaker.

29. Note Kaplan, "Demonstratives," 494: "[T]hink of . . . what-is-said in a given context . . . as propositions. . . . For each occurrence of a singular term in a

sentence there will be a corresponding constituent in the proposition expressed. The constituent of the proposition determines . . . the object relevant to evaluating the proposition. . . . In general, the constituent of the proposition will be some sort of complex, constructed from various attributes. . . . But in the case of a singular term which is directly referential, the constituent of the proposition is just the object itself." Also 523: "Indexicals are directly referential. The rules [specifying the referents of occurrences of the indexicals] tell us what it is that is referred to. That is, they *determine* the content (the propositional constituent) for a particular occurrence of an indexical. But they are not a *part* of the content (they constitute no part of the propositional constituent)." And his "Afterthoughts," 569: "The directly referential term goes directly to its referent, *directly* in the sense that it does not first pass through the proposition. Whatever rules, procedures, or mechanisms there are that govern the search for the referent, they are irrelevant to the propositional component, the content. When the individual is determined, . . . it is loaded into the proposition."

30. See note 26. Such points can be defended on various other grounds.

31. I am putting this point fairly loosely. A detailed discussion of (III) would show that there are a number of ways in which the distinctive use of "I" might be held to be irreducible, unanalyzable, and ineliminable. By "a rule analogous to (R)" I mean, roughly, a rule that governs some other linguistic expression or feature (for instance, "*ich*" or "*je*" or various verb endings) in the way that (R) governs "I."

32. As noted, versions of (III) occur in Castañeda; he was well aware of the issues raised under (I) and of the bearing of (I) to (III) on Kant. So too is Chisholm. Strawson made a version of (II) (and (III)) and its connections with Kant prominent. In my "Apperception and the 1787 Transcendental Deduction," 386, and *KTD*, 160 and 371, I very briefly note the presence of matters connected with (I) in Kant. Brook discusses Kant on self-awareness in detail and considers issues relevant to (I) to (III). However, Brook does not focus on direct reference, and I don't think he appreciates the relevance, for example, of the direct-referentiality of "I" to Kant's important point (iii) about the propertylessness of self-reference via "I." (See pages 494–5 of my review of Brook's book in *Mind* 105:1996, 491–5.) Various individual points related to (I) to (III) have been widely noted, but I know of no systematic study of these points that takes into account the crucial importance of the direct-reference idea and relates that idea in a detailed way to Kant. Henrich has been a pioneer in defending nonlinguistic versions of (III), and his recent articles "Subjektivität als Prinzip," *Deutsche Zeitschrift für Philosophie* 46:1998, 31–44, and "Self-Consciousness and Speculative Thinking" in *Figuring the Self*, 99–133 (originally published in his *Fluchtlinien* [Frankfurt/M.: Suhrkamp, 1982]), offer a wealth of arguments and proposals about the irreducibility of self-awareness. He is well aware of the relevance of current discussions of linguistic self-reference to understanding self-awareness, but he has not, to my knowledge,

commented on the direct-reference idea or considered in any detail its bearing on Kant. M. Frank's *Selbstbewußtsein und Selbsterkenntnis* (Stuttgart: Reclam, 1991) contains a comprehensive survey of the literature and the issues, with attention to Kant among many other historical figures. He discusses direct reference and points by Kaplan and Perry, but he does not provide a detailed study of Kant or bring out the close connection of Kant's core views with the idea of direct reference. The papers collected in *Figuring the Self*, ed. D.E. Klemm and G. Zöllner, ignore that idea, as do the contributions in another valuable collection, *The Modern Subject*, ed. K. Ameriks and D. Sturma (Albany: State University of New York Press, 1995), with the exception of Frank's paper "Is Subjectivity a Non-Thing, an Absurdity [*Unding*]?" 177–97, who discusses the related idea of *de se* attributions. However, Frank does not make a detailed study of Kant. In his essay "Understanding Apperception Today," in *Kant and Contemporary Epistemology*, ed. P. Parrini (Dordrecht: Kluwer, 1994), 331–47 (see also his contributions to *Figuring the Self*, and *The Modern Subject*), Ameriks focuses on recent "empiricist" approaches to apperception and, like Frank, makes many interesting points; but neither he nor those approaches consider the relation of direct-reference ideas to Kant. Nor does Sturma in his stimulating discussions of matters bearing on (III), although he is clearly aware of such ideas. See his "Self and Reason," *The Modern Subject*, 199–215, and "Self-Consciousness and the Philosophy of Mind," in *Proceedings of the Eighth International Kant Congress*, ed. H. Robinson (Milwaukee: Marquette University Press, 1995), vol. 1, pt. 2, 661–74. H.F. Klemme's thorough study of Kant's account of self-awareness in his *Kants Philosophie des Subjekts* (Hamburg: Felix Meiner, 1996) is not focused on matters like direct reference. Neither is Pierre Keller's recent *Kant and the Demands of Self-Consciousness* (Cambridge: Cambridge University Press, 1998), nor is James Van Cleve's *Problems from Kant* (New York: Oxford University Press, 1999). Patricia Kitcher's "Kant on Some Functions of Self-Consciousness," in *Proceedings of the Eighth International Kant Congress*, vol. 1, pt. 2, 645–60, proposes an explication of Kantian self-consciousness but does not consider its first-person characteristics. This literature is rich, interesting, and helpful. What is not being recognized, however, is that much of Kant's core account of the *I think* turns on matters analogous to direct reference via "I"; and the direct-reference idea provides the best available framework for approaching his treatment of self-awareness.

33. Allowing for this analogue is, of course, not explaining why the possibility of "I" ascription or *I think* accompaniment exists. (Contrary, I think, to Kant, I view that question as an empirical matter but cannot discuss it here.) Sometimes Kant talks as though the *I think* always accompanies our representations (B350, A362–3). However, this is not his dominant view, and it is not, I believe, a plausible position in itself. Nevertheless, and as I note in *KTD*, 162–6, the plausible, dominant view of the mere possibility of *I think* accompaniment leads to difficulties when Kant appeals to it as a premise of the Transcendental Deduc-

tion. I ignore those difficulties here, for they show, not that the mere possibility is implausible, but that there are problems with using it, plausible as it is in most cases, to undergird the Deduction.

34. The term "dthat (the actual entity who is now producing these very words)" might be argued to achieve direct *reference*—to the person producing the words in parentheses—in a way analogous to the reference achieved via "I." (In Kaplan's "Demonstratives," "dthat" functions as a demonstrative that, prefixing a description, converts that description into a directly referential term whose denotation is the denotation of the description.) Nevertheless, this term doesn't function to express first-person *self-awareness* as does "I."

35. Since it is easy to get confused at this point, note that the B429 passage and the propertylessness of the I, insofar as the I is represented by the *I*, do not mean that (a) through the representation *I I* positively represent myself as not having any properties. Rather, what is meant is that (b) through the *I I* do not represent myself as having any properties. (a) would mean that the *I* represents me as some sort of bizarre, propertyless being; (b) means only that the *I* representation by itself does not specify any properties as belonging to me. (B429: "in the consciousness of myself . . . nothing in myself is *thereby* given for thought"; my emphasis.) Even though I thus gain no information through the *I* about any properties that I may have, I still might get that information from other sources. According to Kant, I in fact get it from inner sense, although given transcendental idealism I thereby learn only of properties of myself as I appear to myself. Were he to reject transcendental idealism, Kant could grant that through inner sense I might know properties of myself as I am in myself—and so properties of the referent of *I*.

36. *Anthropology* §1 (7:127): even if someone "cannot yet say 'I' [*das Ich*] . . . he still has it in mind. So any language must think 'I' when it speaks in the first person, even if it has no special word to express this I-hood." This fact is accepted—and is extended in *I*-representation-disavowing ways that Kant would reject—by writers like Perry, *op. cit.*, 62–4, 145–9. See note 38 below.

37. Compare M. Dummett, *Frege and Other Philosophers* (Oxford: Clarendon Press, 1991), 318.

38. Perry argues also that "I" sentences serve to classify and describe certain self-oriented ways of perceiving, thinking, and behaving; he appeals to all these facts to conclude, *inter alia*, that "there is no need for a self-referring component of our belief, no need for an idea or representation of ourselves" (*op. cit.*, 219; see also 61–4 and 145–9). However, in a later article, "Rip van Winkle and Other Characters," *European Review of Philosophy* 2:1997, 13–39, Perry recognizes nonlinguistic cognitive roles; such roles include the self role of "being the owner of the thought" (21). So Perry may now accept the full reach of first-person self-awareness.

39. For example, the animal is not in a position to reflect, in the first person, on all this monitoring and attunement to the environment as being *its own*. Such

reflection can be important for survival. It allows, among other things, for the animal to plan for breakdowns in the monitoring apparatus itself and to deal with them if they occur. ("Is that a real growling that I hear and is it directed at *me*? Are my eyes malfunctioning?" and so on.) If it lacks such a means of self-reflection—means not tied to current properties of the animal and its location—then to deal adequately with such problems the animal will have constantly to update some non-first-person self-image ("the *x* with the current monitoring system *S* and body *B* located now at point *p*," etc.). This self-image would itself be subject to reference failure and to other malfunctions in ways that the representation *I* is not.

40. The first basic question thus asks (a) how the designation relation between *I* and the *I* is established (and its nature) and (b) how its establishment works to ensure that the content of the thought that the *I* thinks by producing *I* is the structured proposition whose subject slot is occupied by the being, *I*, and is not occupied by any properties of that being. The present, second basic question asks how the *I*'s production of *I* ensures that the *I*'s grasp of that being, in grasping that proposition, is of the first-person sort. (Why does the *I*'s production of *I* ensure that the mode of presentation, to the *I*, of that proposition is first-person?) By a proposition, here, I mean a structured entity that gives the content of what we say, think, or judge. This entity may be thought of as an ordered collection of slots filled by the properties that our thought ascribes to objects and by either those objects themselves (in the direct-reference case) or by combinations of properties specifying those objects (in the "man wearing a black hat" case). The metaphysics of propositions is not found in Kant, but he is committed to the idea that, in thinking via concepts and the *I*, we arrive at judgments that claim, for example, that *I* am working. His account of the *I think*, in effect, characterizes what is thought in such judgments. I use the notion of a proposition in order to elucidate his characterization, without meaning to attribute that notion to Kant himself.

41. Compare note 34.

42. One major inconsistency—and well-known problem—lies in Kant's various attempts to classify the *I* or *I think* as a representation. (Note that this problem, like others that arise, isn't part of Kant's (i)-to-(xi) core group of claims about the *I think*.) (a) Given A320/B376–7 and related texts, the *I* can be only an intuition or a concept. As a representation of thought, it would seem to be a concept. But thought through concepts is discursive and represents objects by means of general properties common to those objects, unlike the *I think*. (b) Perhaps, then, it should be regarded as an intuition (as Fichte proposed, and as Kant himself says—in a non-Fichtean manner—in various places). But that classification also seems impossible. For Kant, all human intuitions are sensible, not achieved via thought (intellectual); and intuitions provide us with distinguishable, individuated objects, unlike the *I think*. In my "Apperception and the 1787 Transcendental Deduction," 428–9, 446–7, and *KTD*, 192–8, I observe that not only does Kant in various places describe the *I* or *I think* as not-a-representation, a judg-

ment, a concept, and an intuition, but also his varying accounts of the *I think* are, at least in part, a response to conflicting intellectual pressures that he is under. To those observations it is worth adding that Kant's difficulties in classifying the *I think* are not merely an artifact of such pressures. The difficulties reflect, also, the problem of understanding the nature of first-person awareness. Such awareness is not ordinary sense awareness and (for the reasons just given) is not in Kant's sense an intuition. Yet, equally, it achieves direct reference, in a distinctively first-person fashion, to a specific being, ourselves. So it is unlike ordinary thought through concepts. As noted above in connection with (B), it also goes beyond cases of simple direct reference governed by just (R). So what is it? If Kant has no satisfactory answer, then neither does anyone else. (Kant in addition says that the *I think* "expresses an indeterminate empirical intuition," B422 note; or a "feeling of existence without the slightest concept," *Prolegomena*, §46 note (4:334); or "a perception of one's self" (A342/B400); compare also A848/B876. But these phrases, however suggestive, do not solve but merely label the problem. For Kant an act of thought like the *I think* cannot literally be a perception, for one thing.)

Perhaps here is the place to note that Kant's core account of the *I think*, as set out in (i) to (xi) (and as extended in the first, second, or third answer below), does not involve the idea that first-person awareness arises through the self's making an object of itself in consciousness. (Such an idea appears circular—the self could not arrive at the first-person awareness that that object was itself unless it already had identified itself as such and so already had such awareness.) So I agree with Ameriks (in *The Modern Subject*), Sturma (in the second paper cited in note 32), Klemme (*op. cit.*, 377 ff.), and Keller (*op. cit.*, 104–5) that Kant does not hold the sort of reflexion theory of self-awareness criticized in Henrich's article "Fichte's Original Insight," in *Contemporary German Philosophy*, ed. D.E. Christensen, et al. (University Park, Pa.: Pennsylvania State University Press, 1982), vol. 1, 15–53.

43. For a longer discussion, see *KTD*, 193–7.

44. Note, further, the B418 association of the *I think* with "the concept of a thinking being in general," and the A346/B404 implication that the *I think* represents "a transcendental subject of thoughts = X." Of course Kant also speaks of "the consciousness of myself in thought in general" (B409), and of "the thought of myself" (B430). Such talk reflects his continued recognition that—contrary to the view I am here sketching—the *I think* doesn't just relate us to some thinker in general, it relates us to ourselves.

45. Of course the concept of a subject of thoughts relates to the objects in its extension—and so to our own self—in the conceptual way indicated in section I. But that concept does not thereby relate specifically to our self alone, and its relation is not of the propertyless, designative sort that Kant attributes to the *I*. Moreover, application of the concept to the manifold of inner sense then yields a relation to the self with all its inner-sense-knowable properties rather than such a designative relation.

46. To use contemporary terminology, Kant's talk of the *I* as the concept of the subject of thoughts in effect focuses on properties of the self that are used to fix the reference of that representation. (The referent of the *I* is fixed as the subject that produces the act of thought, *I*, among other thoughts.) Kant then introduces those properties into the content of what is thought about that referent via *I*. This procedure is analogous to defining "I" as meaning "the speaker." (Kant's talk of the subject of thoughts could also perhaps be understood as saying that the *I* is a directly referential thought whose referent is fixed by the fact that, roughly, that referent is the subject that produces various thoughts. However, if this answer does not simply return us to Kant's original view of the *I think* as a directly designative, first-person thought, it leaves the first-person nature of the *I* still unaccounted for.) The present *de dicto* approach to the *I think* also undermines Kant's attempt to answer Pistorius. (See *KTD*, 195–6 and 380–1.)

47. I see this implication as implicit in Henrich's "Subjektivität als Prinzip," although he does not discuss Kant in any detail in that paper or address the specific questions (A) and (B). (But see his "Self-Consciousness and Speculative Thinking," 111–15.) His discussion involves numerous subtleties and needs much more attention than there is space for here. He also would not necessarily accept my direct-reference framework. By noting other notions that he holds to be explanatorily primitive, he tries to remove the impression that what I have called genuine first-person self-awareness is a *Spuk*, a mysterious surd in an otherwise comprehensible world. I would like to thank Henrich for a copy of this paper and for a lively, stimulating conversation about Kant on Unter den Linden in May 1998.

48. In Kant's eighteenth-century German "*vor*" sometimes occurs for "*für*"; see, e.g., I. Heidemann in the Afterword to the Reclam edition of the first *Critique* (Stuttgart, 1966), 994. So it is tempting to read *Vorwort* in the first quote as *Fürwort* (pronoun), which would make the case even stronger for seeing the third answer in Kant's remarks. Kant's use, here and in the first *Critique*, of "designate" (*bezeichnet*) to describe the operations of the *I* also suggests that he thinks of the *I* as a mental linguistic term. Note, further, A727–8/B755–6 on empirical concepts, where Kant associates designation with linguistic terms. (He says a word like "water" is more a *Bezeichnung* than a concept of the thing, water. This passage also hints at the current idea that natural-kind words are rigid designators and so directly referential.)

49. A350; note also A364: "the only permanent appearance which we encounter (*antreffen*) in the soul is the representation *I*."

50. As Kant's remarks on the people who have *I* in mind but cannot yet say "I" show, he apparently supposes that the internalized *I* comes first, then the external linguistic "I." Alternatively, we might hold that we first learn "I" and (R) and then internalize to *I* (or we do both at once). Or we first use our own proper name in an "I" fashion and then switch to "I" (compare Perry, *The Problem of the Essential Indexical*, 61–2). We need not settle such matters here.

51. Recall, again, that "I" reference is direct. The content of what my words say (as against what you may be able to gather from hearing my words in a particular context) is specified by first using my property of being the speaker to fix the referent of "I" as I. I (and not any of my current properties, including that of being the speaker) then occurs in the proposition that constitutes what is said in my statement.

52. The *I think*'s thus taking itself could perhaps be causally generated by some sort of feedback from part of a neural pattern to other parts of that pattern. The causal explanation is not our concern here, and it seems possible that, as the second answer suggests, at least at the phenomenological level the reflexive nature of the *I think* experience is not further to be explained. (The third answer need not disagree with this particular possibility.) It is also conceivable that the apparent reflexivity of the *I think* actually amounts to some other mental process's somehow monitoring our thinking. I doubt this (and it in any case conflicts with Kant's claims about the originality of the *I think*). However, these questions involve empirical issues that cannot be settled here.

53. Moreover, this *I*-act awareness is surely what recent authors have called awareness from the inside. The *I*-act grasps itself not in the way in which I imagine watching myself, from a distance, step into the woods, but in the way in which I imagine actually shouldering a pack and setting out on the trail. (I here assume that awareness from the inside is not inherently first-person. Anyone who disagrees should ignore this present note.)

54. Compare B157–8 note, with my emphases: "The *I think* expresses the *act of determining* my existence. Existence is already given thereby, but the mode in which I am to determine this existence, that is, the manifold belonging to it, is not thereby given. . . . [Through the *I think*] I cannot determine my existence as that of a self-active being; all that I can do is represent to myself the spontaneity of my thought . . . and my existence is still only determinable sensibly . . . as the existence of an appearance." Observe also B423 note, as quoted in note 7 above. Note carefully that, on the present version of the third answer, the content of the *I*-act's reflexive awareness of itself is different from the propositional content belonging to (or thought through) the *I*-act itself. On this version we distinguish five things: (1) the content of the thought that one thinks through the *I* or *I think* (as observed in note 40, this content is the structured proposition whose subject slot is occupied simply by the being that is I); (2) the first-person mode of presentation of that content to the I, through the *I* or *I think*; (3) the reflexive thought-awareness that the *I*-act gives; (4) the content of that reflexive thought-awareness; (5) the fact that the production of the *I*-act, following the analogue of (R), establishes the designation relation between that act and the I and, given (3), ensures that the I, as its referent, occurs in the content of thought in (1). (A more refined account would, I think, connect (1) and (4) in a reflexive way. But doing so is beyond the scope of my project here.)

55. In this spirit the defender might say that the core of first-person self-awareness is just an awareness of self as the source or center from which come all the mental acts, desires, perceptions, and so on, of which the self is aware—including this very awareness itself that the self has of those mental acts. The reflexive account of the *I*-act regards the *I*-act as a mental indexical that is taken by itself to refer back to its producer. Because that act also thereby takes that producer to produce *it*, that act also takes itself to emanate *from* that producer. So the reflexive account makes the *I*-act into precisely such a core state of awareness. These thoughts call to mind the deep, fascinating reflections on reflexivity and first-person thinking in J.D. Velleman, "Self to Self," *Philosophical Review* 105:1996, 39–76, especially 48–51. That paper, which I read on publication and then put aside, has influenced my formulation of the third answer.

56. This third-answer analysis will obviously be rejected by many or most defenders of the second answer. However, and as I noted earlier, this analysis involves the ineliminable use of a mental indexical "I" (namely, the *I*-act functioning reflexively as such an indexical) and in that sense still preserves something of the irreducibility ideas. See also the second sentence of note 52. It seems clear that such ideas, as applied to first-person reference and thought, need much further discussion.

The Problem of Time in Kant

Gerold Prauss

To the present day the problem of time remains unresolved. Recognized as a problem already by Zeno, Aristotle, and Augustine,¹ time emerges ever anew as problematic because its continuity is conceived in exactly the same manner as the continuity of space. It may appear justified to conceive of it in that way inasmuch as this continuity of space and time stands for the continuity of their extension. The extension of space and time is assumed to be their common feature. The continuity of extension must be common to space and time, or it is at least commonly believed to be so.

In the case of space, however, the relation between such a continuity of extension and discreteness is always one of alternatives, regardless of the considered dimension. This becomes clear when we realize that the continuity of extension of a three-dimensional body always lies outside the two-dimensional surface that forms its boundary. Correspondingly, the continuity of extension of a two-dimensional surface lies always outside the one-dimensional line that bounds that surface. Furthermore, the continuity of extension of a one-dimensional line lies also always outside a zero-dimensional point, which forms its boundary.

It is precisely this relation of alternatives between continuity and discreteness—which in the last mentioned case of the line consists in the relation between extension and point—which in the case of time has produced this persistent and unresolved problem. If the “either—or” relation between point and extension is posited as fundamental in the case of time, the specificity of time is forfeited. It appears, therefore, that the specific nature of time must consist in the relation of “as well as” of extension and point; yet this seems to create a contradiction, to which time must fall victim always anew. All these problems have appeared in Zeno, Aristotle, and Augustine, and numerous times after them.

It is remarkable that to the present day it has never been seriously considered whether this difficulty is perhaps not self-made. Does it not arise only because of the arbitrary assumption that continuity and dis-

creteness in the case of time are related in the same way as in the case of space? The relation of alternatives between continuity and discreteness is too firmly taken as self-evident to be arrived at on the basis of critical considerations. Would it not, then, be monstrous to attempt to criticize it in any sense? How could this genuine monstrosity—that a point does not have extension outside but rather inside itself, and also that discreteness does not have continuity outside but rather inside itself—ever be imagined without contradictions?

As far as I can see, Kant is the first who, in order to solve the problem of time—the structure of which is well known to him—undertakes to consider whether this monstrosity could really be stated without contradiction. His attempt, however, is never fully carried out, so that it has hitherto been overlooked. I myself took notice of it only when I was in possession of the model with the help of which it can be carried out.

At the latest in 1781, when the first edition of the *Critique of Pure Reason* was published, this much was clear to Kant: the “lack” of a proper solution to the problem of time “we attempt to compensate for by means of analogies,” as when for instance we strive to represent time as a spatial one-dimensional “line.” This is only an analogy precisely because “the parts” of such a line are “simultaneous,” whereas those of time are, on the contrary, “successive” (A33/B50). What consequences this ultimately has, Kant does not articulate here: minimally they are that, in order to arrive at a conception of time through such an analogy, we have to disregard precisely the fact that the parts of this line are “simultaneous”; thus we have to disregard the line altogether. For even the shortest line must have parts that are “simultaneous.” This makes it questionable whether anything should remain—even in a merely analogous way—by means of which we can conceive of time. Even if it were meaningful to say that what remains is a point,² the question of whether time can be represented as a point—even in a merely analogous way—would nonetheless loom large because, in contrast to a point, time does have parts that are “successive.”

It is first in 1787, in the second edition of the *Critique of Pure Reason*, that there is evidence that Kant attempted to release himself of these difficulties and put forth a new and better suggestion for the conception of time. But the novelty of the suggestion remains at first concealed because he once again employs the image of line. Once more it was claimed that the conception of time is possible for us only “in the image of [a] line” with the caveat that this is only a “line, insofar as we draw it.”³ In comparison to the first edition of the *Critique*, the novelty consists just in the

addition of the subordinate clause, which Kant begins with the help of “insofar as.” Yet next to the sentence from the first edition that ends with the word “line,” this new and added “insofar as” seems so barely visible that Kemp Smith, for instance, does not even translate it.⁴ Moreover, he replaces it with “which”; according to Kemp Smith, Kant here talks not about a “line, insofar as we draw it” but rather of a “line, which we draw.” Through this sort of translation, however, the decisive feature of the novelty does not find expression. By means of this novelty Kant attempts for the first time to do justice to time as something essentially dynamic; namely, he attempts to find for time an “image” or “analogy” of something dynamic rather than something merely static. From the preceding text, it is already clear that with this formulation something can be expressed as a demarcation from something else, which is yet to be completed. What can be meant is only “a line, insofar as we *draw* it,” and not something like “a line insofar as we *have* drawn it and insofar as it is already drawn.” This second way of presenting a line—as something static—is precisely what does not correspond to the dynamic nature of time. It is exactly these difficulties that Kant from then on strives to avoid: not as something finished and thereby static, but only as something in the process of becoming and as such dynamic, can a line be used as an “analogy” or an “image” of time.

That this must be what Kant means follows—as has been mentioned—from the context. Shortly before that, he makes it clear that he is not at all concerned with the line as such, as the already drawn and thereby static line, but exclusively with the “drawing” of a line and thus with its corresponding dynamics (B154). He formulates here an extremely strict requirement: the drawing of a line is a proper analogy or a proper image for the purpose of presenting time only when, in the process of drawing, we pay “attention” to this “activity itself.”⁵ This can only mean that, in the process of drawing, we must pay attention to the drawing itself. That what Kant demands is in fact the decisive point of his novel suggestion is even further emphasized in the following way: in order to arrive at a conception of time by means of drawing a line, we cannot treat this line as if it were space, and that means that we have to neglect the line itself. Only when we completely succeed in doing so can we, according to Kant, pay attention exclusively to the act of drawing itself, and only thereby can we accomplish the proper conception of time.

With that kind of demand, Kant seems to ask for the impossible, because it is unachievable for us. For how could it be possible for us that, in

the process of drawing, we pay attention only to this act itself, when—according to Kant's own assumption—this act of drawing is an act of drawing a line? It seems out of the question that we can pay attention to the process of drawing without at the same time *also* paying attention to what is being drawn, that is, without paying attention to this line as spatial. Even if that were possible, however, it seems that the factual and objective problem of time would thereby be transformed by Kant into a subjective and psychological problem, in the negative connotation of these words. It seems that with this move Kant would make himself vulnerable to a complaint that could be formulated as follows: "My dear Kant, you seem to have an ability to imagine such things; I, however, lack it."

All of this appears that way only because his suggestion is presented intuitively without being thought through. Yet his suggestion can be operationalized and de-psychologized. For the realization of what Kant suggests is that there is a procedure which is precisely regulated and can, as such, be exactly described. This procedure moreover leads to the realization that there is in advance no possibility of paying attention in the process of drawing a line to anything else but to the act of drawing itself. Through this process a model of time is generated, which allows us to comprehend what Kant here has in mind. It was only after I had proposed this model of time that it occurred to me that it makes Kant's suggestion understandable. Clearly to understand both of them, the only thing that is needed is a reflection on the well-known idealization which every geometrician makes and which Kant tacitly assumes as well.

Drawing as the sketching of a line is in fact nothing other than a certain extension of pigment. For the geometrician it is, nonetheless, the depiction of an ideal geometrical object in the sense that a line as an ideal geometrical object is different from extended pigment in the same way that an ideal geometrical point is different from a dot. Conceived dynamically, this depiction consists in the generation or construction of an ideal geometrical object here and now. And as with the extension of pigment, the sketching of a line can only be considered as the depiction of extending an ideal geometrical point into an ideal geometrical line, which is being constructed or generated in this manner. This much, at least, is without doubt what Kant meant here by drawing of an ideal geometrical line.

I assume this in order to construct or generate an ideal geometrical object that is an intermediate between point and line. If the dynamic generation or construction of an ideal geometrical line can, indeed, be depicted as an extension of an ideal geometrical point, then I pose the ques-

tion: When I carry out this operation on a blackboard by means of a piece of chalk and a sponge, what does it lead to? With a piece of chalk in one hand, in one motion I undertake to do what I do when I draw an ideal geometrical line; with the sponge in the other hand I immediately follow behind the piece of chalk, so that all that remains is the drawing of an ideal geometrical point and that it never becomes a drawing of an ideal geometrical line.

The answer must come out to the following: what I thereby draw and depict is an ideal object, just as it is an ideal geometrical point or an ideal geometrical line that I generate or construct. But this ideal object is neither an ideal geometrical point nor an ideal geometrical line in the above-mentioned sense. For this ideal object is neither a point in contradistinction to a line, nor a line in contradistinction to a point. As an intermediate between the two, it is in a sense both of them. As the process of its construction shows, this ideal object is nevertheless a possible object; as such, it is like an ideal point and an ideal line existent in the geometrical sense.

Precisely this procedure, it seems to me, is capable of operationalizing and thereby also de-psychologizing Kant's suggestion. For a spatial one-dimensional line cannot at all arise by these means. Furthermore, from this process no other possibility can arise but to pay attention to the drawing itself. And for this reason no other ability is required which one person has and others may not. This operationalization leads furthermore to an objectivization of precisely that which we actually gain as an ideal object when we only pay attention to the drawing itself, namely that ideal intermediate between point and line.⁶

We have thereby also addressed the issue of that monstrosity, which supposedly must be a contradiction, but which demonstrably is no contradiction at all. The object corresponding to a self-contradictory concept cannot be existent; the object in question is however existent, since it is constructible in the outlined operationalized way. Consequently, neither its concept nor its presentation can be self-contradictory. How could it be thought that something like time is not existent, that is, not real in some sense? Therefore, neither the concept nor the presentation of time can be self-contradictory. And indeed, what can be constructed in the pattern of this ideal object is, in the end, absolutely nothing else than a model of time that makes the representation and the concept of time possible for us.

It is thus precisely in the sense of that monstrosity that the constructed ideal object is structurally congruent with time. After all, both are a point

with extension not outside but rather inside itself, and thus a discreteness that does not have a continuity outside but rather inside itself. Precisely in this consists the specificity of time that makes it so different from the specificity of space, so that the two are mutually incommensurable. This complete sequential dynamism is time not only as the extension which is inside it as a point, but also as the continuity which is inside it as discreteness. By contrast, space is that static simultaneity only as the extension that is outside of time, understood as this type of a point, and only as the continuity that is outside of time, understood as this type of discreteness. This last relationship between continuity and discretion in the case of space recurs for this reason again for every dimension of space.

The condition of the possibility that extension can also occur within a point, or that continuity can also occur inside of discreteness is, accordingly, that this occurrence can be only something essentially dynamic, i.e., time, not something static, like space. As monstrous as time is, this essential dynamic is quite possible, as much now as ever. Self-contradictory it certainly is not; it is real and existent precisely as this monstrosity. And the condition for its reality is completely met by the explained model of time.

The ideal object that has the structure of time exists only *while* I set the piece of chalk and the sponge in motion in the above-mentioned way and continuously keep them in motion; that is, it exists only *while* there is this sort of motion. If there is no such motion, there is also no ideal object as a model for time. Extension is then not within the point, and continuity not inside discreteness, but only outside.

The construct of this time-model is thus in fact its constructing. For this ideal object can be given only as one that is being constructed. And it cannot actually be given as one that *has been* constructed, and that is *finished*, or as one for which the process of its construction has therefore been made *superfluous*. On the contrary, this process is necessary, for this ideal object can be given only insofar as it is created always anew, and never as something actually persistent.

Precisely in this respect is it a model of time: time comes and goes, as we are justifiably accustomed to saying. The difference is only that it does not come and go in the way in which any other thing would, i.e., any other thing that also already persists before it comes and still continues to persist after it goes. Time comes only in the absolute sense that it comes into being, and it also goes only in the absolute sense that it passes away. Time always does both at once, so that it in no sense actually persists, not

even between the two states. For this reason, as an existent, time is always existing toward desisting, and never existing toward persisting. Time is thus precisely an absolute change, in which there is nothing that does not change, and in which *panta rhei*.

We now see that the problem of time is self-made. Time, the peculiarity of which is that it is completely dynamic and cannot be counterfeited into space, is thereby free of self-contradiction. It only becomes self-contradictory when it is counterfeited as space, insofar as one treats time as a spatial relationship between point and extension, and thereby eliminates its dynamism and turns it into something static, which time is not. This essential dynamism of time becomes gradually clearer when, through the presented model, one conceptually secures that which is intuitively comprehensible. And the concept needed for this unique model must be correspondingly unique as well.

As an “absolute change,” in which “there is *nothing* that does *not* change,” time is initially determined only negatively, in two ways. In a first step that can also allow a positive determination, the second negation can be replaced by something positive. Time is an absolute change in the sense that in it “there is nothing which is a substrate”; thus, time is a substrate-less absolute change. Time is thus a motion without anything that moves. That means that time is that which is motion without anything that would be *in* this motion, namely, *in* motion which would be a *state* that must, in any sense, be differentiated from that which is *in* motion or *has* it. To replace the first negation as well, which is now to be found in that “without,” only the following remains as something positive: instead of simply being *something in* motion, time is rather *something as* motion. Time is thus something that *is* motion, not something that *has* motion, that is, not merely as its *state*, compared with which it would be something still *other* than motion.

Despite this monstrosity, which is now being conceptually unfolded, time can be intuitively grasped in the form of the presented model. Only the chalk that is being continuously rubbed off belongs to the drawing of my model of time, and not the piece of chalk, or the sponge, or the blackboard. They are only the means for the depiction of this model of time. It can now even be imagined that we have a transparent blackboard, so that I can manage to depict this model of time from the opposite side. It can also be imagined that this blackboard is transparent only in the sense that the chalk being rubbed off is visible, and not the piece of chalk or the sponge. In that case, everyone who is not aware how this motion is pro-

duced, must take it for the relative external motion of a chalk-point; everyone must take it as something identical that is *in* motion across the blackboard and, with reference to this blackboard, as something moving, and *vice versa*.

Yet everyone who is properly informed can take this motion only for what it is: for the constant coming into existence and ceasing to exist of a continually new chalk-point. This point, however, is precisely not something identical *in* motion across the board and thereby also not something *moving* against that board. Nor is it the other way around: the blackboard is not moving *against* the point. It is exactly through this, however, that this motion continuously becomes a *sign of* the very peculiar motion of that ideal intermediate of point and line, or point and extension. If this very peculiar motion cannot be a relative external motion, this can in a positive sense only mean that it must be an absolute internal motion. It is that point which possesses extension only inside itself, and therewith this complete dynamism of something *as* motion.

However structurally congruent this intermediate of point and line may be with time, it could nevertheless remain questionable whether this structure is informative: is not every sort of information lost in the process? Is not this unity—which brings together point and extension or, correspondingly, discretion and continuity—equal to that proverbial night in which “all cats are gray”?

Not at all. For this sort of unity builds an inner structure that repeatedly shows itself as asymmetrical and one-sided when it is further conceptualized. As a way of conclusion, I would like to demonstrate that this is so with a few examples.

The generated intermediate between point and extension should not be understood in the sense that this would be just as much a case of extension in a point as well as a case of a point in extension. Only the former is characteristic of time, which is not counterfeited into space: extension in a point. But the latter (to wit, a point in extension) does not hold for time at all. The latter is characteristic only of space (and of time when it is mistakenly regarded as space). This is made clear by every point in extension—such as a body, or a surface, or a line—where point always has extension outside and never inside itself. In the case of time, however, the point takes precedence over extension. In this case, the point is presupposed for extension, not the other way around; for only in this way can it be extension *in the point*. By contrast, in the case of space extension is presupposed for the point, not *vice versa*, for only in this way can it be a

point in *extension*. Therefore, in the case of space, extension takes precedence over the point.

A further example: the time, which is not counterfeited as space, is extension in a point as a complete dynamic succession. This differentiates it from the static simultaneity of space, which is also characteristic in the sense that it is known as “isotropic,” while the former is, however, “anisotropic.” Time that is not counterfeited as space is in fact of that asymmetrical one-sided direction that we know as the “anisotropy” of time, but which we have so far not been able to define. By means of the presented model we are now in a position to do so: time is asymmetrally and one-sidedly oriented in the sense that it is only oriented toward coming into being, and not also toward ceasing to exist. For time is passing away only insofar as it also comes into being, and by no means the reverse. Time is absolutely not coming into existence because it also passes away. Only its coming into existence is the prerequisite for its passing away, and not at all the other way around. Thus, coming into existence takes precedence over desisting, even though it is not a matter of coming into existence in order to persist. And to that precedence which the point has over extension corresponds this precedence in the most exact way.

This asymmetrical relation reaches, however, still further. Through the construction of the intermediate between point and extension emerges the model of time, but by no means time itself. It would be preposterous to assume that something or someone generates time in the same way in which I generate the model of time. I would make this assumption only when all less pretentious assumptions have been shown to be untenable, for example the following: in the pattern of time nature behaves, where possible, as in the pattern of the Big Bang, that is, as a point that extends *itself*, and which is not extended through something *else*. Nature would correspondingly behave in the pattern of time just as my model of time behaves, if it could also come about without me.

If it could be only a matter of a point extending itself in the presented way, this would have yet another asymmetry as a consequence. The opposite would again be excluded. What would then emerge can be described exclusively in the sense that a *point extends* itself to an *extension*. And this could not be equally well described the other way around, namely as an *extension* that *contracts* itself to a *point*, in that it practically shrivels up. An extension of which this could be meaningfully claimed does not appear at all, just the opposite: only a point does appear. What appears in this process is, again and again, just one single point and never a

still further point, and thus also never yet another point. And nothing is changed by the fact that this point constantly has extension in itself, through that absolute inner motion of its auto-extension. In the pattern of its auto-extension the point takes precedence over its extension, and not the other way around.

This can be conceptually expressed in a suitable way only as follows: as an entity that also must have an identity, this point can only have an identity insofar as it has its identity as a continually renewed difference. For it is only through this that it has the identity of the entity—known to us as time—that via continuous difference constitutes its continuous identity. This difference, however, which this point of time has *in* itself, it has only *for* itself, and not for something other than itself. For anything other than itself does not appear here at all.

This point of time thereby differentiates itself fundamentally from everything that has an identity only in the sense that it differs from other things, that is, in the normal sense of both “identity” and “difference.” Such normal identity and difference this point of time has only insofar as there is something else, something other than the point of time itself. Prior to this normal sense, this point of time has a very specific identity and difference in a quite abnormal sense of both words; it has them in the sense that it possesses difference as something *in* itself as identity, and difference *for* itself as identity, and not the other way around. This is once more an asymmetrical relationship, this time of identity and difference, which exactly corresponds to that of point and extension: not only that the point takes precedence over its extension, so does its identity over its difference, in the entirety of its self-extension.

This asymmetrical relationship, however, is all the more significant as point and extension appear in it as equally original with respect to each other: the point takes precedence over extension not in the sense that first a point appears by itself alone, and then in addition and subsequently it is followed by extension. As the model of time attests, the point appears only along with the appearance of extension. Nevertheless, the point takes precedence over that extension precisely because it appears as a self-extending point.

If all of this is indeed valid, it would have the results to which I can here only briefly allude.⁷

Much too late Kant noticed that “synthesis” cannot be literally comprehensible as “composition,” but only as “extension,” as is especially the case with time and space.⁸ From the point of view of time, however,

this would still have to be understood for space: synthesis would, according to the previous account, occur as the spontaneous auto-extension of a point. Its auto-extension would lead not only to an extension that it would possess—like time—merely inside itself, but also to an extension that it would possess outside itself, like space. Even the extension of space would also be a result of the auto-extension of this point, but in exact opposition to the extension of time. The presupposition for this respective point and this respective extension is also a respective capacity, to which Kant refers as the faculty for “understanding” and for “sensibility.” For both of them would nature, in the form of a highly complex organized body, respectively be a capacity and a possibility.⁹ And there, where nature made these capacities or possibilities for understanding and sensibility real in the existence of both, on the basis of a highly complex organized body, nature would appear as a subject.

This would prevent the unsolvable problematics of schematism from arising in the first place, because, then, point and extension, understanding and sensibility, would become real ever anew only in a respective unity with each other: already in the case of subjective time, but also in the case of subjective space that goes along with subjective time, and so on. New ways would thereby be opened for the deduction of the various forms of a priori elements, insofar as they are the conditions of the possibility of empirical cognition. These a priori elements would be different from each other only through the respectively different relationship between point and extension, understanding and sensibility. They would, nevertheless, in each of these different relationships have in common that understanding and sensibility, or point and extension, and would possess the complete dynamism of time; that is, they would ever again be only a constant coming into existence and passing away.

In spite of this, point and extension would, as in the case of space, initially stand in opposition to each other, according to their form, which Kant calls respectively the “concept” and “intuition” of an object. As concept and intuition are not only formed but also employed in a judgment about this object, this would again be a case of how, in a new way, point and extension enter into a relationship to each other. This time, as well, they would do it despite the fact that as concept, intuition, and judgment, a subject is nothing other than the complete dynamism of time.

Time, until now uncomprehended, that monstrosity, is actually this until now uncomprehended subject. Time as this subject, however, is then a real monstrosity, that is, not an unreal one like that “ghost in the ma-

chine." The subject is an entity and yet not a "res . . .," because it is the complete dynamism of that substrate-less absolute change. ". . . *extensa*" is the entity of the subject only as ". . . **temporaliter extensa**," which Descartes did not take into consideration. ". . . *extensa*" to him means directly ". . . **spatialiter extensa**," and with that immediately "**res spatialiter extensa**" as well. And precisely in that is the reason that a "res . . ." cannot be comprehensible as ". . . *cogitans*." However, as one that is only ". . . **temporaliter extensa**," an entity can perhaps still be comprehensible as a ". . . *temporaliter extensa cogitans*" as well. As such, this entity could be related to a "**res spatialiter extensa**" as the *body* of a subject in a way in which, according to Aristotle, the *anima* behaves toward the *corpus*: as *forma corporis*.¹⁰ As an entity of time it would then be precisely the form of motion of a body. For as the subject in a form of one, namely, its own body, the subject would be exactly that which through itself as that completely special type of constant motion would *place* its body *in* motion or *at* rest: already as a cognizing, and thus first and truly as an acting subject.¹¹

Notes

1. See, for instance, Aristotle's *Physics*, bk. 4, sec. 10, and Augustine's *Confessions*, bk. 11, sec. 14.

2. See, however, A439/B467.

3. Kant's original text says, "*Linie, sofern wir sie ziehen*"; B156.

4. This addition is, however, noticed and translated by Werner Pluchar, in whose rendering the whole sentence reads: "For as regards time, which after all is not an object of outer intuition at all, we cannot present it to ourselves except under the image of a line insofar as we draw that line; without exhibiting time in this way, we could not recognize the singleness (*Einheit*) of its dimension." Pluchar's translation of the *Critique of Pure Reason* is published by Hackett (Indianapolis, 1996).

5. Kant uses phrases like "*bloß*" *auf diese "Handlung" selbst "achthaben"*; they are repeated at B154 and B155.

6. He for whom obtaining this model of time by means of a piece of chalk, a sponge, and a blackboard is not sufficiently precise, can generate it for himself in an absolute and exact way by means of a simple postulate. It involves no contradiction to posit the following: let us assume the dynamic generation of an ideal geometrical line in one motion by means of the dynamic extension of an ideal geometrical point. Such an extension would fix a direction of this extension as well as the direction opposite to it. Since such an extension is contingent, we can also allow the following assumption: let such an extension take place in one motion, so that—at the same time—precisely as much extension arises in one direc-

tion as vanishes in the opposite direction. This postulate leads absolutely and exactly to the same result of an ideal geometrical intermediate between point and line, as does the time-model discussed in the text.

7. For detailed discussion, see my *Die Welt und wir*, vol. 2, pt. 1 (Stuttgart: Metzler, 1999), §§10–12.

8. Compare A360, *Metaphysical Foundations of Natural Science* (4:508; line 1), and Kant's letter to C.L. Reinhold from May 12, 1789 (11:35; lines 10 f. and 27 ff.).

9. The text in which this is made clear—*Über eine Entdeckung, nach der alle neue Kritik der reinen Vernunft durch eine ältere entbehrlich gemacht werden soll* (8:221–3)—apparently is not widely understood in this sense.

10. See *De Anima*, II, 412a20.

11. Translated by Predrag Cicovacki.

Kant and Short Arguments to Humility

Karl Ameriks

I

There is an unusual relationship between Kant's own arguments for transcendental idealism and the way that his position has been presented by later philosophers. In our own time, many important *systematic* philosophers who are not very concerned with the details of Kant's texts have presented his position as a kind of "global" idealism, that is, an idealism that immediately makes ideal *everything* that human beings can sense, or think about, or at least know in any way. Sometimes this kind of interpretation is accompanied by a *sympathetic* attitude, a desire to present Kant as merely rejecting a kind of transcendental realism that is committed to an allegedly *absurd* positing of entities that are in principle wholly beyond rational comprehension. (Critiques of "metaphysical" realism by Hilary Putnam and Richard Rorty are examples of this tendency in contemporary theoretical philosophy; in practical philosophy the work of John Rawls and his students often appears to promote such an interpretation in its preference for "Kantian constructivism" over moral realism.¹) At other times, this kind of interpretation is presented by philosophers who are *suspicious* of idealism of any kind, and who aim to tar Kant with the brush of making everything much too ideal, as if he believed, for example, that our concepts and the properties they refer to are *nothing but* human constructions. (Nicholas Wolterstorff and Roderick Chisholm have suggested versions of this view.²)

These kinds of tendencies can also be found in Kant's *first* readers. It is well known that Garve, Feder, and other early critics used the "tar brush" approach and presented Kant's philosophy as a *bad* version of a global idealism of a "subjective" Berkeleyan variety.³ A lesser known but even more influential development in Kant's time involved an opposite tendency: followers such as Reinhold and the early Fichte sought to "save" Kant by presenting or slightly modifying his position in such a way that it would clearly assert what they took to be an *attractive* global form of

idealism that would leave no significant room at all for a metaphysical “thing in itself,” i.e., a reality wholly beyond the posits of transcendental subjectivity.⁴ A somewhat (but only somewhat) similar division of interpretive tendencies can be found in the work of current scholars—e.g., Paul Guyer, who has criticized Kant for holding onto untenable substantive claims of ideality, and Henry E. Allison, who has aimed to “save” Kant by proposing that transcendental ideality has no immediate negative ontological meaning, and is primarily just a methodological instrument for reminding us of the importance of focusing on conditions of human knowledge as such.⁵

An interpretive approach that *differs* significantly from the “global” reading of idealism noted above can be found in several of my own earlier treatments of this issue, which called for paying closer attention to the specific *basis* from which Kant actually argues for his idealist claims. This interpretation was presented under the heading of a critique of what I called “the short argument for idealism”; the terminology derives from a passage in Reinhold, who thought such an argument was a very good thing systematically even while he allowed that it did not follow the “letter” of Kant’s work.⁶ I use the phrase “short argument” to designate any of a variety of “global” arguments that, like Reinhold’s, attempt to establish transcendental idealism (or at least what is alleged to be its true “spirit”) without going through the actual long and complex steps that Kant lays out in the Transcendental Aesthetic, Transcendental Analytic, and Transcendental Dialectic of the *Critique of Pure Reason*. The crucial feature of these steps is that they all involve starting from reference to *specific* aspects of our forms of intuition, space, and time (e.g., their involving particular synthetic a priori claims). The metaphysical and epistemological arguments in the Aesthetic, as well as the central arguments of the Dialectic in the First Antinomy, focus directly on the features of space and time. The strategy of the Analytic is more indirect; its arguments all center on an investigation of our nonintuitive faculty of cognition, namely, the understanding; but they conclude that, for all our theoretical purposes, the use of this faculty is tied down to a reference to space and time, and hence the claims of our understanding, as well, are essentially restricted by the ideality of these forms of intuition. On this approach, the arguments for transcendental idealism are necessarily “long,” complex, and in part indirect—whereas a “short” argument moves to idealism relatively quickly and directly by beginning not with our specific forms of sensibility but with some much more general or “global” feature of human cognition.

The paradigmatic version of the short argument approach can be found in Reinhold, who began his *Elementarphilosophie* with the bare notion of representation, and claimed that reflection on this notion alone can reveal that all that we can think and know must be ideal, must be only what is “for us” and not “in itself.” Other possible starting points for a short argument emphasize the “global” fact that in knowing we are, in part at least, essentially “receptive”; or, in contrast, that we are, in part at least, also essentially “active.” Similar short arguments can start simply from the fact that we use sensibility at all; or, in contrast, from the fact that we use concepts at all. The first fact might seem to imply that what we are given does not come to us as it is “in itself” but only in a limited and filtered way. The second fact might seem to imply that what we know is in part molded by our own creative ways of organizing data, and so is not left as given “in itself.” (Hegel discussed these contrasting options right at the start of the Introduction to his *Phenomenology* and suggested that they were the most natural ways of moving toward what *he* took to be Kantian idealism—a form of idealism that he still found much too subjective.⁷) There are, of course, many different ways in which arguments need to be, and have been, filled in from these starting points, and there is no reason to assume that in the end they must always turn out to be literally “shorter” than Kant’s own text. The key point here, however, has nothing to do with the terms “short” or “long” taken literally. The main issue, as Reinhold already saw, is simply whether Kant’s *specific* views about space and time, and all of their complexities, play a *necessary* initial role in the argument, or, rather, whether there is a way to try to get to something appealing and at least very much like Kant’s idealist conclusions by abstracting from these specific views from the very *beginning*, and by focusing on something supposedly simpler and more basic, such as representation as such.

Although the distinction between what I am calling “short” and “long” arguments is based fundamentally on a distinction that concerns the argument’s premise, it naturally has implications for how the conclusion of the argument for idealism is to be understood. Sometimes, however, the conclusions need not look very different, and one has to do some work to keep in view the underlying structure of the argument. On a long argument, it is simply space and time—and all and only all that is (epistemically) dependent on them—that turns out to be ideal. On a short argument, on the other hand, the conclusion asserting ideality concerns a realm defined more globally, such as the whole realm of the represent-

able—or the sensible, or the conceivable, or the knowable. One might thus be tempted to assume that long arguments concern only spatio-temporal things, whereas short arguments always concern much more. For many philosophers who use these arguments, however, matters are not this simple, because on some versions it can turn out that there are not, and perhaps cannot be, any more things (or ontological “aspects”) that are known, or sensed, or even “clearly” conceived or represented, than the things that are spatio-temporal. In that case, the actual scope of what is ideal can turn out to be coextensive for the two kinds of arguments. Nonetheless, there is still a huge difference in methodology, if not in ontology, if someone says that what we represent is ideal *because* it is spatio-temporal, while someone else, such as Reinhold (and other German idealists) says that, on the contrary, what is spatio-temporal is ideal *because* it is represented. (Reinhold—and he was not alone—went so far as to argue that not only the ideality but also the specific nature of space and of time can be derived from what is required by a form of representation as such.⁸)

II

Although attempts to find a short argument in Kant, or to construct an improved version of this kind for oneself, can be highly interesting and have a long and ever growing popularity, they remain very questionable for both historical and systematic reasons. Historically, they ignore and make strangely pointless large sections—indeed, the obvious central sections—of Kant's own work. And yet, whatever its interpretive oddities, the move to a “short” approach is understandable, and it is no doubt motivated by more than simply the notorious difficulty of Kant's own writing. There appears to be a systematic payoff that attracts philosophers to the short argument, a hope of connecting Kant's work in a more direct and appealing way with traditional idealist doctrines and debates on skepticism, representationalism, and realism. Nonetheless, precisely because the claims of the short argument are so global, it is hard to see how they could be any easier to establish than conclusions like Kant's own more restricted claims—unless bold global claims such as “all that is represented must be ideal” are meant to be understood in terms of a Pickwickian sense of “ideal” (e.g., as meaning “all that we represent must be representable by us”) that leaves only a straw man to oppose them.

New light on this old issue may arise in the aftermath of a very recent and challenging metaphysical study by Rae Langton, *Kantian Humility: Our Ignorance of Things in Themselves*.⁹ It is a striking fact that from the time of Kant's first readers to our own era, most of the advocates of the short argument approach have had an historical and systematic attitude that is highly *unsympathetic* to mainstream work in classical metaphysics. Working from a presumption that Kant (as the "Critical" philosopher, or, as Mendelssohn called him, "*der Alleszermalmende*," the "all-destroying") had a similar attitude, they have presented their short arguments precisely as an efficient way of moving to a wholesale dismissal of any outdated or "pre-Critical" metaphysics of "ultimate reality." Langton, however, has a very different background and motivation. She is a student of David Lewis and a very able practitioner of contemporary metaphysics. This alone is rare for a Kantian. Until recently, the last half century of English work on Kant was dominated by philosophers who were uninterested in, and not only unsympathetic to, traditional metaphysics as such—and the top contemporary metaphysicians returned the compliment by generally being equally uninterested in Kant. To be sure, Strawson's elegant work generated an important (indeed, dominant) and extensive wave of literature on Kant that might seem to be an exception to this trend. In fact, though, this work kept its focus primarily on epistemological issues (e.g., analogs of the private language argument) and was most influential because of its discussion of empiricism, skepticism, and themes close to positivist doctrines such as the verification principle. It is typical of this work that it provides little discussion of the main issues of Kant's Dialectic ("rational," i.e., philosophical, psychology, cosmology, and theology), or of the central issues in current metaphysics that can be found in more recent "textbooks" by figures such as Chisholm, Armstrong, Lewis, or Loux.¹⁰

Langton opposes this trend by arguing vigorously that Kant can be understood and defended best when he is taken to be a classical metaphysician of contemporary relevance. She takes the "acid test" for this approach to lie in finding a way to make sense of Kant's doctrine of "our ignorance of things in themselves." This doctrine is usually called transcendental idealism, but it is no accident that Langton does not use this terminology explicitly in her title. She aims to save Kant's doctrine by taking it to be simply a kind of quasi-Lockean claim about the unknowability of the *intrinsic* qualities of objects. Since, on her version of Kant, this kind of unknowability goes hand in hand with an assertion

that there *are* such qualities “out there,” it has little to do with idealism as that term is often understood in English, namely, as a globally negative or skeptical notion. Hence, she prefers to say that Kant’s position is a combination of some *realist* metaphysical views (about intrinsic properties of external things) and various sensible *epistemological* doctrines that entail a fundamental “humility” on our part, i.e., a resistance to traditional claims that there is nothing significant at all beyond what we can determine, at least indistinctly, through our own representations. Realism and humility go together. Precisely because one is such a realist about something “out there,” it is only proper not to be overconfident about what we can *know* about it. Nonetheless, this Kantian humility and ignorance is not simply a matter of realism; it also involves the metaphysical claim that certain features are such that we cannot know them as inhering in things precisely *because* they cannot ever belong to things in themselves as such. Hence, they must be “merely phenomenal” or “ideal,” even if they may always appear “real” in an intersubjective experiential sense. Kantian arguments for humility are thus, at least in part, arguments for ideality as well, and so I will sometimes speak interchangeably about arguments for humility and for ideality, with the understanding that neither term by itself entails a rejection of all realism.

I must confess to finding Langton’s notion of Kantian humility to be of considerable interest, not least because several years ago I stressed a similar (but only similar) point, namely, that “transcendental idealism can be expressed as not so much a metaphysical extravagance as rather a principle of modesty, as a reminder that things in their intrinsic character need not be the way that our specific modes of knowing must take them to be.”¹¹ I backed this approach by extensive reference to Kant’s earliest metaphysical remarks (and especially passages concerning our ignorance of the “intrinsic”), a strategy that Langton also adopts. She provides helpful citations of many neglected metaphysical passages that overlap with or correspond to ones that I also thought worthy of emphasis, and so it might appear that her work could be welcomed as fundamentally a reinforcement for those committed to bringing out the metaphysical dimension of Kant’s philosophy.¹² Nonetheless, despite several points in common, and the similarity of key terms such as “modesty” and “humility,” there remain significant differences between Langton’s reading and my own. A consideration of these differences may help to make clearer the main options that remain for those who would pursue a metaphysical approach to Kantian idealism. Only after an analysis of the main claims of her study will

it become clear how much is at stake—how close and yet how far her valuable reading comes to what I take to be the truly Kantian approach.

III

A first and very crucial feature of Langton's approach concerns what she calls the fact of "receptivity." What she means by this fact, first of all, is simply Kant's insistence that we actually are affected by particular external things that we claim to know theoretically. Unlike philosophers (e.g., Berkeley, Leibniz, Malebranche) who allow that something external exists but then deny that it really affects us, and those (skeptics, strict phenomenologists) who deny that there is anything external at all (so that there is nothing that is even in a position to affect us), Kant stresses that we are sensible beings, and that this means that metaphysically and epistemologically we are dependent on how things affect us. The attempt to move immediately from this fact to idealism involves a strategy that I once called "the passivity argument." I explored this strategy as a popular and understandable way of at first approaching Kant, but concluded that it was so obviously deficient historically and systematically that it could not seriously be ascribed to him (*KTM*, 255–9). Langton, however, still seems to have a sympathy for something like the passivity argument since she follows Strawson in regarding the claim that receptivity entails ignorance of things in themselves to be a "fundamental unargued for premise of the *Critique*."¹³ In her own words, "Kant believes that humility follows from this fact of receptivity. . . . If this is correct, then our ignorance of things in themselves is not supposed to be a special [NB] consequence of the arguments about space, time, or the categories; it is to be a general consequence of the fact that human knowledge is receptive" (*KH*, 3).

This passage reveals the fundamental nature of Langton's interpretation, and the fact that she is one more reader who remains very much in the grip of key features of the "short argument" approach to Kant. What is new and interesting about her interpretation is her refusal to be content with a "very short" argument, and her ability to spell it out very clearly in terms of a host of neglected metaphysical details. She sees that a move *directly* from receptivity to ignorance and ideality is hardly self-evident, and so she sets out to find extra premises that could be added to the fact of receptivity in order to give a plausible warrant for the conclusion of what she calls Kantian humility.

An alternative approach, of course, would be simply to resist Strawson's unsupported statement about Kant's so-called "fundamental premise," and to deny that the crucial aspect of Kant's argument rests at all on an argument basically from mere receptivity.¹⁴ It is striking, for example, that receptivity considerations would seem to tie ideality to specifically *empirical* features of knowledge, whereas the arguments of the *Critique* that Kant emphasizes (e.g., the "geometry argument" of the Aesthetic, or the Analogies in the Analytic) proceed by laboriously establishing controversial a priori features of knowledge, and then by asserting always that it is these unusual features that require ideality. This point is quite compatible with his arguing later that all our empirical representations involving space and time—and thus all our sensible representations—are also ideal, since this is precisely because they are determined by the ideal pure representations. Kant's famous letter to Herz, and his discussion in the *Critique* about the need for a Transcendental Deduction, make clear that the issue of the ideality of representations emerges for him from questions specifically about concepts that carry with them necessity and a priority (viz., the non-sensible representations that are the categories), and these are not a matter of our mere receptivity.¹⁵ All this is ignored by Langton's starting point, and this reinforces the usual suspicions about any version of the short argument approach.¹⁶

Langton's analysis is, nonetheless, unusually intriguing even if its most general form follows an old and disputed pattern. Unlike other advocates of a short argument, she believes that Kant has a substantive and understandable notion of something beyond our theoretical knowledge. In this way she is much closer, at least in her conclusions, to "long" interpretations that stress the nonideality of properties not tied to our specific forms of space and time. For her, the Kantian "in itself" is not a mere analytic or methodological notion, not something that we cannot reach simply because it is *nothing* at all, let alone something that has *no meaning* at all. It is rather something that we *properly* "yearn to know," something that Kant speaks of as involved with real features such as affection, and for which he reserves "distinctive and inner predicates" that he does not allow in the realm of appearances.¹⁷ Unless we hold that Kant believes things can exist without "inner predicates"—something he resists in all periods of his writings and lectures—there is no reason to *deny* a substantive metaphysical content for (what Kant meant by) the "in itself," even if we cannot affirm positive (theoretical) *epistemological* claims about what it is like.¹⁸

This strategy raises an old and obvious problem that Langton confronts right from the start. Since Kant expressly denies *knowledge* of things in themselves, how can he consistently allow an affirmation of the “inner predicates” and causal relations that have just been noted? Langton’s answer is a variant of a strategy that is itself old and obvious—and still perfectly respectable, even if strangely neglected. The strategy is simply to say that Kant does not mean to block *all* kinds of knowledge of things in themselves, but only certain types. This strategy seems forced on us by minimal charity, for without anything like it, Kant would be talking wild nonsense in his many general discussions of things in themselves, in his many positive characterizations of them from a practical perspective, and in his many slightly more specific remarks about them as having “distinctive and inner predicates.” As long as we do not claim to know positively the “inner predicates” of specific things, it is not clear that Kant’s general concern with humility and idealism is done any serious harm; there is still quite a lot left to be “humble” about, and also a lot that can be called “ideal” in a specifiable and significant sense. Langton’s proposal is that Kant is concerned simply with denying our knowledge of the *intrinsic* properties of things, and thus with denying the “transcendental reality” of relational claims. When he says that things in themselves affect us, or that we have detailed knowledge about the domain of experience, all we have are many bits of *relational* knowledge, and thus precisely not illegitimate (for a Critical philosophy) knowledge of things “in themselves,” i.e., things in their *intrinsic* nature.

This interpretive strategy suggests an interesting way that Kant can be taken to make a claim about our necessary ignorance that is at least meaningful and can seem, on a sympathetic reading, to have some chance of being consistent with the rest of his philosophy. Moreover, it reveals how Kant’s position contrasts very dramatically with that of his main opponent, Leibniz. *If* Kant held, as Leibniz generally was taken to hold, that relational claims are in principle *reducible* to intrinsic claims about individual substances, then he (Kant) would be caught, after all, with a dogmatic commitment to strong knowledge claims—and ones clearly inconsistent with the rest of his philosophy—about the natures of particular things in themselves. The relational knowledge that we possess about phenomena *would entail* (albeit in a way that we need not be able to see clearly) intrinsic truths about things in themselves. The crucial antecedent of the conditional claim cannot be asserted, however, because Kant clearly emphasizes that on his view there is precisely *no such reducibility*

of relations.¹⁹ And it is precisely this insistence that Langton uses as the essential supplement for the receptivity argument. Her argument is that the relational nature that characterizes our knowledge as receptive, combined with Kant's doctrine of the nonreducibility of these relations, is what implies humility: All we can know is relational, but the relational does not reveal the intrinsic (given the nonreducibility doctrine), so what we know are, in this sense, "mere phenomena," not things in themselves. Or, to be more precise, Kant holds that what there is in itself must have *some* intrinsic nature, so we do have that kind of indeterminate knowledge *about* things in themselves, but this is still not any kind of illicit theoretical knowledge *of particular* things, or even of the *determinate* nature of things in general (*KH*, 13). This reading has some obvious advantages. It can easily supply an understandable and modest meaning for Kant's doctrine that our standard knowledge claims have "mere phenomenality" as well as "empirical reality," and it has no problem applying to the claims that involve the essentially relational attributions that occur within the framework of modern theoretical science and that are of special significance for Kant.²⁰

IV

Langton finds grounds supporting her approach in considerations about interaction that go back to Kant's very earliest texts. In his 1747 essay on *Living Forces* Kant defended the claim that "substances can [NB] exist and nevertheless have no external relation toward others at all."²¹ And from his 1756 *Physical Monadology* on, Kant argued that matter (as a play of forces) is not strictly speaking substance, but only *substantia phenomenon*, something treated *as if* substance,²² and that "besides external presence, i.e., relational properties, there are other intrinsic properties without which the relational properties [forces] would not exist, because [otherwise] there could be no subject in which they inhered."²³ As Langton notes, Lewis White Beck and others have thought that these claims are still basically Leibnizian, and were written to express a view (of Kant's at that time, supposedly) that relational properties such as force and space are *fully grounded* in substance in a Leibnizian way, as a *phenomenon bene fundatum*.²⁴ In fact, however, all that Kant borrows from Leibniz here are the ideas that a substance is something that ultimately exists on its own, i.e., is a subject or bearer and in no way an accident, and that "sub-

stances in general must have an intrinsic nature which is therefore free of [i.e., not dependent on] all external relations.”²⁵ Unlike Leibniz, Kant holds that the relational properties of finite substances *cannot* be reduced to, or “fully grounded” on (i.e., be shown to supervene on), their intrinsic properties. Moreover, although for Kant a substance is independent, it need not be absolutely independent in being or determination. That is, not only does Kant say a substance does not have to be responsible for its own existence (contrary to Spinoza), but he also allows—*unlike* Leibniz—that there is real interaction of finite substances, and he affirms that the properties of mundane substances are in part due to causal effects of other substances. In an early argument (in the *Nova dilucidatio*, 1755) for his “Principle of Coexistence,” Kant contends that substances do not exist together (“in the same world”) unless they are in some causal interrelation.²⁶ He adds to this a “Principle of Succession” that goes even further (and was eventually given up) and asserts that *no* change, not even in thought, can happen within a particular (finite) substance without some cause from *outside* that substance.²⁷

All these considerations point to a fundamental and enduring Kantian distinction between intrinsic and relational properties. Intrinsic properties characterize what the substance is “in itself,” and a substance could not exist without *some* of these. Relational properties, on the other hand, are such that at least some substances *could* exist without any of them, and also such that, although they represent the actual relational truth about substances, they do not logically, let alone epistemically, entail (through reduction or supervenience) what is true about the substances intrinsically. Since the relational properties exhaust Kantian “experience,” or what we know about the phenomena, i.e., the items that we can actually empirically determine,²⁸ and since the intrinsic properties seem to be in principle beyond our access (since any candidate we offer for them appears to be replaceable by a set of relational properties that we can determine empirically, and there remains a posited “substrate” that is always beyond whatever determinations we make, since these always turn out to be relational), it can seem understandable and attractive to equate the relational with what is phenomenal, and the intrinsic with what is in itself. The claim of our ignorance of the “in itself” then appears fairly innocent and compatible with all sorts of causal or “affection” claims about things in themselves. Whereas previously it might have seemed dogmatic to speak of any causal relations of things in themselves, now, on Langton’s interpretation, we can say easily enough that precisely by speaking of causality in terms of relations, we are not at all speaking illegitimately

about what things are like “in themselves,” and yet, at the same time, we need not say that our causal knowledge is sheer error or illusion. On the contrary, it is a completely proper representation of the relational—but merely relational—truths about what there is.

V

All this can seem to leave us with an appealing dissolution of some old mysteries, and a nice metaphysical version of a short argument for Kantian idealism. But can it be all that simple? Can these early texts already reveal the essence of the transcendental idealism that Kant was to argue for much later and in a more complicated but (supposedly) substantially unrevised way?

Although these texts do involve some important characterizations of the notions of substance and the “in itself” that Kant showed no signs of retreating from in his later works, they also seem to me to be in some very significant ways *unlike* his views later in his Critical period. It is striking that, unlike the *Critique*, the early works move directly to strong metaphysical claims, rather than starting from epistemological or “merely experiential” ones, and that at the start they concern mere sensible rather than a priori and necessary relations. Langton may think that it is a good thing (part of what she regards as Kant’s “empiricist” background) if little more than our mere receptivity points to a line of thought that seems only a short step from making intelligible a strong distinction between knowable phenomenal items and unknowable things in themselves; for her, “humility does follow from receptivity, given irreducibility” (*KH*, 126). But this step still seems to me too short and easy to be valid, especially since it leaves key changes in Kant’s doctrines through the 1760s and 1770s, as well as nearly all the details of his main arguments in the Critical period of the 1780s and after, without any clear relevance for his most important doctrine, the thesis of transcendental idealism.

But if something has gone wrong here, what is it? I would be the last to counsel against giving more attention to Kant’s metaphysical statements and early works. Nonetheless, it is fair to ask exactly how far one can go with only these works and the kinds of statements to which Langton limits herself. The doctrine of transcendental idealism is so difficult that one should not turn away from considering any serious hypothesis that might shed light on it.

In this spirit I once examined, under the heading of an argument called “the essence theory” (KTM, 267–70), a hypothesis somewhat similar to Langton’s. This theory involves equating phenomena with contingent properties, and the unknowable aspect of things in themselves with essential properties. Given Kant’s unqualified rejection of any Leibnizian notion that we can come to a (theoretical) determination of the substantive essences of particular things, the essence theory offers at least one relatively simple way of explaining why Kant might have thought that *we cannot* know things “in themselves.” Nonetheless, I noted (as I have here with regard to Langton’s interpretation) that such a theory does not seem to correspond to the complex epistemological considerations that Kant in fact emphasized in the course of the development of his idealist doctrines. Moreover, there seems to be no reason to deny that (even for Kant) there could be contingent (and unknown) rather than essential properties that could also belong to a thing in itself. Given the apparent possibility of contingent properties of things in themselves, mere essentiality cannot be assumed to be a clear necessary condition of nonphenomenality.

A similar problem can be pressed in the end against Langton’s approach, even if that approach has features that keep it free from some problems that defeat the essence theory. She stresses tying the thing in itself to intrinsic properties, and tying the phenomena that contrast with the thing in itself to nonintrinsic, i.e., irreducibly relational properties. This strategy has implications that *partially overlap* with the essence theory insofar as essential properties *can* be possessed as intrinsic and nonphenomenal, and (for Kant) there are specific relational properties that obviously *cannot* be essential for a thing in itself as such. Nonetheless, although it is essential that a thing in itself has *some* intrinsic properties, it is not clearly the case that all the intrinsic properties that it has (or perhaps any of its distinctive properties) are essential to it. This leaves the essence theory with a problem that Langton’s approach does not have to worry about. The contingent but still nonphenomenal properties that give that theory trouble can be ascribed to things in themselves without any problem in her view.²⁹

Despite this advantage, there remains a difficulty for Langton’s view. The view appears to entail two controversial theses: (1) that there can be no intrinsic but merely phenomenal properties, and (2) that there can be no relational but nonphenomenal properties. Against the first point, one might try to argue by supposing, in a way common to some views in contemporary debates, that “sense data” as such, or “qualia” of some

sort, could be counted as both intrinsic and manifestly phenomenal. A contemporary advocate of qualia might stress precisely that they have a simple, immediate, and not essentially relational character, and that they have a presence that remains certain however things might go with the development of our ordinary relational and scientific knowledge. This is an interesting objection, but I would concede that it is not one that *Kant* himself would press very far. The immediacy that "qualia" advocates emphasize is something that Kant (for better or worse) probably would not regard as having any *cognitive* character as such. And, insofar as we can move on to know what qualia are, Kant seems to presume that they should fit into a general and relational theory about the properties of the space-time (and hence, on Kant's view, merely phenomenal) domain.³⁰ Hence, if qualia are completely nonrelational, then they are not knowable, and so they are not an instance of the Kantian phenomenal domain, which is exhausted by what can in principle be known through (what are at least in part) empirical means. And if they are relational and knowable after all, then they lose their character as sheer intrinsic qualities.

Even if objections to thesis (1) could be turned back in this way, thesis (2) remains vulnerable to a more serious and genuinely Kantian objection. The key issue here is whether there could be relational properties that ever attach to Kantian things in themselves as such. This might seem to be ruled out by definition; once a property is relational, then one might say it just is what Langton thinks Kant means by "not in itself." But this strategy threatens to appear like the "analytic" and "anodyne" high road that Langton criticizes in others. Why should we think Kant holds to the stipulative definitional route of making relations contrary to the very meaning of what can be possessed by any thing in itself *as such*? There is, after all, at least one possible property of things in themselves that seems relational and of great concern to Kant, namely, freedom. He clearly wants to be able to say, with at least possible truth (whatever the warrant may be; nothing here rests on our being able *theoretically* to know that we *actually* have the exercise of such freedom) that a subject is free, absolutely and "in itself." This freedom must be more than an intrinsic feature entirely within a subject; it must be an actual relational capacity, even if it is also said to be within the subject "in itself" because there is no room for it as such (given Kant's empirical determinism) in the subject as phenomenon. That is, it must be a capacity of the subject really to affect something other than it, to be an uncaused cause of phenomenal events and some kinds of effects on other substances, especially other people.

Otherwise, Kantian notions of morality and freedom are toothless beyond belief, no better than the fantasy of pre-established harmony.

Freedom is not an odd or atypical example of a possible property of a thing in itself; it is clearly the very kind of property for which Kant designed his theory. Hence, making room for it must override the benefits of a merely definitional strategy that would keep a clean division of relational properties and properties of things in themselves. The fact that, on Kant's mature theory, this property is one that in principle cannot be theoretically known, in no way disqualifies it from being a property of a thing in itself as such (such unknowability is a feature that belongs in any case to all the properties of the subject in itself on Langton's theory as well). Langton herself emphasizes that the relationality of a property possessed by a substance can be "conceptual" and need not depend on the actual existence or state of another substance (*KH*, 38). Thus, regardless of whether there exists something else that a subject's freedom affects, it still can be the case that the subject in itself has the (relational) property of being a free being, of having a nature that *would* allow it to act freely and effectively in particular ways in an actual situation with distinct substances. The fact that this property in turn might require some more basic intrinsic properties does not mean that the property cannot be relational and be possessed by the thing it itself as such.

The example of freedom is only one of a group of similar properties that are highly relevant to Kant's system. Even though Langton makes extensive use of Kant's earliest work, she, like most interpreters, has little to say about Kant's views in the 1770s, after he had already introduced a doctrine of transcendental ideality in his *Inaugural Dissertation*. In the *Dissertation* Kant makes his initial argument for the transcendental ideality of space and time (as "forms of the sensible world"), but he still holds to the doctrine that there is a kind of nonideal theoretical knowledge that we can have of the "form of the intelligible world." It is striking that this knowledge has as its prime example an argument concerning the interaction of substances and the unity of the world of finite beings (§19, 2:408). It is very clear from the text that Kant still regards this knowledge as at once relational and concerning things in themselves as such. In his later work, in the *Critique of Pure Reason* and after, he severely tightens his limitations on what we can know about things in themselves, and he argues that not only the spatio-temporal form of the objects of experiences, but also all the causal laws that we can know for them (which in the end also indirectly cover all the content in our theoretical determina-

tion of the self) are essentially mediated by the forms of space and time and thus turn out to be transcendently ideal. Nonetheless, this change does not affect Kant's *definition* of transcendental idealism, but only his conclusions about some of the features of our faculties of knowing and the scope of claims about what it is that is merely ideal. He never indicates that he wants to go back on what he has argued in the *Dissertation* in such a way that it would become *immediately* entailed by the mere relationality of a property that it cannot belong to a thing in itself. It is true that all the items Kant claims to be ideal are also relational, but all the weight in the actual arguments that the *Critique* makes for the ideality of various relational features (such as space and time, or self-knowledge) rests on *specific* characteristics of the relations involved. In each case ideality is concluded only after a discussion of a particular spatial or temporal relation that is involved, and only after noting that there are specific reasons for saying ideality is the best explanation for the spatio-temporal dimensions of the things of experience, or of our a priori knowledge of their spatio-temporal form.

Matters are complicated somewhat by an unusual argument in the *Critique* that includes the statement that "a thing in itself cannot be known through mere relations" (B67). Note, however, that this is a statement about *knowledge*, not about the properties of a thing as such, and it is compatible with, and probably aimed primarily at expressing, the innocent point that a thing cannot *have mere* relational properties without any intrinsic properties at all. The statement by itself does not rule out the possibility of a relational property possessed by a thing in itself, or even of a whole set of such properties had by things composing an "intelligible world."³¹

This is not to play down in any way the argument beginning at B67, for it is one of the most important and least understood passages in the *Critique*.³² The argument is one of many instances where Kant added a point in the second edition in order to clarify his theory of self-knowledge and to meet the vehement objections that were made to his conclusion that we have only ideal theoretical knowledge about the self. This is a conclusion that went against Kant's own earlier claims in the 1770s, and he did not reach it without careful reflection. He surely understood that the strong claim in the *Critique* that all our determinate theoretical knowledge is merely ideal would have to meet the main challenges of his predecessors. The main challenge here went back to Leibniz and his claim that we do know at least our own self as a substance relatively directly through inner intuition. Kant set out an argument at B67 to meet this challenge head on,

and to bring “closure” to his idealism by asserting that even the last “Cartesian” reserve of knowledge did not violate his doctrine about how our theoretical philosophy is limited to merely ideal determinations. Kant’s discussion is quite complicated, and one can easily misunderstand its conclusion that the mind “then intuits itself not as it would represent itself if immediately self-active, but as it is affected by itself, and therefore as it appears to itself, not as it is” (B69).

There are a number of “red herrings” in this argument that our self-knowledge is ideal *because* it involves “self-affection.” Read in isolation, one could imagine that this passage vindicates the passivity argument, or Langton’s argument from the fact of receptivity. But the slightest reflection reveals that it is crucial to ask precisely what it might be about “affection” that leads Kant to link it with ideality. A number of suggestions arise immediately from this text and related passages. Our self-knowledge is “affected” insofar as it is in general passive and sensible, and also insofar as (as Kant stresses) we have to exercise, and be subject to, a specific act of “self-affection” that gathers our thoughts into a synthesis of apperception, and thereby raises them from mere inner sense to knowledge.³³ There is also a distinct kind of affection that Kant takes to be present in self-knowledge insofar as he holds that, at least sometimes, we need not only the judgmental aspect of apperception but also special acts of attention (see B157n.) and reflection—acts that can be distinct from each other and insufficient for apperception. Kant stresses that these acts tend both to add distortions and to undermine the possibility of an introspective science of mind. Also, in contrasting our self-knowledge with that of a creative intellect, Kant seems to be suggesting that the limited activity of our finite mind testifies to another way in which our involvement with affection introduces a possible distorted character in what we know.³⁴ What is remarkable, however, is that all the different aspects of “affection” that I have just mentioned clearly do not by themselves amount to a convincing reason for saying that we are *in principle* limited to knowing the self “only” as it appears. Passivity and activity, reflection and attention, can all introduce incidental distortions, but they do not amount to a proof—and were never anywhere offered by Kant as a proof—that we *cannot* (theoretically) know the self in itself *at all*. People have been aware of these features for a long time without being led into anything like a doctrine of the ideality of all self-knowledge.

One cannot help but ask if Kant has some other line of argument in mind—and, in fact, he does, and the text makes the relevant line quite

obvious. Right before the passage cited from B69, Kant says that “*the form of this intuition*, which lies antecedently in the mind, determines, in the representation of time, the mode in which the manifold is together in the mind, since it then intuits itself not as it would represent itself if immediately self-active, but as it is affected by itself”(B68–9).³⁵ The same general point is made earlier in the argument when Kant concludes “that the subject can be represented through it [i.e., the “form of inner sense,” time] only as appearance, not as it would judge itself if its intuition were self-activity only.” (B68; cf. A37/B54 f.) In other words, the only way of giving any cogent sense to Kant’s use of “affection” here is to take it to stand for the specific kind of affection or limitation that we are subject to because we are subject to the specific *form of time*, a form whose ideality has been argued for earlier on a variety of specific epistemological grounds. Our being merely affected as such, even inwardly, in just any ordinary way, provides no ground at all for concluding that we *could not* know our self in itself as such, and it is no wonder that Kant never fell back on such a simple kind of argument.³⁶ This is not to say that we must regard as convincing or clear Kant’s own complex argument, in either its claims about the ideality of the form of time or about the ideality of all our self-knowledge insofar as we are subject to being “affected” by this form. But whatever the difficulties of Kant’s long arguments for ideality, it is very clear that he made them, and made them repeatedly, and it is not clear why they should be ignored for the sake of arguments that have severe problems of validity and obvious problems of interpretive relevance.

VI

The problems with the short argument to humility are not limited to the difficult area of the self and self-knowledge. There are fundamental objections to accepting this argument *in general*. The specific passages on self-knowledge reveal the basic inability of the argument (as long as it does not resort to being stipulative) to address the claim that our knowledge must be ideal *in principle*. The fact of receptivity and some other elementary reflections can, to be sure, show that there is a distinction to be made between the qualities that we ascribe to things themselves and the qualities that we can happen to have in our own mind simply as the result of our contingent sensible situation. At any given moment it is always possible that our sensibility can give us only a distorted view of how

things are in themselves. Moreover, if we accept the rejection of a reducibility relation between sensible properties and intrinsic ones, we can say, as Langton does against Leibniz, that we see through a “broken mirror,” and we cannot *assume* that even in an “indistinct way the relational properties of our phenomenal knowledge reflect the inner natures of things” (KH, 127). Nonetheless, this does not by itself rule out the possibility of our making lucky true guesses, or even some kind of indirect arguments, from our sensible data, to those inner natures. Hence, insofar as the short argument approach relies only on considerations of receptivity and irreducibility, it is still far from presenting an argument that historically or systematically can begin to match the grounds that Kant had for blocking *in principle* our claims to know things in themselves, and it is very far from vindicating any claim to blocking even the possibility that we might have some true hypotheses about what those things are like.

In addition, there are general reasons for supposing (against what was called the thesis [2] of the short argument) that Kant himself was committed to properties that are simultaneously inherent in things in themselves and in part relational. Langton focuses almost entirely on Kant’s very early discussion of causality, in the 1740s and 1750s, but his interest in this issue came to its peak in his *Dissertation* of 1770. Throughout his early work, and despite numerous terminological complications, Kant sought to defend a modified version of a “real influx” theory that would simultaneously respect the independence of finite substances and provide for a more than nominal (pre-established harmony) or wholly externally imposed (occasionalist) unity.³⁷ His *Dissertation* solution (given up later solely for epistemic reasons, not because of any noted metaphysical difficulties) was to propose that the unity of mundane substances can come about only from an underlying common cause, namely, God.³⁸ However, Kant stresses that even God cannot provide a true unity of such substances in just any way. He must respect a condition of what I have elsewhere called the “Restraint Argument,” namely, that there is something in the natures of particular things that distinguishes them and makes it the case that the properties that God causes are not simply His own.³⁹ As Kant notes, if God causes a thought in me, for that thought to be mine and not His, there must be something in me that can receive the thought.⁴⁰ This point implies not only that I must have, as has been stressed so far, some simple intrinsic properties, but also that, given the way the world actually is, I must have some properties in myself that are fit to be acted upon by God and appropriate for fitting together with other substances in

one world; in other words, properties with the right kind of relationality.

Langton does not explore this possibility for things in themselves although she has an extremely useful discussion of the powers of substances. She notes that they do not directly entail particular effects on other substances because, given different laws of nature, the exercise of the powers will naturally have different effects, and perhaps no effects at all. A typical causal power thus has relationality built into its very *concept* (e.g., a power to bleach terrestrial things can be “in” the sun, even if an obstacle such as clouds could keep the sun from ever actually bleaching such things), without this entailing a dependence of it, or the substance bearing it, on any other actual finite substance (*KH*, 116–18). After noting this point, however, Langton rejects the thought that we should therefore characterize Kantian powers as *intrinsic* properties. She points out that if the laws of nature were to change radically, then a thing that has the same intrinsic properties (properties it would have if it existed alone) as it has now (when it has, for example, the power of attraction in this world) could, in a different world, have a very different power instead. A causal power’s properties thus seem to be “extrinsic,” because they are manifestly subject to the contingent laws of nature, which laws in turn, in Kant’s philosophy, are subject to the infinite free power (and what might be called the “superadding” habits) of God.⁴¹

This view leaves an enormous “looseness” between intrinsic properties and causal powers, a looseness that Langton welcomes because it fits the sharp line she draws between things in themselves, defined by intrinsic properties, and merely phenomenal objects, which are understood in terms of powers, i.e., relational properties that (given the point just made about the laws of nature and the whims of God) do not “reduce” to the intrinsic properties. Nonetheless, I do not see why a line has to be drawn so sharply in this way, nor why the nonreducibility of relations should lead to a dismissal of all possible (i.e., weaker than reductive) “grounding” relations between relations and things in themselves. Given Kant’s own Restraint Argument, it seems fair enough to say that changes in the laws of nature, even by God, need not make *all* the difference.⁴² Even if I am put into another world, so that I attract things in a different way than by Newton’s laws, it seems possible to believe that there is something about me, even in that world, that has something to do with the power of attraction that I have taken on. Consider another example (inspired by the Restraint Argument): If the way that I think something is *simply* a matter of the laws of the world that God puts me into, then it seems that there is

nothing relevant left to me as such, and I am a mere blank tablet on which absolutely anything can be drawn in a way in which I make no difference at all. A way out of this difficulty would be to allow that each thing has a set of relational properties already within it as a thing in itself. Very roughly speaking, these properties would allow for different powers to manifest themselves differently in some different contexts, so that if substance S is switched from world 1 to world 2, then it could take on form S2 rather than just any strange form, and this property S2 would be a *joint product* of its own complex nature and the world in which it is placed. In a similar way (only as an analogy, since the properties here are all phenomenal), we might say that when a paper looks yellow in a certain light, this is of course very much the product of the light, but it is also to some extent a product of the paper “in itself,” for with a different color “in itself,” it could have looked somewhat different in this light (even if in some other cases this would not make any difference). Similarly, for Kant a subject might have a good will in itself, as a relational disposition to the moral law and its implications, and yet how this will manifests itself—e.g., in extensive charity or mere justice—can be allowed to differ significantly depending on the world involved, even while the person “in itself” remains an ineliminable source of the actions.⁴³ In other words, there may be a way of building relational properties into a thing in itself without having to fall back into saying that all these properties involve a reducibility or mere phenomenality of relations.

VII

A related and final difficulty for Langton concerns the well-known general problem of *transcendental* affection. A traditional objection to Kant’s philosophy was that he could not consistently claim to be ignorant about things beyond experience and then also say (as he does) that such things affect us as receptive beings—for claims about this kind of affection would seem to imply knowledge about what is, in part, beyond experience after all. One traditional reply to this objection (and one to which that I am still sympathetic) would make use of a general strategy that Langton also adopts in taking Kant’s claim about ignorance as meant not really to rule out all kinds of claims beyond experience, but only those that involve some kind of relatively specific (theoretical) determinations. Along something like this line, one might argue that there is nothing wrong with positing sim-

ply that there is something “underlying” experience—because the very notion of a “mere phenomenon” in some way implies that there is something nonphenomenal.⁴⁴ Langton uses a variant of this argument that seems to rely not on a mere analytic implication (of “mere phenomenon”) but rather on a fundamental presumption that, for things to exist at all, there must be substances, and each substance must have some qualities that are intrinsic rather than merely relational. One might well ask where this presumption comes from, but Langton gives plenty of evidence that Kant never questions it (*KH*, 32); and Kant is hardly alone in holding on to the presumption.

This reply does not solve all problems, however, and there is a difficulty that arises precisely because of the main feature that Langton’s approach relies on, namely, the sharp parallel distinctions between intrinsic and relational, and “in itself” and “merely phenomenal.” The difficulty has to do with the problem in principle of finding a systematic Kantian “location” for transcendental affection. Kant clearly asserts, and Langton can hardly deny, that for him there is *transcendental* affection, and not merely the kind of relations that occur within the causal context of phenomenal items. But this is to say that transcendental affection is at once nonphenomenal and relational—and Langton’s short argument leaves no room for such a possibility. She can easily call ordinary relations (of matter, force, space, etc.) relational and merely phenomenal; and she can just as easily say that the intrinsic and nonrelational characteristics of a thing in itself are nonphenomenal. But transcendental affection (which is simply a more indeterminate instance of the general structure underlying the property of absolute freedom discussed earlier) cannot be either simply phenomenal or simply nonrelational.

In desperation, one might contend that Kant ought not to have spoken about this kind of affection, but his reasons for doing so are no stranger than those for the other principles that Langton allows. Moreover, accepting such affection still need not lead into an undermining of Kant’s general Critical stance, at least as long as one has the charitable and relatively loose way of reading Kant’s limits on knowledge that Langton also encourages. To allow the possibility or even the actuality of some transcendental affection is still not to claim to know illicit determinate facts about specific nonphenomenal things. It is not to violate the very bounds of sense. Someone might say that nonempirical causation of any type, for example of nonspatial-temporal items on spatial-temporal items, is immediate nonsense; but the fact is that such talk was allowed for centuries in rational discussions of broadly Platonic and theological contexts, and

to claim now that it is dogmatic nonsense is to make a dogmatic metaphysical claim oneself. The main problem here is not with the *sense* of Kant's idealistic theory and its involvement with a doctrine of nonempirical affection. That theory can have sense even if it has obvious problems of vagueness and justification. The main problem is with a line of interpretation of the theory that cannot make clear room for even the possibility of this kind of affection (the "acid test" of Kantian hermeneutics). For this reason, the difficulties of Langton's impressive efforts force us back once again to seeking an alternative to the whole short argument approach—an alternative that is alive to Kant's metaphysical roots but that stays closer to the central sections of Kant's Critical texts.

Notes

1. On Kant and contemporary metaphysical anti-realism, see my "Hegel and Idealism," *Monist*, 74:1991, 394–6. On Kantian moral constructivism, see my "On Schneewind and Kant's Method in Ethics," *Ideas y Valores* 102:1996, 28–53.

2. On Wolterstorff, see "Hegel and Idealism," 394; on Chisholm, see my "Chisholm's Paralogisms," *Idealistic Studies*, 11:1981, 100–8.

3. See the materials contained in Johann Schultz, *Exposition of Kant's "Critique of Pure Reason,"* tr. and ed. J.C. Morrison (Ottawa: University of Ottawa Press, 1995).

4. See my "Reinhold and the Short Argument to Idealism," in *Proceedings: Sixth International Kant Congress 1985*, eds. G. Funke and T. Seeböhm (Washington: The Center for Advanced Research in Phenomenology and the University Press of America, 1989), vol. 2, pt. 2, 441–53; "Kant, Fichte, and Short Arguments to Idealism," *Archiv für Geschichte der Philosophie*, 72:1990, 63–85; and *Kant and the Fate of Autonomy* (Cambridge: Cambridge University Press, 2000), chapter 2; cf. Kant's eventual rejection of Reinhold and Fichte in his letter to Tieftrunk, October 13, 1797 (12:207–8) and his public declaration, August 7, 1799 (12:370–1).

5. See my "Kantian Idealism Today," *History of Philosophy Quarterly*, 9:1992, 329–42.

6. "Thus one can begin to understand in a shorter way the impossibility that Kant established of knowing the thing in itself. The thing in itself is not representable, so how could it be knowable?"; K.L. Reinhold, *Versuch einer neuen Theorie des menschlichen Vorstellungsvermögens* (Jena: 1789), 255 (cited in my "Reinhold and the Short Argument to Idealism," 442).

7. See my analysis in "Hegel's Critique of Kant's Theoretical Philosophy," *Philosophy and Phenomenological Research*, 48:1985, 1–35.

8. See my "Kant, Fichte, and Short Arguments to Idealism."

9. Oxford: Clarendon Press, 1998; henceforth *KH*.

10. A very clear and typical review of this approach is provided in Michael J. Loux, *Metaphysics: A Contemporary Introduction* (New York: Routledge, 1998), 7–10.

11. See my *Kant's Theory of Mind* (Oxford: Clarendon Press, 1982; 2nd ed. forthcoming), 7; henceforth *KTM*. It should be noted that in saying that transcendental idealism "can be expressed" in this modest way, I was not endorsing the view that this was the most accurate way to define it, and in fact I went on to argue that Kant himself seems committed to the bolder claim that certain features (of "our specific modes of knowing") not only "need not" be, but also cannot be, ascribed to things in themselves.

12. Langton has independent reasons for focusing on many of the surprisingly neglected Kantian metaphysical doctrines that I have stressed in my earlier work. See, for example, her treatment of "phenomenal substance" and matter as not substance in the pure Kantian sense (*KH*, chapter 3; cf. *KTM*, chapter 1, 299 n. 79, and *passim*; this issue is also discussed in James van Cleve, *Problems from Kant* [New York: Oxford University Press, 1999], 120); her stress on the importance of the doctrine of the irreducibility of relations in Kant (*KH*, chapter 5; cf. *KTM*, 36, 65–7, 267–77); and her stress on the contrast between Kant's "scientific" notion of experience and phenomenalist definitions of it, a contrast based on what Kant says about how things of experience are known "through" representations (and are not equivalent to collections of them) as well as on his insistence that unobservable items, known through laws and not sensations, can be paradigms of "empirical reality" (*KH*, 159, 143; cf. *KTM*, 113).

13. *KH*, 4, referring to P.F. Strawson, *The Bounds of Sense* (London: Methuen, 1966), 252.

14. Langton cites from the *Critique of Pure Reason* one passage that is apparently meant to support Strawson's claim: "Properties that belong to things in themselves cannot be given to us through the senses" (A36/B52, cited in *KH*, 23). Note that this passage does not say that we cannot know things in themselves simply *because* we use sensible representations. It thus does not show that the unknowability in itself of what is given in our receptivity is a "fundamental premise" rather than a conclusion in Kant's sequence of argument.

15. See Kant's letter to Herz, February 21, 1772 (10:131–2): "Thus the passive or sensuous representations have an understandable relation to objects. . . . However, the pure concepts of the understanding must not be abstracted from sense perceptions . . . if such intellectual representations depend on our inner activity, whence comes the agreement that they are supposed to have with objects?"; translation from Immanuel Kant, *Correspondence*, ed. and trans. A. Zweig (Cambridge: Cambridge University Press, 1999), 133. Cf. A85/B117 f. For more on interpretations of Kant's letter, see Predrag Cicovacki, "An Aporia of A priori Knowl-

edge: On Carl's and Beck's Interpretation of Kant's Letter to Markus Herz," *Kant-Studien* 82:1991, 349–60.

16. A major difficulty of Langton's approach—as with short arguments in general—is that it almost totally ignores Kant's considerations about space, spatial knowledge, geometry, and the antinomies, which on most traditional accounts are clearly central to the systematic and historical development of Kant's transcendental idealism and the development of his Critical philosophy. See e.g., typical studies by Ted Humphrey, "The Historical and Conceptual Relations Between Kant's Metaphysics of Space and Philosophy of Geometry," *Journal of the History of Philosophy* 11:1973, 485–512; Günter Zöllner, *Theoretische Gegenstandsbeziehung bei Kant* (Berlin: Walter de Gruyter, 1984); and Henry E. Allison, *Kant's Transcendental Idealism* (New Haven: Yale University Press, 1983). It is easy to understand how Langton fell into this approach if she was taken by Kant's very earliest writings, which do not contain anything like the Critical account of space. Harder to understand is her response to this problem at the end of her book, where she seems to shift course and to tie the arguments for ideality only to space and matter rather than to force, and to speak of bodies as if they had a privileged reality. Her final sentence is: "the ideality of space does not, in his [Kant's] own opinion, undermine the reality of what is presented in space, namely bodies, constituted by forces that are mind-independent properties of absolutely independent substances, things in themselves" (218). Unless this passage is meant only to reassert the "empirical reality" of bodies in Kant—which no one would deny—it appears to suggest a notion of "body" independent of space that is very hard to understand and has no place in Kant's final Critical work. By the time of the *Metaphysical Foundations of Natural Science* (1786, 4:503–8, and 520–21; cf. *KTM*, 40), Kant had explicitly completed the argument for the ideality of not only space but also all the matter and the bodies that fill space. This leaves room for the empirical—and only empirical—reality of bodies as such, but it implies that bodies are metaphysically ideal and that there is a nonempirical reality of the nonmaterial thing in itself that underlies matter and bodies (cf. below, note 42). From the *Critique* on, there is no nonequivocal sense in which space is merely ideal while bodies are real.

17. A565/B593, cited in *KH*, 12.

18. Langton (*KH*, 27–30; cf. *KTM*, 82 n. 100, and 269) makes good points here against suggestions (by J. Bennett) that Kant may have confusedly thought that the unknowability of the "inner" was a matter of our not being able to know a "bare" substrate. There is considerable material on this and related topics in *Lectures on Metaphysics/ Immanuel Kant*, eds. and trans. K. Ameriks and S. Naragon (Cambridge: Cambridge University Press, 1997).

19. See the references above in note 12.

20. See the references above in note 12.

21. *KH*, 121, citing 1:21.

22. On the claim that matter is phenomenal, see A277/B333 and the references above in note 12, and especially *KH*, 50–63.

23. *KH*, 101, citing 1:481.

24. *KH*, 102; cf. above, note 12. See L.W. Beck, "Introduction," in *Kant's Latin Writings: Translations, Commentaries and Notes*, ed. L.W. Beck et al. (New York: Peter Lang, 1986), 12.

25. A274/B333, cited at *KH*, 32.

26. *KH*, 107–8; 1:412–15.

27. *KH*, 104–6; 1:410. Cf. my "The Critique of Metaphysics: Kant and Traditional Ontology," in *The Cambridge Companion to Kant*, ed. P. Guyer (Cambridge: Cambridge University Press, 1992), 249–79; as well as Eric Watkins, "Kant's Theory of Physical Influx," *Archiv für Geschichte der Philosophie* 77:1995, 285–324; and his "The Development of Physical Influx in Early 18th Century Germany: Gottsched, Knutzen, and Crusius," *Review of Metaphysics* 49:1995, 295–339; and Alison Laywine, *Kant's Early Metaphysics and the Origins of the Critical Philosophy* (Atascadero, CA: Ridgeview, 1983). A point that needs to be stressed here is that Kant does not accept the Principle of Succession later (although he offers an epistemological analog for it in the empirical domain in his "Refutation of Idealism"), no doubt because for his Critical philosophy it is too metaphysically dogmatic and would cause problems for his own theory of transcendental freedom.

28. Langton properly emphasizes that while Kant's notion of "experience" corresponds to relational properties defining our modern scientific image of the world, it contrasts sharply with standard phenomenalist notions of immediate "experience" as well as with arcane suppositions of certain physicists—including Newton—that supposedly have no empirical explanatory power. See above, note 12.

29. I have argued that these contingent properties could be easily lost sight of by an advocate of the essence theory, since all the contingent properties to which we have access are, according to Kant, filtered through the forms of space and time, and thus (supposedly) inherit their ideality (*KTM*, 271). Thus, the intrinsic but *contingent* characterization of things, including ourselves, is an especially opaque realm—whereas we can easily enough, through the pure categories, at least imagine some nontheoretical but thinkable *necessary* features of things, things in their nonphenomenal essence as pure substance, cause, etc.

30. See *KTM*, 293. It is no accident that Kant does not work out a separate science of psychological knowledge.

31. Langton (*KH*, 97) argues that a substance cannot "essentially have relational properties," since this would undermine Kant's belief in the independence of substances. But note that this is still compatible with allowing that *in fact* things in themselves as such could have relational properties. Moreover, even if Kant's earliest work does seem to disallow the essentiality of relational

properties for things in themselves (in order to safeguard the independence of substances in a system that still makes no use of notions such as transcendental freedom), my own guess is that all that the Critical Kant held, or needed to hold, is that things in themselves cannot consist *solely* of relational properties, essential or not.

32. See *KTM*, 272 f., on “the Relation Argument.”

33. On the distinction of inner sense and apperception, see *KTM*, 241–52.

34. For a very helpful critical analysis of the “myth of the giving,” the thought that we somehow could know the self (or anything) simply through being active, rather than receptive, see Susan Hurley, *Consciousness in Action* (Cambridge: Harvard University Press, 1998), chapter I. 2, “Self-Consciousness, Spontaneity, and the Myth of the Giving.”

35. The passage is discussed in *KH*, 44–6. On affection, cf. *KTM*, 252–5.

36. What can make matters confusing here is the fact that it is true both that all our spatio-temporal data are ideal and that they are “affected” in the sense of being receptive—but this does not mean that they are ideal *because* they are affected by receptivity as such rather than by our form of spatio-temporality.

37. See the works cited above, note 27.

38. Langton notes (*KH*, 137) that “Kant simply gives up on the problem, thus conceived. Abandoning the question of how a substance can be endowed with relations to form a world, Kant says that what is relevant as far as we are concerned is this: if we are to have experience of anything, then we must already be part of a world of causally interacting substances.”

39. “The Critique of Metaphysics,” 263.

40. See *Metaphysik Herder*, 28:52; and *Reflexion* 3581, 17:71.

41. See *KH*, 117, citing 1:415: “A substance never has the power through its own intrinsic properties to determine others different from itself. . . . It has only the power in so far as substances are held together in a nexus through the idea of an infinite being.”

42. Cf. *KH*, 119: “Intrinsic facts don’t constrain relational facts in any way, else God would not be unconstrained,” and 175: “powers don’t depend on intrinsic properties at all.” But why not “at all”? Part of the problem here may be the fact that Langton may not realize that Kant, following Baumgarten and the scholastics, was quite aware of complex causal situations where a variety of different causal factors can “concur” in an effect. (See *Lectures on Metaphysics*, 534–5, for references in the Concordance to many passages on concurrence and related topics.) Such situations leave room for not always tying the notion of causality to strict laws, and so it is no accident that Kant uses the language of concurrence in discussing instances where human freedom, along with other factors such as God, may play a role in something’s coming to be. For a discussion of some typical passages, see my “Kant on Spontaneity: Some New Data,” *Proceedings of the VII International Kant-Kongress 1990*, ed. G. Funke (Berlin: Walter de Gruyter, 1991), 436–46.

43. By sharply severing relations between things in themselves and phenomena (see above, note 42), Langton leaves very mysterious the general relation between the physical (and phenomenal) world and things in themselves: "If physical properties are the properties physics discusses [viz., relational ones 'all the way down'], the conclusion is that there are non-physical intrinsic properties of which we can have no knowledge" (*KH*, 180). She notes, fairly enough, that Kant may still be in a better position than "contemporary orthodoxy" here, since Kant at least (in insisting on a reference to underlying intrinsic properties) is "not purporting to give an *explanation* of a physical phenomenon" (176). This point is well taken; I have used very similar arguments to suggest that Kant's own position in philosophy of mind is at least no worse than Colin McGinn's "transcendental noumenalism," which insists that there is an ultimate natural explanatory ground for all mental events but that it is principle unknowable by us. (See my "Kant and Mind: Mere Immaterialism," *Proceedings of the Eighth International Kant Congress 1995*, ed. H. Robinson [Milwaukee: Marquette University Press, 1995], vol. 1, 675–90.) Nonetheless, the fact that we cannot use features of things in themselves as reductive bases or sufficient "explanations" (let alone physical explanations) for phenomena does not force us into the conclusion that they play no role at all. To use an analogy: on a traditional theistic view (spelled out in a Plantingian manner), God's freedom can put us into an endless variety of worlds, and in each of these His freedom may play the major role in determining the actual laws of nature and the shape that our evil actions take on; but that the actions are evil can still be due to us and our free essence (in a nonsupervening way). It is perhaps no accident that Langton has been most influenced by very early works of Kant that still involve a deterministic perspective and do not express the main ideas motivating Kant's later turn to transcendental idealism.

44. See Bxxvii. A natural way to fill out this idea would be to argue that once we learn that space and time cannot without contradiction be ascribed to things in themselves, then, unless all is "mere illusion," there must be some things characterized in themselves in some non-spatio-temporal way.

Part Two

Kant's Rousseauistic Revolution

Which Freedom?

Ralf Meerbote

I. Introduction

Many a topic would be appropriate for a volume dedicated to Lewis White Beck, but none more so than that of the nature of the freedom of rational actions as envisaged by Kant during his Critical period, together with the companion issue of the compatibilism of such freedom with the causal necessitation of events of nature championed in the Second Analogy of the *Critique of Pure Reason*. Beck himself devoted much thought to these matters throughout his long life, both in his works on Kant and in his own systematic writings such as the early essay, “The Psychophysical as a Pseudo-Problem,”¹ the later “Conscious and Unconscious Motives,”² and the still later Cassirer lectures, *The Actor and the Spectator*.³

Beck also has chronicled Kant’s own, slowly developing and changing conceptions of freedom and its compatibilism, in particular Kant’s youthful quasi-Leibnizian effort (in his *New Elucidation of the First Principles of Metaphysical Cognition*, of the year 1755) and his later, mature conception of these issues in the *Critique of Pure Reason* and in his Critical moral-practical works. Superficially, both of these attempts of Kant’s have a number of elements in common. Both hold that the conditions which make for the freedom of rational actions are in some sense wholly inner or internal to the agent.⁴ Both also hold that such actions and their freedom are indeed in some sense compatible with natural causal necessitation.⁵ Kant always strove to be a compatibilist. But during his Critical period his conception of the internal grounds of action is more refined than it had been in 1755; Kant now portrays free actions as having a first, or absolute, or unconditioned beginning in the agent, whereas such a portrayal cannot be found in 1755. What is more, Kant in the 1780s and thereafter regards free actions as in an important sense indeterminate, whereas *New Elucidation* treats actions of this sort as deterministic in the fashion of all other states or events subject to the Principle of Sufficient Reason.

The just-mentioned three criteria—internality, unconditionedness, and indeterminateness of free actions or their grounds—are definitive of Kant's mature action theory. I will first take up these three criteria and will then discuss a number of interpretations of Kant's theory, in the hope of making some progress in identifying the view of freedom and its compatibilism which our philosopher held during his Critical period.

II. Three Criteria of Freedom

The leading idea behind Kant's Critical account of rational actions is that agents act from their conception of their reasons for action, not just in accordance with reasons. Agents apperceive, or represent their possession of, their reasons. As has been argued in the literature,⁶ Kant, furthermore, takes apperception to be the key to the freedom or spontaneity of actions of this sort. Apperception is necessary for the rationality of an agent's actions, since it is necessary for the agent's deliberating about possible courses of action. On the strength of such deliberations, a Kantian agent then appropriates some reason as his reason for action. Deliberation and appropriation alike involve apperception of reasons, and, for Kant, apperception of this sort makes for freedom of rational actions. Kant hence adopts the following criterion of internality: the proper internal conditions of rational actions that make such actions free are furnished by the agent's appropriation and apperception of his reasons for action. Free actions are actions that are initiated within the agent on the strength of propositional, apperceived states of his that express those reasons.

This model of deliberation, appropriation, and apperception maps the role of practical reason in actions. Hence we can also say that (the use of) practical reason makes for freedom of rational actions. But the practical reason employed by an agent can be either pure or empirical.⁷ Kant identifies cases of acting on the strength of pure practical reason with moral actions. He holds that in cases of this sort, the agent is acting from his conception of the categorical imperative. He labels the freedom of this sort of action "autonomous" (see *Groundwork*, 4:440, 446–7) because he holds the imperative in question to be legislated a priori by pure practical reason. In contrast to such autonomy, actions performed on the strength of empirical practical reason possess a heteronomous freedom or spontaneity. They possess spontaneity because they satisfy the above model of deliberation, appropriation, and apperception. In nonmoral contexts, how-

ever, this spontaneity is heteronomous, because the agent is in possession of interests that are not legislated a priori by reason. (I shall return to autonomy and heteronomy.)

An agent pursuing (either autonomously or heteronomously) free actions is a member of the “intelligible world,” or—as Kant also puts it in his *Groundwork of the Metaphysics of Morals* (4:462)—a member of the world of intelligences. Reason—which makes us free—is viewed here as an ability or a power possessed by members of such a world, and the freedom due to reason that Kant is talking about is a positive freedom, or freedom positively conceived, i.e., a freedom *to* perform actions. (In this connection, he also speaks of transcendental freedom and of a causality of reason or of the will.⁸) But Kant also believes that, as a necessary condition of freedom positively conceived, rational agents must be free in a negative sense as well. Freedom negatively conceived is the independence of rational actions *from* what Kant in the *Critique of Pure Reason* calls pathological necessitation, that is, necessitation due to the senses. More broadly, freedom negatively conceived is the independence of rational actions from any natural causal necessitation, that is, from the causal necessitation that Kant defends in the Second Analogy of the first *Critique*.⁹ Exactly how to understand such independence is a question that goes to the heart of the compatibilism issue and hence warrants examination. Our way with this question has consequences for any understanding of freedom positively conceived—another matter that needs to be looked into—since any interpretation of what it means to call reason a power will be influenced by what we take independence from natural causation to be.

Kant’s second criterion of the freedom of rational actions—unconditionedness—addresses itself to freedom negatively conceived and asserts that, in the order of their coming to be, actions, to be free, have an absolute or first beginning (see A448/B476). Part of Kant’s conception of unconditionedness can be readily appreciated by considering the last paragraph of the *Groundwork* (4:463). In the last sentence of that paragraph, he refers to the unconditioned practical necessity of the moral imperative and claims that we do not comprehend this necessity but that we do comprehend its incomprehensibility. His defense of this claim shows that conditionedness (of a fact or of a necessity) is to be understood as derivability (of that fact or necessity) from some antecedent condition. Unconditionedness is hence the lack of such derivability.

The categorical or moral imperative is a paradigm of an unconditioned condition, since Kant holds this imperative to be underivable from any

other antecedent fact or requirement. His positive explanation of the moral imperative is that it is the content of an originary or creative (hence a priori) act of pure practical reason. He consequently takes this imperative to be necessary. He also makes it clear that the sort of comprehensibility which he is talking about requires conditionedness or derivability. Hence the moral imperative and its necessity are incomprehensible since they are unconditioned. The fact that we, at the same time, comprehend this incomprehensibility is due to the fact that we can derive the unconditioned character of the moral imperative from the originary and a priori nature of the act of pure practical reason in question.

Kant builds on this conception of unconditionedness in his characterization of the performance of rational actions. Any free action, according to him, has an absolute or first beginning, or starts a new series of events or consequences, because he holds such actions to fail to be derivable from any antecedent conditions in accordance with any causal laws of nature (see *Groundwork*, 4:456–9). (As was the case for the categorical imperative a moment ago, any free action is hence also incomprehensible, at least with respect to causal conditions of nature, while, at the same time, we ought to be able to comprehend such incomprehensibility.)

Since free actions are initiated on the strength of appropriation and apperception of reasons, Kant's view of the unconditionedness of free actions turns out to be that it is such appropriation and apperception that are unconditioned since they are not derivable in accordance with any causal laws of nature. The first and the second criterion hence work in tandem. Free actions are initiated by the agent's unconditioned appropriation and apperception of reasons for action. The freedom of moral actions depends on pure practical reason's autonomously legislated categorical imperative. (Recall that agents act from their conception of this imperative.) The unconditionedness of actions of this sort is that of an autonomous first beginning. Heteronomously free actions find their unconditioned beginning in assertoric hypothetical imperatives from the conception of which agents act. Even though any hypothetical imperative may in principle be derivable from some antecedent imperative of this sort and hence may not itself be unconditioned, an agent's apperceptive representation of some particular hypothetical imperative is, according to Kant, in every instance an unconditioned beginning of an action. Heteronomously spontaneous actions do not find their beginning in an a priori originary, legislative act of pure practical reason, but in a reason which, although a posteriori grounded, is in its status as apperceived reason nonetheless unconditioned.¹⁰

Just as appropriation and apperception of reasons, or the causality of reason or of the will, are to express freedom positively conceived, so Kant's conception of free actions as failing to be derivable (in accordance with any causal laws) from any states or events of nature is to capture freedom negatively conceived. Kant's negative conception of freedom, then, moves into the limelight the nature and role of causes and of causal laws (as one would expect in light of the compatibilism issue). Questions concerning the relationship and the difference between causes and an agent's reasons are crucial for Kant since, by the 1780s, he had learned to differentiate between causes and reasons (all lumped together earlier under the Principle of Sufficient Reason). But any question after causes is, for Kant during his Critical period, deeply embedded in his conception of the determinate nature of our cognitive judgments about nature. With Kant's rejection, on behalf of free actions, of causal derivations or explanations comes a wider conception of actions of this sort as falling under indeterminate descriptions. Hence our third criterion of free actions—indeterminateness—now becomes a focus of attention. It will emerge that, for Kant, descriptions of an agent's reason for action are indeterminate, whereas descriptions of natural causes are determinate. The issue of compatibilism hence becomes one of compatibilism between reasons and actions viewed as indeterminate and natural events viewed as determinate.¹¹

For Kant, an object, state, or event is determinate provided it is the object of a determinate judgment. In his fullest sense of "determinate," judgments of this sort possess the following features.¹² (i) They have all of the schematized categories as part of their transcendental-logical form. (ii) They contain an empirical intuition (if they are singular judgments) or else (if they are universal judgments) they have been arrived at by inductive generalization from judgments that contain an empirical intuition. (iii) They contain (one or more) "given" empirical and schematizable concepts which either are predicated of what is presented to or in empirical intuition or else are predicated in the fashion of universal judgments, in accordance with the transcendental-logical form(s) in question. More concretely, a determinate judgment applies "given" empirical concepts to what is perceived or to what stands in a causal connection to what is perceived. Such a judgment treats its objects as spatial and temporal and as extensively and intensively measurable, and it takes some of its objects to be substances which exhibit lawlike causally governed changes of state (events) and which stand in relations of mutual interaction.

In this characterization, I am assuming that the Second Analogy specifies that it is event-types which are causally related so that natural causal necessitation is nomological. This requirement is stricter than one that calls merely for singular causal connectivity. This is a contentious issue. There is a school of thought according to which Kant in the Analogy establishes only that every event has a cause,¹³ and there is a school¹⁴ according to which he in addition establishes that like events have like causes (or, if you prefer, that like causes have like events as their effects). Both Beck¹⁵ and Henry Allison,¹⁶ for example, believe that in the Analogy Kant asserts only singular causation. Both hold that inductive generalization and nomological necessitation get established by Kant on the strength of some third-*Critique* considerations concerning generalization and object- and event-types. The difference between these two schools of thought has in turn been thought to make a difference to the issue of compatibilism. We shall see that the weaker of the two readings encourages Allison, for one, to think that free actions are subject to the Analogy's (purported) singular causation even though not to nomological necessity.¹⁷

But even if Beck's and Allison's reading of the Second Analogy is correct and nomological necessity is not a component of the categorical form of determinate judgments, there still are plenty of regards in which judgments that ascribe free actions to agents fail to exhibit the logical form of their determinate cousins. For one, Kant denies that purposes, hence intentions and reasons for action, can be empirically intuited or perceived. In the Preface to his *Metaphysical Foundations of Natural Science* (4:471), he expresses doubts concerning the possibility both of self-observation (of one's mental states) and of perception of the mental states of others.¹⁸ (Apperception of reasons for action could hence not be perception.) In the so-called "First Introduction" to the *Critique of Judgment*, as well as in the main body of that work, he holds that the purposiveness of natural purposes is not perceived or experienced. It is but a small step to argue that therefore agents' purposes and reasons for action likewise fail to be perceivable.¹⁹ Judgments that ascribe free actions to agents do not essentially contain, or inductively rest on, empirical intuitions, and the concepts that they employ are not schematizable. In the aforementioned Preface, Kant, furthermore, denies that empirical psychology can be properly mathematized.²⁰ He means by this that the principle of the Axioms of Intuition is not part of the logical form of empirical psychological judgments. *A fortiori*, the same would hold of judgments that ascribe apperceived reasons and actions to agents. In addition, the

principle of the Anticipations of Perception does not apply to apperceptions, since in propositional states there is nothing to be measured intensively.²¹

Leaving mathematizability aside, at least two more categories clearly fall by the wayside for judgments that ascribe free actions to agents. In several versions of his *Refutation of Material Idealism*,²² Kant argues that the “permanent” cannot be met with in any of our mental states, and in the Third Analogy he treats substances that are in mutual interaction as spatial. Hence both the principle of the First and that of the Third Analogy are withheld from empirical psychology. (The former is withheld because the “permanent” in question is substance; and the latter, because Kant conceives of the mental as temporal without conceiving of it as spatial.) But once again, as goes empirical psychology, so goes our ascription of apperception.²³ Judgments that ascribe free actions to agents do not count the principles of the First and of the Third Analogy among their logical form.

This leaves the Second Analogy. Many a commentator would argue that if the First Analogy falls, so does the Second;²⁴ but Allison, as intimated above, maintains that the latter (as interpreted by him) does apply to free actions. If, on the other hand, the Second Analogy is interpreted as requiring nomological causal necessity, then it will readily seem not to be part of the logical form of free-action judgments, if only because lack of proper mathematizability has negative consequences for any view that champions lawfulness.²⁵ The fact that free-action judgments fail to obey any principle of nomological causal necessitation is particularly important since it signals the absence of causal laws in our explanations of free actions. The unconditionedness of actions of this sort hence arises as an integral part of Kant’s theory of determinateness or indeterminateness.

In most or all respects, then, judgments that ascribe free actions to agents fail to be determinate judgments. It is clear why this is so: they fail to be determinate because in his theory of determinate judgments, Kant is analyzing our empirical knowledge of nature, but when an action is considered to be free, it is not considered to be among the proper objects of such knowledge.

This brief sketch of Kant’s conception of determinateness and of his reasons for considering ascriptions of apperception not to take the form of determinate judgments must suffice for showing that he believes an agent’s appropriation and apperception of reasons for actions to be indeterminate. Free actions and natural events, then, differ with regard to our third criterion.²⁶

They also differ with regard to the first and the second criterion. As far as the latter—unconditionedness—is concerned, the reason for this should be obvious. The causes of which Kant speaks in the Second Analogy give rise to a regress *in indefinitum* (although not to a completed infinite regress) of antecedent causes. This is so because all causes here are themselves events. There is according to the Second Analogy no first or unconditioned beginning in the chain or chains of natural events. (The causal necessity of every member of such a chain can be comprehended since it is always a conditional necessity, calling for, and being assigned, an antecedent member of the chain as its antecedent condition.) Natural events find no unconditioned beginning in any member of the chains to which they belong, whereas free actions are not viewed as consequences subject to any causal laws. The reason why free actions and natural events differ with regard to the remaining criterion—internality—is also not difficult to fathom. Natural events occurring in a substance typically are not viewed as initiated in the substance in question (and, needless to say, they are typically not initiated by propositional states). Rather, they are viewed as being ultimately initiated by a condition external to that substance.

An excellent way to appreciate externality is to reflect (again) on the Third Analogy. There, Kant holds that natural substances stand in relation of mutual causal interaction. But such relations obviously make for external natural events: events in each of the substances involved induce events in the other(s).²⁷ In addition, the concrete interpretation which Kant assigns to his three Analogies in his *Metaphysical Foundations* makes it clear that motion is a fundamental property of material substances and that moving substances require the application of external forces. Hence, when substances are viewed first as natural and then as material, externally initiated events get ascribed to them. Hence at least some, if not all, natural events are externally conditioned.

An often-mentioned discussion of Kant's in the *Critique of Practical Reason* (5:93–8) sheds further light on the externality of natural events.²⁸ There, Kant distinguishes determinate descriptions of mental states in general from proper ascriptions of transcendental freedom (i.e., of freedom positively conceived or the causality of the will) and argues that under the former the agent is being treated as falling under natural causal necessitation (even though it is his mental life that is being considered). But by being treated in this fashion, the agent is not being treated as free. Kant's view, in this passage, that empirical psychology affords us determinate descriptions is, as we saw, by no means consistent with what he has to

say elsewhere about that discipline and about the determinability of the mental. Nonetheless, his present discussion clearly indicates that any ascription of transcendental freedom calls for indeterminate descriptions and that, if we assume that mental life in general (in contrast to what gets described as apperception of reasons for action) gets described determinately, it would, under those descriptions, be viewed as externally conditioned. This more than suggests that Kant is thinking of all determinately described events, hence of all natural events, as described as externally conditioned.²⁹

Free actions satisfy our three criteria whereas natural events do not. In particular, free actions do not fall under the Second Analogy (at least in its strong reading) since they are indeterminately initiated by unconditioned apperceptions of reasons for action. And yet, as I mentioned above, Kant speaks of a causality of reason or of the will, intending thereby to be characterizing the efficacy of reason or will, or freedom positively conceived. Such causality is to express the fact that apperceived reasons move the agent to action.

Kant works out his causal principle of the Second Analogy in great detail;³⁰ but the same can hardly be said of any causality of reason, and there is no consensus in the literature as to how a causality of this sort is to be understood. The puzzle of freedom positively conceived is compounded by that of freedom negatively conceived: as we have seen in some detail, the kind of efficacy that Kant is talking about, or its exercise, is to be independent of natural causation. Kant seems to be talking about uncaused causes, not causes in the spirit of the Second Analogy.

As a matter of fact, many a commentator has felt moved to think of Kant's conception of free action as "contracausal," to borrow a term of Allison's, to the point of saddling Kant with some sort of incompatibilism between freedom and natural causation. One may indeed wonder what negative freedom is to be if not some sort of "contracausality." If Kant's overall metaphysics is to be shown to be consistent, both the negative and the positive conceptions of freedom require further explanation and defense, and this need must be met in full view of the Second Analogy (as well as the other requirements that make for the determinateness of our cognitive judgments about nature). Kant needs to explain how there can be in the world both natural events, externally related, always conditioned, and determinate, and free actions, properly internal, unconditioned, and indeterminate. He confronts this challenge repeatedly during his Critical period, in particular in the third Antinomy of the *Critique of Pure Reason* and in a number of discussions of freedom and reason in his moral-practical writings.³¹ (In all of these contexts, he explicitly addresses the issue

of compatibilism.) But what room can there be in his overall metaphysics for the independence and the efficacy of the rational will, given, for example, the standard view that free actions are temporal events and that all temporal events fall within the scope of the Second Analogy?

III. Nontemporal Agency

I will now take up several views that attempt to interpret major features of Kant's action theory, with particular emphasis on positive and negative freedom, hence on the issue of compatibilism between free agency and natural causal necessitation. The first view to be discussed rejects the just-mentioned standard view of such agency and treats rational actions as timeless and hence as not temporal events at all. The basic idea would be that free actions are noumenal, and that what is noumenal is nontemporal and hence not part of nature. The Second Analogy, in contrast, deals only with natural or phenomenal, hence temporal, events and would therefore not apply to free actions. Natural events are causally determinable but free actions would not be. According to the timeless-agency view, nontemporality safeguards the internal character of free agency, and the unconditionedness of free actions or their independence from natural causal necessitation is guaranteed by the inapplicability of the Analogy to actions of this sort. All of this would have to be understood in the spirit of a two-worlds interpretation of Kant's ontology, so that free actions would not be numerically identified with natural events.

The desired compatibilism between free actions and natural events, according to this view, resides in the (purported) fact that free actions obtain in a world numerically distinct from that in which natural events occur. Kant's characterization, in the *Groundwork*, of rational beings as members of an "intelligible world" may seem to provide support for such a reading; and it may seem that, as members of such a world, rational agents wield a noumenal or intelligible, hence nontemporal, efficacy, to be distinguished from the causal efficacy that is at work in the phenomenal world. More broadly, the timeless-agency view seems to derive support from the fact that Kant, both in his practical and in his theoretical philosophy, invokes a distinction between what is noumenal and what is phenomenal, where what is noumenal gets identified with things as they are in themselves. He puts free actions on the side of things as they are in themselves. Since the latter are nontemporal according to him, and since

a two-worlds interpretation of his distinction between things as they are in themselves and things as they appear (phenomena) has seemed to many commentators to be the correct way of reading him, the timeless-agency view seems to be securely anchored in Kant's general ontology.

Allen Wood has presented such a view of Kant's theory of action.³² According to Wood, a free, rational agent timelessly chooses his noumenal or intelligible character. This choice in turn makes for the agent's empirical and temporal character and, with this, for the series of natural events that is causally necessitated by the latter character. The timeless choice in question is efficacious in bringing about a series of natural events since a different choice would or could have made for a different series. Nontemporal agency, according to Wood, is at least a necessary condition for some of what occurs in the phenomenal world.

But Wood's interpretation faces perplexing and serious difficulties. The dependence of any nontemporal-agency view on a two-worlds interpretation of Kant's ontology is a double-edged sword since that interpretation is itself highly debatable. But more immediate objections can be brought against Wood's view. As Hud Hudson has pointed out, there is no evidence that Kant countenances timeless choices.³³ What is more, the notion of such choice is intrinsically puzzling, since it is the notion of something coming about in an agent but timelessly so. We normally view any coming-about as temporal, and we normally take any reason from which an agent acts to obtain temporally before or during his (temporal) performance of the action. Even though our conception of free actions is indeterminate, such actions are not thereby regarded as nontemporal. This is what Allison, for example, has in mind when he holds that free actions are not timeless even though independent of the conditions of time.³⁴ Another objection that can be brought against Wood's view is that it assumes that the agent's empirical character is straightforwardly causally integrated with the rest of nature, i.e., that an agent's temporal psychological states are describable in terms of determinate psychological descriptions. But we have seen that this may not be Kant's considered view.

However, there is even worse to come for the timeless-agency view. The very point of speaking of an agency here is that through his timeless choice of his intelligible character, the agent nontemporally *acts on* portions of the phenomenal world. As noted a moment ago, such agency is to be efficacious and, at the minimum, necessary for some series of natural events. (The view in question does not claim some mere pre-established harmony between the two worlds.) But, as Hudson has also pointed out,³⁵

this means that Wood's reading denies the sufficiency of natural causes for at least some natural events. This, however, conflicts with the Second Analogy. According to the Critical Kant, it is a necessary condition of the empirical knowability of temporal, natural events that whatever causes them be itself temporal. Since at least some of the effects of the agency in question are temporal and empirically knowable, this agency must then itself be temporal. According to Kant, only temporal events can be causes of temporal events. All temporal events find their sufficient causes among nothing but temporal events. Since Wood believes there to be a nontemporal agency at work, he denies this sufficiency. His view hence jeopardizes the empirical knowability of at least portions of the phenomenal world. (The argument just sketched also works for "natural," with the same negative consequences. Only natural events can be causes of natural events. Natural events find their sufficient causes among nothing but natural events. But according to Wood, there is a second, nonnatural efficacy which operates at, even though not in, the phenomenal world.³⁶)

Wood's view is the temporal analogue of Euler's virtual-presence view of the location of the soul. Wood has Kant arguing that there is an agency that acts at a time without acting in time. But since Kant, during his Critical period, rejects Euler's view on the grounds that it provides no proper metaphysical solution to the puzzle concerning the seat of the soul,³⁷ it is hardly credible that he would accept any temporal-virtual-presence view of agency. The epistemic requirement that the Second Analogy introduces and, more generally, Kant's overall Critical metaphysics rule out Wood's view.

It will be objected that this last claim (concerning Kant's overall metaphysics) cannot be right, if only because of Kant's repeated identification (see above) of what is free and noumenal with nontemporal things as they are in themselves. Since Kant speaks of a noumenal or intelligible character or causality of the will, does this identification not show at least part of Wood's reading to be correct (with Kant risking an inconsistency between the Second Analogy and the conditions of free agency)? I would urge three replies to this objection. One is that things as they are in themselves are not efficacious for Kant.³⁸ If this is right, then the claimed identification will not support Wood's reading (unless Kant is again inconsistent). The second reply is that Kant's conception of "noumenal" is a complicated business indeed, and may not support any identification of the ontological status of the efficacy of the rational will with the status of things as they are in themselves.

For example, consider Kant's negative or limiting employment of the concept of a noumenon (A251–2, B307–8). Such a use is, as Kant puts it,

to “curb the pretensions of sensibility” (B311). To be noumenal is to be other than phenomenal, i.e., it is to be supersensible, but this characterization declares to be noumenal more than things as they are in themselves. For example, we view free actions as noumenal when we do not view them as phenomenal, but it would have to be argued, not just assumed, that such actions are hereby viewed as nontemporal. “Noumenal” in its broadest meaning covers any object of reason, free actions as well as things as they are in themselves, but only some of what is noumenal is nontemporal. Let us think of free actions or of the exercise of free agency and of intelligence or rationality as noumenal or intelligible properties of some entities. This characterization does not force us to treat such actions, exercise, or properties as automatically nontemporal. Similarly, the “supersensible substrate” that Kant invokes in the *Critique of Judgment* (see 5:412, 422, 429) is not necessarily nontemporal, and the real purposiveness that he discusses in that work is assuredly noumenal, i.e., supersensible, and equally assuredly not nontemporal (even though not properly schematizable).³⁹ In short, Kant uses the term “noumenal” very broadly indeed, too broadly perhaps for his own good.

My third reply (which perhaps is no more than a restatement of the second) is that Kant does often put free actions on the side of things as they are in themselves, in order to “save” freedom, but he does this, not in order to hold that we act nontemporally, but rather in order to provide some conceptual room, as it were, for a point of view from which we do not describe ourselves as determined by any causal laws of nature, including any laws of sensuous inclination. (Unconditionedness, again.) If any of this is right, then Kant’s talk of things as they are in themselves and of a noumenal character of free actions will not support Wood’s nontemporal-agency view. Wood may want to urge some revisions of the first *Critique* (and of the third) in order to support his reading and still have Kant adopt some (revised) overall consistent view. In urging revisions, he would be in good company, as we shall see. But, as matters stand, it is difficult to see how his interpretations of Kant’s view of free agency can be consistent with the main tenets of Kant’s metaphysics.

IV. Three Temporal-Agency Views

In light of the foregoing, we had better hold rational actions to be temporal. Free actions, then, are temporal events initiated by a noumenal, yet

temporal agency. Hence the nature of the efficacy of an agency of this sort needs to be accounted for. Hence equally in need of explanation is the independence of free, temporal actions from natural causal necessitation. The ultimate test of temporal-agency views will be their success in dealing with Kant's compatibilism issue.

I will now take up three attempts which treat free actions as temporal and which address themselves to these matters. First is Beck's interpretation of Kant;⁴⁰ the second, Allison's,⁴¹ and the third, an interpretation that follows in the footsteps of Donald Davidson's reading of Kant.⁴²

Beck is an open revisionist. He defends the view that practical reason is efficacious and that free actions are unconditioned, and then argues that unless Kant or we drastically revise our understanding of the content and role of the Second Analogy, no progress can possibly be made in reconciling freedom with that Analogy. He ascribes to Kant a two-descriptions view of sorts rather than a two-worlds view,⁴³ and he believes that the temporal character of free agency has such agency and its consequences compete with natural causation. His fundamental objection to a variety of compatibilist rescue efforts boils down to the belief that *no* event that is taken to be subject to causal necessitation can also be taken to be a free action (not just that *some* events taken in the first way cannot also be taken in the second). He believes that noumenal freedom is incompatible with phenomenal causal necessitation.

An incompatibilist in his own thinking about these matters, Beck, however, knows fully well that Kant is groping for some sort of compatibilist reconciliation between freedom and the Second Analogy. Beck's prescription or recommendation—both to Kant and to us—is that we demote the principle of that Analogy from the status of an *a priori* constitutive principle (essential for all of the objects of nature and for all of our empirical knowledge of them) to that of a principle merely regulative of the methods that we practice as far as possible and of the types of explanation that we settle for as far as possible, when we attempt to acquire as much knowledge as we can. Beck hence recommends we countenance there being uncaused events, or there being events not caused by natural causes in order thereby to allow for the possibility (and actuality) of free actions. (This, then, is his interpretation of the independence of free actions from natural causation.)

His argument winds its way through his interpretation of Kant's Antinomy of Teleological Judgment in the third *Critique*. He subscribes to the view that Kant's resolution of *that* conflict consists of treating the prin-

ciple of the teleological assessment of some objects of nature and that of the mechanistic determination of all such objects as being equally regulative, and he then reads this result back into the Second Analogy of the first *Critique*. Allison and others have found fault with this resolution of the third *Critique*'s Antinomy and with Beck's translating his resolution into a revision of the status of the Second Analogy.⁴⁴ I only want to second Hudson's observation that Beck's revision of the Second Analogy asks us to pay a very high price indeed for a reconciliation of freedom with natural causation.⁴⁵ Empirical knowledge *à la Kant* would no longer be possible since such knowledge requires the Analogy to be universal in scope and to be a constitutive principle. According to Beck's interpretation, some events are not natural and, one might think, some natural events are not naturally caused but rather are the effects of a nonnatural, although temporal, agency. If he were to accept the latter of these two claims, his view would be subject to one of the arguments advanced against Wood above: since at least some of the consequences of free agency are natural and empirically knowable, this agency must itself be natural. The only way for Beck to escape this argument would be for him openly to count all of the consequences of free actions as nonnatural as well, removing more than just those actions themselves from the scope of empirical knowledge. But as a result of all of this, it would be difficult to see what remains of the Second Analogy. (Leaving nontemporality and free actions themselves aside, the only difference between Wood and Beck would be that the latter openly declares nonnatural and unknowable what the former takes to be natural but is driven to put into epistemic peril.)

Beck's way is also less than what one might hope for because, according to him, we can be free only in the gaps, as it were, in or between nature's causal chains. At the same time, his move illustrates how enormous the tension is or is perceived to be between freedom and universal causal necessitation since this tension moved Beck, for one, to pay the price in question as well as to settle for being free only when nature suspends her operations.

More recently, Allison has written extensively on freedom, causation, and compatibilism.⁴⁶ I have commented on his views on other occasions and will sketch here only a few essentials of his position, as I see them.⁴⁷ In contrast to Beck, Allison has compatibilist instincts, but the view that he eventually arrives at may not be any more satisfactory than Beck's. He very much sounds like a defender of a compatibilist interpretation of Kant in his declaration that (according to our philosopher) "we are both caus-

ally conditioned and free . . . albeit *qua* considered from two different points of view.”⁴⁸ No lover of two-worlds interpretations, Allison defends and utilizes a forthright two-descriptions approach in his argument against Beck, an approach that is designed to show (as he implies in that argument) that the same event can be viewed both as causally conditioned and as free as long as it is viewed from two different perspectives. (He chides Beck for believing that an event can be taken to be either caused or free but not both.)

Allison's analysis of free actions turns on the appropriation and apperception of reasons and on the unconditioned character of such appropriation and apperception, as explained above. (He also refers to our conception of free actions as “indeterministic.” This is something I shall return to.) Our earlier model of the exercise of practical reason hence serves as our characterization of free actions from one of Allison's two points of view. (He identifies this point of view with that of the agent's noumenal or intelligible character.) At the same time, from the other of the two points of view (that of the agent's empirical character), we characterize actions of this sort as being causally governed in the phenomenal world. Hence, in his argument against Beck, Allison affirms that “the intentional actions of rational agents” are subject to the principle of the Second Analogy. As already mentioned, he interprets that Analogy to assert singular causation, that is, he asserts the “every-event-has-a-cause” reading, not the nomological-necessity interpretation. (He holds that physical-object descriptions do express nomological causal necessity, but that such necessity goes beyond what the Analogy requires.) In short, free actions are subject to singular causation. This is made true because we are applying two different “models”—that of the agent's intelligible character and that of his empirical character—to the same action-events. So Allison's compatibilist reading consists of holding that free actions are temporal events and that some events both are subject to singular causation and exhibit freedom “albeit *qua* considered from two different points of view.”

This, however, is not the whole story. When we consider an action to be free, we consider it to have an unconditional beginning in the agent's appropriation and apperception of some reason, i.e., to have a beginning that is not derivable from any antecedent condition in accordance with any causal laws of nature. But since Allison does not believe that the Second Analogy enforces nomological causal necessity and since he believes that, according to Kant, whatever is mental is at best subject only to singular causation (or, perhaps, because he does not take seriously

Kant's present denial of *lawfulness*), this view of Kant's reduces for Allison to a rejection of the only sort of causation that he thinks is in contention, i.e., to a rejection of singular causal connectivity. Hence he interprets Kant's lack of derivability to mean failure to obey the Second Analogy or failure to exhibit causation *überhaupt* as he (Allison) understands that Analogy and that causation. He hence holds that the beginning in question, or an agent's apperceived reason for his action, is uncaused since not subject to singular causation. But then, of course, if the agent's reason does not fall under the Second Analogy, then neither does the free action itself, since that Analogy requires that if an event is caused, it is caused by an event which is itself caused. Hence free actions are not subject to the Second Analogy's singular causation.

Allison consequently dubs our conception of freedom "indeterministic" and "contracausal," meaning by this that free actions are uncaused. (This, then, is his explanation of how such actions are independent of natural causation.) Only the intelligible-character model allows us to ascribe freedom to some events, and this model does so by denying that the events in question are caused. Since this way of considering those events has them not fall under Allison's Second Analogy, his overall position is that there are temporal events (free actions) that both do and do not fall under that Analogy, "albeit from two different points of view." We conceive of free actions as causally conditioned and we conceive of them as not causally conditioned since free, from two different points of view.

Allison takes these two ways of considering the events in question to be compatible, on the grounds that the freedom asserted by the intelligible-character model is "empirically vacuous," i.e., does not "enter into an empirical [causal] . . . explanation of . . . action."⁴⁹ Hence he portrays himself as a compatibilist in his argument against Beck. (He also refers to his "indeterministic" and "contracausal" conception of freedom as "incompatibilist,"⁵⁰ apparently meaning by this that this conception cannot be accounted for just in terms of the empirical-character model, and he makes approving mention of Wood's description of Kant's project as one which aims at the "compatibility of compatibilism and incompatibilism."⁵¹ Allison's version of this description is that there are two different points of view of some events, both as caused and as free, which are compatible and equally legitimate, even though one of them invokes an "incompatibilist" conception of freedom.)

Allison's resolution is quite as puzzling as his two-conflicting-views-of-things solution to Kant's problem of things as they are in themselves

and as they appear.⁵² It would seem that the empirical-character model and the intelligible-character model conflict since one asserts what the other denies, viz., the character of all events as caused. Allison seems to be implicitly revising the Second Analogy by narrowing its scope (like Beck), since he is countenancing uncaused nonnatural events. (He is withholding events of this sort from the empirical-character model but makes them the centerpiece of his other model.) From the point of view of the agent's intelligible character, the Analogy is denied of nonnatural temporal events. But this makes the relation between the two Allisonian points of view extremely puzzling, since actions and consequences explained one way in terms of the causally governed empirical character of the agent are also explained another way in terms of uncaused acts of appropriation and apperception, hence in terms of a nonnatural efficacy. This, however, renders Allison's resolution problematic for a by now familiar reason, namely, the argument already brought into play against both Wood and Beck. It is a necessary condition of the empirical knowability of natural events that whatever causes them be itself natural. Since at least some of the effects of free agency are natural and empirically knowable, this agency is itself natural. Only natural events can be causes of natural events. This is an argument that can be made from the point of view of the agent's intelligible character. But according to Allison, from that point of view, the agency in question and its efficacy are considered to be nonnatural. Hence, in light of the argument just given, Allison's conception of freedom jeopardizes our empirical knowledge of nature in the same sort of way as Wood's and Beck's view before him. When Allison claims that his two points of view are compatible because the freedom asserted by the intelligible-character model is "empirically vacuous" since it does not "enter into an empirical [causal] . . . explanations of . . . action" (see above), his defense misses the fact that that model questions the very possibility of the empirical explanation of events, including actions.

It will not do for Allison merely to reply that free agency according to him is both natural and nonnatural. It will not do for him to claim that, since he (against Beck) accepts the Second Analogy as constitutive, there are two different types of "causes" at work on natural events, one natural and the other free, hence uncaused. Even granting free causes, nothing would explain how these two types bring about the same natural events (if they do), since pre-established harmony is not an option. But the real crux of the problem lies in the very fact that Allison, in light of his two ways of considering free actions, is, like Wood and Beck, postulating two

different efficacies. But then, again like his two predecessors, he is—by means of his “contracausal” conception of freedom—impugning the sufficiency of natural causes for at least some natural events. The “conflict” between Allison’s two points of view looms large indeed.⁵³ Once again, it seems that, Allison’s argument against Beck notwithstanding, the present interpretation, like Wood’s, will need to consider forthrightly revising the Second Analogy (and live with the consequences), something Allison is not likely to want to do.

Allison’s solution, perhaps like the next and final view to be discussed, is a case of having one’s cake and eating it, too. Beck holds that no event can be both free and causally conditioned. Allison replies that some events can be considered to be both from two different points of view which, however, turn out to conflict. I finally turn to another interpretation which, I believe, captures Kant’s view more nearly than does Wood’s, Beck’s, or Allison’s. It is an interpretation which, like Allison’s, utilizes a two-description approach to Kant in order to furnish a compatibilist reading. Stimulated by his reading of Kant’s third Antinomy of the first *Critique*, Donald Davidson a number of years ago developed anomalous monism as an account of the mental,⁵⁴ and anomalous monism as a reading of Kant has since then been developed further by others.⁵⁵

This reading buys into the view that appropriation and apperception of reasons for actions make the actions in question free or spontaneous (the first of our three criteria of freedom). Anomalous monism nicely accounts for unconditionedness (the second criterion), as we shall see. More generally, what makes this interpretation of Kant’s view noteworthy is its consequent two-descriptions interpretation of his third criterion, the indeterminateness of free actions. According to the anomalous monist interpretation, determinate judgments (as I described above, including the claim that the Second Analogy enforces nomological causal necessity) and indeterminate judgments ascribing free actions to agents are applied to the same entities. The view is a monistic one in that it asserts token-identity between all instances falling under free-action descriptions in indeterminate judgments, and instances falling under descriptions in determinate judgments. As far as action-events are concerned, token-identity is to capture Kant’s claim that one and the same event is being regarded in two different ways (see, e.g., *Groundwork*, 4:456). The view makes for anomalousness of the relations among free-action types as well as of the relations between them and determinate types in that it takes all free-action types (whatever they may be) to be indeterminate, in particu-

lar, not to lend themselves to nomological causal necessities. (Neither can such types be analyzed in terms of determinate types.) Hence there are, for example, no causal laws that bridge the gap between the two families of types, and indeterminate descriptions of the relation between apperceptions and actions, for example, lay claim to no causal lawfulness. Determinate and indeterminate judgments are irreducibly different from one another, both in form and in content. Anomalous monism puts itself forward as an elaboration of Kant's account of the difference between his two types of judgment and hence takes anomalousness to be safeguarding his third criterion of freedom, indeterminateness.⁵⁶

According to this interpretation, free actions are natural events and agents' reasons for actions are natural causes. (At the same time, it is not claimed that all natural events are free actions, only that all free actions are natural events. Hence only some causes will be reasons for actions.) Anomalous monism does not demote the Second Analogy to a merely regulative principle and denies that there are nonnatural or uncaused events (or nonnatural or uncaused causes). (Timeless agency is also rejected.) Our empirical knowledge of nature is hence not in peril. The monist's avoidance of views to the contrary turns in part on his straightforward understanding of Kant's second criterion, unconditionedness. According to monism, since indeterminate judgments that contain free-action descriptions do not obey the Second Analogy, there stands nothing in the way of having judgments of this sort specify a first beginning of a free action, even though determinate judgments containing causal descriptions always specify or presuppose antecedent causes of the action in question. The anomalousness of free-action types sees to it that there is no way of deriving a free action (described as an action of this sort), or an agent's reason for action (described as a reason), from any causal antecedent (described as a natural cause) in accordance with any causal laws of nature. Anomalous monism treats all reasons and actions under free-action descriptions as unconditioned *vis-à-vis* any natural events under causal descriptions, since there is no way of deriving the former (as described) from the latter (again as described) in accordance with any causal law. Monism, in short, takes seriously Kant's reference to causal laws of nature in his conception of unconditionedness and denies that there are any laws which would make for the derivability in question. The anomalous monist interpretation of Kant's view explains unconditionedness in a forthright and natural fashion (and it does so without postulating uncaused causes). Our second criterion of free actions is hence accounted for as well.

But unconditionedness makes only for freedom negatively conceived and one may still wonder whether anomalous monism can account for freedom positively conceived while still avoiding nonnatural causes or events. Token-identity, however, is to do the rest of the job. In particular, the monist deals with the remaining (first) criterion of free actions—appropriation and apperception of reasons—by claiming token-identity between appropriation and apperception, on the one hand, and what is phenomenal or natural, on the other, and by claiming that this does not obviate the proper internal character of what moves the agent to action. This is the most controversial aspect of this interpretation since monism treats the efficacy of the will as itself token-identical to natural causal efficacy. (I shall come back to this token-identification of the will.)

Anomalous monism as a reading of Kant avoids, then, the difficulties encountered by Wood's, Beck's, and Allison's views. Token-identity of free actions with natural events, including token-identity for appropriation and apperception, together with the anomalousness of free-action types, allows the monist to deny that free actions are timeless or nonnatural or uncaused or (*qua* free) conditioned and that natural causes are insufficient for any events. In particular, this reading of Kant avoids Allison's difficulties. Because of token-identity, the monist denies that free actions are contracausal. The monist, as a matter of fact, holds the strong thesis that, again because of this identity, events that are indeterminately described as free actions are by means of causal descriptions described as nomologically necessitated. Allison to the contrary, to be subsuming an event under a free-action description is not thereby to be calling the action uncaused. To be subsuming an event under a free-action description is merely on that occasion to be withholding causal descriptions. In an indeterminate judgment that some action is based on some reason, no causal claim is either made or denied (even though reasons are causes). The anomalous monist interpretation of Kant avoids Allison's two conflicting ways of considering events. (Because of the monist's token-identification of the rational will, this view also avoids Allison's—and Wood's and Beck's—nonnatural efficacy. See below.)

I have sketched the monist interpretation's way with Kant's three criteria and have highlighted the differences between it and Wood's, Beck's, and Allison's readings. Some of these differences clearly are to monism's advantage. But several fundamental challenges to the monist reading come to mind immediately, and I will conclude by discussing two of them.⁵⁷

The first issue concerns the matter of fit between indeterminate and determinate conditions of action. How can we explain that and how some

particular free action is token-identical to some particular natural event? Can we explain this or is this left mysterious? According to anomalous monism, there is no derivation of any free-action description from any causal descriptions (in accordance with laws). The view treats nothing when regarded as a free action as conditioned *vis-à-vis* anything regarded as causally necessitated, as explained above. Monism sharply distinguishes between the determinate and knowable and the indeterminate and unknowable aspects of those events which are free actions.⁵⁸ As a result, the monist treats the fit in question as itself indeterminate, hence unknowable.

Free actions as unconditioned are incomprehensible according to the monist reading, just as they are for Kant. They are incomprehensible because of the underivability of free-action descriptions from any descriptions in accordance with any causal laws. What is more, the monist reading also makes this incomprehensibility comprehensible, as required by Kant. It does so by pointing to the indeterminateness of free-action types or to the anomalousness of the relations among free-action types and between types of this sort and causal types. In the same vein, the fit or compossibility between some event as a natural event and the same event as a free action—when there is a fit—is itself incomprehensible in Kant's sense because it cannot be derived in accordance with any causal law of nature. And we again comprehend this incomprehensibility in the same manner as before. (When there is no fit, this, too, cannot be derived in the required manner.) At the same time, we can grasp how there can be fits: the determinateness, external character, and conditional necessity of causes do not in principle compete with the indeterminateness, internality, and unconditionedness of actions or their reasons.

By treating both free actions and their token-identity to natural events as incomprehensible (in Kant's sense), anomalous monism treats what we view as free actions as "theoretically" incomprehensible, a disclaimer that matches Kant's.⁵⁹ Anomalous monism in these regards faithfully maps Kant's own pronouncements. If there is a mystery here, it is his mystery. This, then, counts in favor of anomalous monism as a reading of Kant. What seemed a challenge to this reading turned out to be support.

The second issue is the monist's token-identification of an agent's appropriation and apperception of his reasons (the noumenal, temporal causality of the will) with natural causal efficacy. The negative freedom, or independence from natural causal necessitation, which Kant insists on in behalf of free actions is interpreted by anomalous monism to consist of the anomalous relations between the types in question, not Wood's inap-

plicability of the Second Analogy or Beck's absence of natural causation or Allison's contra-causality or conflicting ways of considering events. Kant's positive freedom, or the efficacy of the rational will, monism takes to be identical in every instance to the efficacy of natural causation, without reducing the conditions of free action to such causation in the fashion of a type reduction or type identification and without invoking any causal laws. In this limited manner, anomalous monism treats the will's efficacy as natural, hence as not mysterious. By not denying the sufficiency of natural causes for all events, this interpretation of Kant avoids any non-natural-efficacy or two-efficacies hypotheses. But it will be objected that this conception of negative freedom is not strong enough and that, as a consequence, the monist's conception of positive freedom is not strong enough either. Allison, for example, has charged that the anomalous monist has not shown freedom to be genuine and has simply failed to capture Kant's view.⁶⁰ Allison objects to token-identity of the rational will as much as he objects to type reductions or the causal lawfulness of free actions.

It may be helpful to compare the anomalous monist reading of compatibilism with Kant's early quasi-Leibnizian effort, mentioned near the beginning of this paper.⁶¹ In that effort, Kant treated free agency as itself deterministic, as did Leibniz himself.⁶² That master rationalist believed that freedom of the will is nothing but a deterministic mental condition of action. Kant, at that early time, stresses that free actions have an essential internal component, but he is also committed to locating other, antecedent conditions of such actions outside the agent, with the internal being no more than a part of an unbroken law-governed chain of necessitating conditions.⁶³ Kant in 1755 is committed neither to the indeterminateness of free actions nor to their unconditioned beginning, and he makes only a halfhearted effort at establishing their properly internal character. This, then, is his early conception of freedom that the Critical Kant contemptuously dismisses as no more than the illusory "freedom of the turnspit."⁶⁴ Without whittling away at the Second Analogy, hence without risking internal inconsistency of Kant's overall metaphysics, anomalous monism portrays Kant as significantly removed from that earlier position.

This point can also be made in a slightly different way. The monist reading is to be preferred on the grounds that it posits only one efficacy, namely, that of natural causation, rather than a second, nonnatural efficacy. This is important because of the difficulties that two-efficacies views encounter and because, as I mentioned, we find in Kant only one reason-

ably clearly worked out conception of causation, namely, a conception of natural causation. Any nonnatural efficacy of the will certainly seems mysterious in comparison to that of natural causation. The anomalous monist reading of Kant attempts to do justice to Kant's "theoretical" views on efficacy, again without having freedom collapse into that of the turnspit.

But nothing I am saying here will convince the opposition. The monistic reading of Kant will be rejected by many because the token-identity of the will with natural causation will simply be found wanting. There is a deep-seated belief in some quarters that freedom must be contranatural, to use Allison's term, and that its efficacy must be nonnatural, and there is the conviction, firmly held by many commentators, that, at bottom, Kant shares this belief. Hudson has convincingly worked out the nature of the fundamental standoff between two very different persuasions which are squaring off here.⁶⁵ I can only hope that my tally of what is commendable and what, questionable in several interpretations of Kant's view has shed some light on the options available both to him and to us.

V. Conclusion

Kant's compatibilism issue turns out to be an interesting conflict between deeply held beliefs about the human power of action and equally deeply held beliefs about the power of natural causation. Each of the views that I have described pays a price for what it holds to be true. Wood opts for a nontemporal agency and cannot satisfactorily explain how an agency of this sort acts on what is natural. In the course of his attempted explanation, he jeopardizes the possibility of empirical knowledge. Beck demotes the Second Analogy to a regulative principle and holds that there must be uncaused events for there to be free actions. This again has enormous repercussions for the theory of knowledge of the first *Critique*. Allison retains that Analogy as a constitutive principle but winds up with two conflicting points of view asserting causation and freedom, respectively, and, it seems, with two efficacies somehow existing side by side, again jeopardizing empirical knowledge. Anomalous monism retains the Second Analogy in all of its power, avoids conflicting points of view and dual efficacies, but consequently treats the efficacy of the will as token-identical to that of natural causation.

Which freedom is the right one and which the right one to ascribe to Kant? Which price are you willing to pay?

Notes

1. *Journal of Philosophy* 37:1940, 561–71.
2. *Mind* 75:1966, 155–79.
3. New Haven: Yale University Press, 1975.
4. *New Elucidation* declares actions free when they are “called forth by . . . motives of the understanding applied to the will” and defines spontaneity as “action which issues from an inner principle” (1:400, 402). Kant’s Critical views on these matters are discussed in this paper.
5. For Kant’s explanation of this in 1755, see my discussion in part IV below (following my exposition of anomalous monism).
6. See my “Kant on the Nondeterminate Character of Human Actions,” in *Kant on Causality, Freedom, and Objectivity*, eds. W.L. Harper and R. Meerbote (Minneapolis: University of Minnesota Press, 1984), 138–63; and H.E. Allison, *Kant’s Theory of Freedom* (Cambridge: Cambridge University Press, 1990).
7. Again see Meerbote, *op. cit.*, and Allison, *op. cit.*, and L. W. Beck, *A Commentary on Kant’s “Critique of Practical Reason”* (Chicago: University of Chicago Press, 1960), among others.
8. See, for example, *Critique of Practical Reason*, 5:96–7 (for transcendental freedom), and *Groundwork*, 4:446 (for causality of the will).
9. A 189–211/B232–56. For freedom from pathological necessitation, see A534/B562; for freedom negatively conceived, *Groundwork*, 4:446.
10. Kant’s discussion of heteronomy, in *Groundwork*, 4:441, does not contradict this explanation. There Kant is thinking of determination of actions by objects (in accordance with causal laws of nature), and this indeed does not affirm spontaneity. See Allison’s discussion of what he calls the “Incorporation Thesis,” in *Kant’s Theory of Freedom*.
11. For one interpretation of the role that the distinction between determinateness and indeterminateness plays in Kant’s resolution of the compatibilism issue, see my discussion of anomalous monism, below. But other interpretations, such as Allison’s, also take this distinction to be of crucial importance.
12. I am not providing any arguments here for this interpretation of Kantian determination in its “fullest sense.” For some textual indications, see Kant’s conception of the object of empirical intuition as undetermined prior to categoriality (A20/B34) and his declaration (at B128–9) that categories determine intuitions of objects, as well as his distinction between determinate and reflective judgments, in the so-called “First Introduction,” 20:211. Reflective judgments are indeterminate.
13. Some defenders of this way of reading Kant believe, in turn, that this renders his defense of the principle of causation inadequate, whereas others disagree with this latter assessment. For references to defenders of the singular-causation reading of the Analogy, see notes 15 and 16.

14. Ably represented by, for example, Michael Friedman.

15. Two good sources are his "A Prussian Hume and a Scottish Kant," in *McGill Hume Studies*, eds. D.F. Norton, N. Capaldi, and W.L. Robison (San Diego: Austin Hill, 1979), 63–78; and "Über die Regelmässigkeit der Natur bei Kant," *Dialectica* 35:1981, 43–56.

16. See his discussion of causation, laws, and the Second Analogy in *Kant's Theory of Freedom*, esp. chapter 2.

17. Such as in his critique of Beck, *ibid.*, chapter 4.

18. It would seem, however, that a denial of the first of these two possibilities puts into question the very possibility of empirical psychology as an enterprise of the sort Kant wants to allow, viz., empirical psychology as merely descriptive, since it puts into question the very possibility of inner sense. The present passage may hence be suspect. However, the claim—which I go on to make—that apperception is not perception is supported by the doctrine of the *Critique of Judgment* which I refer to next. The basic point is that, for Kant, apperception of propositional states is not experiential. See his fascinating "Beantwortung der Frage: Ist es eine Erfahrung, dass wir denken?," in *Kleine vermischte Schriften*, ed. Kirchman et al. (Leipzig: Felix Meiner, 1873), 193–4. In this short exposition, Kant does, however, claim that thinking *produces* some inner state that can be observed (in contrast to his claim in the Preface of the *Metaphysical Foundations*).

19. "First Introduction," 20:220; *Judgment*, 5:359. In the latter, in one of his discussions of the argument from design for God's existence (section 75), Kant claims that God's intentions are not (the sort of conditions that are) observed but are only reflectively (indeterminately) thought. That it is divine intentions which are discussed here is incidental to the general point that intentions are not perceived.

20. *Metaphysical Foundations*, 4:471. This lack, of course, turns on the inner or mental character of the objects of empirical psychology, a character not described by Kant as spatial.

21. To my knowledge, Kant does not assert this anywhere, but surely he subscribes to this claim. The law of continuity that he mentions in the Preface (in the passage referred to in the previous note) is a law governing (the real in) sensation, not propositional states.

22. Such as at B274–6, and in the version (under the title "Widerlegung des problematischen Idealismus") in Kirchman's edition of *Kleine vermischte Schriften*, 197–8. In the latter version, Kant explicitly mentions *commercium*, i.e., interaction, covered by the Third Analogy. But since that Analogy requires spatiality and since Kant's conception of what is mental is not a conception of the mental as spatial, it becomes an interesting question, how what is mental can be in *commercio* with spatial substance. Whatever is the answer to this question does not reinstate either the First or the Third Analogy into empirical psychology.

23. As far as the inapplicability of the Third Analogy to apperception is con-

cerned, we will see below that Kant contrasts apperception with the externality required by that Analogy.

24. If only because B232–3 informs us that all succession of appearances is change of states of substances.

25. Not to mention the belief, shared by Kant, that lawfulness and freedom are, on the face of it, in opposition. (See unconditionedness, above.) But different explanations of this belief have been advanced (e.g., by Allison and by anomalous monism, below).

26. I have provided merely a negative characterization of indeterminate judgments. Kant furnishes a positive account in his theory of teleological reflective judgments, largely in the *Critique of Judgment* and the “First Introduction.” But the negative characterization is all that is needed here. It turns out that indeterminateness is necessary both for appropriation and apperception and for unconditionedness, with unconditionedness itself necessary for appropriation and apperception. (Indeterminateness is necessary for unconditionedness because the latter rejects the Second Analogy, at least in the nomological-necessity sense. Kant also believes that the categories stand or fall together.) Indeterminateness is also sufficient for unconditionedness (but not for appropriation and apperception). Our three criteria, then, are not independent of one another. Still, it is convenient to present them one by one, as I have done. For the insufficiency of indeterminateness (or unconditionedness) for appropriation and apperception, see also note 57.

27. Also recall spatiality, in note 22.

28. This is the discussion that leads to the famous rejection of the “freedom of the turnspit.”

29. For the relations between our three criteria, also see note 26.

30. He does so both in the *Critique of Pure Reason* and in his physicalist interpretation of his causal principle in the *Metaphysical Foundations*, as well as in his resolution of the Antinomy of Teleological Judgment, in the *Critique of Judgment*.

31. Prominent among the latter is the “turnspit” discussion. See note 28.

32. “Kant’s Compatibilism,” in *Self and Nature in Kant’s Philosophy*, ed. A.W. Wood (Ithaca: Cornell University Press, 1984), 73–101.

33. *Kant’s Compatibilism* (Ithaca: Cornell University Press, 1994), 26.

34. *Kant’s Theory of Freedom*, 52.

35. *Kant’s Compatibilism*, 25. Hudson and I are in complete agreement concerning Kant’s claim of sufficiency in behalf of natural causes.

36. We shall see that other interpretations of Kant’s view have him accept and defend a nonnatural, albeit temporal, agency. Hence it is helpful to distinguish the argument for naturalism from that for temporality. It will not help for Wood to protest that reasons (not causes) bring about (not cause) some temporal events. The point of the argument against him is that temporal events require to be caused, and

require to be caused by temporal causes. It also will not help to try to appeal to direct-agent causation. There is a general consensus that, for Kant, reasons move actions. Direct-agent causation would also fail to show any nontemporality of the agent.

37. See Robert Butts' exemplary discussion in *Kant and the Double Government Methodology* (Dordrecht: D. Reidel, 1984), chapter IV. Butts discusses Kant's informative 1796 preface to Sömmering's *Über das Organ der Seele* and shows how the views expressed there by Kant informed his entire Critical period.

38. Double affection, if there is such a twofold relation, is not likely to be specifying two efficacies, and even if it did, it would not specify two efficacies bearing on the phenomenal self or agent.

39. *Judgment*, sections 63–6, as well as in his resolution of the Antinomy of Teleological Judgment.

40. See his *Commentary*, and the later "Five Concepts of Freedom in Kant," in *Stephan Körner-Philosophical Analysis and Reconstruction*, ed. J.T.J. Szrednicki (Dordrecht: Martinus Nijhoff, 1987), 35–51.

41. *Kant's Theory of Freedom*, and *Idealism and Freedom* (Cambridge: Cambridge University Press, 1996).

42. Above all in his "Mental Events," in *Experience and Theory*, eds. L. Foster and J.W. Swanson (Amherst: University of Massachusetts Press, 1970), 79–101.

43. Hudson stresses the "of sorts." See *Kant's Compatibilism*, 31 n. 41. Beck also thinks that Kant moved away from a two-worlds view only belatedly.

44. *Kant's Theory of Freedom*, 75–6.

45. *Kant's Compatibilism*, 38–9.

46. In *Kant's Theory of Freedom* and in *Idealism and Freedom*.

47. I did so in "Allison on Freedom and Idealism," *Kantian Review* 1:1998, 136–50, and in a paper read at a conference in Allison's honor at the University of California at San Diego, in May 1997.

48. *Idealism and Freedom*, 19.

49. *Kant's Theory of Freedom*, 5–6.

50. See, e.g., *Idealism and Freedom*, 126.

51. *Kant's Theory of Freedom*, 28.

52. I have taken up some of the difficulties that Allison faces in this regard in "Allison on Freedom and Idealism."

53. Allison himself believes that his "points-of-view" interpretation deflates such criticisms since, according to him, Kant does not take these points of view to be "ontological" but rather to be just ways of considering events (in this case). I have discussed this Allisonian strategy in the paper referred to in note 47.

54. See note 42.

55. Unbeknownst to some recent philosophers of mind, Kant's view, then, is very much an object of philosophical debate today (if Davidson's reading of him is on target). For those "further" developments, see my "Kant on the Nondeterminate Character of Human Actions," and Hudson's *Kant's Compatibilism*.

56. The indeterminate judgments which are employed when some natural event is described as a free action are judgments which express what is intelligible or noumenal in Kant's best sense of those terms. In Kant's world of intelligences, anomalousness holds sway.

57. There also are two problems that I will not discuss. There is the issue of freedom requiring the ability to do otherwise. (This is a requirement that Kant accepts during his Critical period.) In *Kant's Compatibilism*, 93–6, Hudson has defended David Lewis' account of this ability, in a manner congenial to the monist interpretation of Kant. The other matter concerns a feature internal to anomalous monism as a reading of Kant. According to this reading, indeterminateness (or unconditionedness) is not sufficient for appropriation and apperception of reasons and hence is not sufficient for freedom. (See note 26.) Appropriation and apperception are needed for freedom in their own right. This is so because anomalous monism treats all propositional mental life as anomalous and hence as unconditioned, but not all mental life of this sort is free. The special nature and role of apperception and the relation between apperception and the rest of propositional mental life hence deserve further scrutiny (above and beyond the anomalous structure of apperception).

58. I am using "know" here in Kant's strict first-*Critique* sense in which only determinate judgments are properly cognitive. Teleological reflective judgments and/or practical judgments such as a belief that some agent acts from some particular motive or maxim are not cognitive in this strict sense. (Cf. A551/B579, note.) Hence Kant's famous distinction between what is "theoretical" and what is "practical."

59. Cf. the previous note. Token-identity and unconditioned actions are beyond the pale of "theoretical" understanding.

60. He voices this sort of objection both to my account in his *Kant's Theory of Freedom*, 80–2, and in his review of Hudson in *Philosophical Review* 105:1996, 125–7. For Hudson's reply to Allison's criticisms, see the former's review in *Kant-Studien* 82:1991, 219–22; for mine, "Allison on Freedom and Idealism."

61. *New Elucidation* (1755). One reason why I refer to Kant's early effort as "quasi-Leibnizian" is that Kant, against Leibniz, allows for fundamental "extrinsic denominations" (to use Leibnizian language). For Kant, substances are in fundamental respects externally related. Another reason is that in 1755, Kant, contrary to Leibniz, does not take the ability to do otherwise to be required for freedom.

62. This is "deterministic in the fashion of all other states or events subject to the Principle of Sufficient Reason," which later was to evolve into Kant's "determinate." Deterministic conditions here are natural.

63. See *New Elucidation*, 1:401, 403. (This externality is one of Kant's disagreements with Leibniz. See note 61.)

64. *Critique of Practical Reason*, 5:97. Or that of a projectile in free flight or a wound-up clock.

65. In *Kant's Compatibilism*, 96–8.

Consequentialism and Its Consequences

Robert L. Holmes

Consequentialism, as commonly understood, tells us, in effect, always to do what is best. And how, one wants to ask, can that be improved upon? If it cannot, why bother with the categorical imperative, the divine command theory, natural law, and other pretenders to the throne of supreme moral principle? We know there are acts. We know they have consequences. And we know these consequences are good or bad in varying degrees. We do not know whether there is a God, or if there is one, exactly what he commands. Even less can we be sure what Kant meant, or what cognitively deprived rational beings would agree to in a cloistered original position. So why not simply cast our lot with William James when he says there is “but one unconditional commandment, which is that we should seek incessantly, with fear and trembling, so to vote and to act as to bring about the very largest total universe of good which we can see”?¹ And having done that, why not devote our energies to working together to achieve that end?

This oversimplifies, of course, but it contains the elements of a *prima facie* case in favor of consequentialism.

Two other claims reinforce this case. Both, if correct, have far-reaching implications. They bear in particular upon the relevance of moral philosophy to practical affairs, especially in light of the growing conviction—not shared by all, by any means, but nonetheless widespread—that moral philosophy *ought*, in some sense, to concern itself with the problems of people in what social scientists like to call the “real world.”

The first claim is that, whether philosophers intend it or not, ethical theory over time works its way into the thinking of ordinary persons, where it exerts a subtle and largely unnoticed influence. So that while it may be debated whether philosophers should expressly tackle problems like poverty, discrimination, euthanasia, and war, philosophy is already having its influence in the way of spinning the perspectival web within which those problems are located and understood. Dewey captured the spirit of this outlook early in his career, when he observed:

Not even customary morality, that of respectability and convention, is freed from dependence upon theory; it simply lives off the funded results of some once-moving examination of life.

[Ethical theories] filter into the average consciousness, and their truth becomes, wholly unaware to the average consciousness, a part of the ordinary insight into life,—a part of the meaning of the world of practice in which we live. Life looks different to-day to the man to whom Bentham and Kant are not even names, because of the formulae of the greatest good, the autonomy of the will, and the categorical imperative.²

The second claim is more specific. It assigns a privileged position to utilitarianism. Thus, for example, Sidgwick sees utilitarianism as underlying common sense morality, though not consciously so in people's minds. "It is there," he says, "not as the mode of regulating conduct with which mankind began, but rather, as that to which we can now see that human development has been always tending, that Utilitarianism may most reasonably claim the acceptance of Common Sense."³ The thought here is that an evolutionary process is at work in which one substantive moral outlook, that of utilitarianism—the preeminent form of consequentialism—is gradually emerging as dominant.

Not only does consequentialism have a strong, precritical intuitive appeal, if Sidgwick is correct, it also has achieved—or, at any rate, is well on its way to achieving—a *fait accompli* in the workaday world of practical ethics. It is becoming the dominant moral outlook in human affairs. While philosophers debate the merits of different theories, humankind is issuing the only verdict that really counts. And for better or worse, on his view, that verdict is in favor of consequentialism.

Here serious problems surface. For some who agree that a particular moral outlook has come to dominate much of common sense thinking, and who agree that this outlook is consequentialist, part company with Sidgwick in his approval of this state of affairs. They think, rather, that this development is regrettable and represents a trend that should be reversed.

Stuart Hampshire, for example, contends that, especially since World War I, the prevailing morality, particularly in governmental thinking, has been utilitarian—what he labels, disparagingly, "computational morality." He contends that this "utilitarian habit of mind," as he calls it, "has brought with it a new abstract cruelty in politics, a dull destructive politi-

cal righteousness: mechanical, quantitative thinking, leaden academic minds setting out their moral calculations in leaden abstract prose. . . .”⁴ Just as Dewey argued that the quest for certainty in philosophy has had bad consequences throughout the centuries, the thought here is that the consequentialism embodied in the thinking of experts has likewise had bad consequences. Nor, among those who advance this objection, are consequentialism’s bad consequences limited to those that issue from the thinking of Western quantitative analysts. Arthur Koestler, in a related criticism, sees utilitarianism as having taken hold in a whole ideology, namely Marxism, with similar consequences. Where Hampshire speaks disparagingly of “computational morality,” Koestler does the same of what he calls “Commissar-Ethics”:

The transfer from the physical to the ethical level of the principles of quantitative measurement has probably produced the most disastrous results. The implied paradoxa of this kind of “Commissar-Ethics” are less obvious to us than those of the biological ethics of fascism because we have been so thoroughly trained to think in quantitative terms that the application of mathematical criteria to ethical method appears to us simply as an act of common sense. . . . We accept the argument of Soviet apologists that it is better to keep a thousand innocents in jail than to let one spy go free whose activity might endanger the lives of tens of thousands. And we do not notice the hitch in the argument, namely, that we have no physical instruments to measure the exact amount of harm caused by the detention of the thousand innocents and to compare it with the amount of harm to be expected from the hypothetical spy.⁵

Though neither Hampshire nor Koestler claims that consequentialism represents the mature morality of common sense, as Sidgwick suggests, each alleges that it has taken over important areas of thought and action, with unfortunate, if not disastrous, consequences.

What shall we make of such an objection? Can it be met, and can the promise of consequentialism be realized?

So stated, the objection obviously needs qualifying. Consequentialism itself has no consequences, good or bad (other, perhaps, than logical ones), so we must understand the objection to mean that the *holding* of consequentialism—its acceptance or following—has bad consequences. In order for its presumed bad consequences to be of the magnitude alleged by Hampshire and Koestler, one would have to assume that

consequentialism is, if not the dominant view of commonsense morality, at least the dominant view of some significant portion thereof—at the very least, of a significant number of persons, particularly those who wield power and influence. I shall not try to assess that claim, though a case could be made for it with regard to industrial, technocratic society (there is the further question, of course, of whether, if this claim is true of some societies, it is true of others, which it might well not be). Let us consider, rather, some general considerations relevant to assessing the objection.

Before turning to these questions, let me indicate what I shall take consequentialism to be. For it has come to stand for a variety of different things in recent literature.

I

I have spoken as though consequentialism were more or less identical to utilitarianism, and it is often understood in this way, particularly when contrasted with deontological ethical theories. At other times, however, it is identified more or less with teleological ethics in general (which itself is sometimes used more or less interchangeably with utilitarianism). Used in either of these ways the term represents an unnecessary proliferation of “isms,” since it fails to mark out any distinctions not already covered by utilitarianism and moral teleology.

Consequentialism can, however, be understood in a way not covered by either utilitarianism or teleologism, and a way that is helpful in defining those other positions. In this sense it is simply the thesis:

C: Rightness and wrongness are determined solely by the consequences of actions.

Left open is what it is about consequences that is relevant. This means that consequentialism may take axiological or deontological forms, depending upon whether the comparative value of consequences is their relevant feature or putatively nonvalue considerations, such as whether they occasion rights violations, and if so, in what comparative measure (the latter type theory sometimes being called a “utilitarianism of rights”). Utilitarianism can then be understood as a consequentialist theory with a typically axiological orientation.⁶ But it is not the only such theory. Ethical egoism is a consequentialist theory with an axiological orientation, as

are various forms of what I call macro ethics—theories that take the relevant beneficiary of the consequences to be collective or abstract entities like the community, nation, or state. All such theories—theories, that is, that are consequentialist and take the resultant comparative value of consequences to determine right and wrong—may then be understood to be teleological (in the sense that they implicitly or explicitly identify an end to moral conduct, the maximization of value for the relevant entity). As such, they can be contrasted with the various deontological theories.

The assessment of the foregoing objection in the next section will apply primarily to utilitarianism, as that is the specific consequentialist theory Hampshire and Koestler attack. Thereafter, in presenting a different objection to consequentialism, I shall take consequentialism to stand for any theory that affirms C, whether it be utilitarian or not.

II

Let us now return to the objection cited earlier. Does consequentialism have the alleged bad consequences, and if so, should it be rejected on those grounds?

I do not think this objection as it stands is sustainable, and for the following reason.

Let us begin by distinguishing among *accepting*, *following*, and *conforming to* a principle. On the face of it, one might accept a principle, in the sense of believing that it is the correct principle, without trying to live by it or doing so consistently.⁷ Arguably, many Christians accept Christ's teaching in the Sermon on the Mount in this way, as do many Marxists the teachings of Marx. One might even concede that a certain principle is the correct principle of conduct and yet resolve not to act in accordance with it. To *follow* a principle, on the other hand, is not only to accept it, but to try to implement it as well—not necessarily invariably, but by and large and for the most part. It is to incorporate it in some significant way in the guidance of one's conduct. To follow a principle, in this sense, however, guarantees nothing about one's success in trying to implement it. *Conforming to* a principle, on the other hand, signifies acting in accordance with it—that is, actually doing as it prescribes. Normally that would imply accepting and following it, but it need not. Somewhat improbably, one could in fact do as a particular principle prescribes purely by chance, without consciously following any principle at all. Somewhat more plau-

sibly, one might conform to one principle while following another. This could happen if the two principles were extensionally equivalent in the sense of yielding the same prescriptions (it has sometimes been thought, for example, notably by Plato, that prudence and morality coincide in just this way). One might even conform to a principle without in fact knowing that he was conforming to it, but one could not accept or follow a principle without knowing that he was doing so. We may say, then, that accepting a principle does not entail following or conforming to it; following a principle entails accepting and at least trying to conform to it.⁸ And conforming to a principle, while it probably normally involves accepting and trying to conform to it, does not strictly imply either accepting or following it.

If we assume that consequentialism is generally accepted, or accepted by a significant number of persons, and if the conditions in the society in which this is true are deemed bad, the Hampshire-Koestler objection might mean that this unfortunate state of affairs is a consequence of the fact that consequentialism is held as widely as it is. But if this is how the objection is understood, the consequentialist has a reply to it. It is that this unfortunate state of affairs is not due to the holding of consequentialism, but rather to the fact that there are so many who do not hold it. So long as the *acceptance* of the principle is significantly less than universal for the community in question, it may be that the bad consequences are the result of the conduct of those who do not accept consequentialism, not a result of the conduct of those who do, even if the latter are in positions of power. One can hardly hold against consequentialism the conduct of those who do not subscribe to it. It is, after all, the point of advocating consequentialism as a moral principle to encourage people to follow it, presumably in the conviction that their so doing will have the best consequences. If *not* following it has consequences that are less than ideal or even comparatively bad, that is precisely what one should expect. Indeed, the consequentialist might even cite the critic's facts as tending to support the case for consequentialism.

The objection cannot mean that conforming to consequentialism has bad consequences, since the principle (in its utilitarian form, as we are understanding it) calls upon us to do what is best. If everyone does that, we cannot improve upon what we have done.⁹ If not everyone does that, then the consequentialist can revert to the previous reply and lay the alleged "bad consequences" at the feet of those who do not conform

to consequentialism, whether their nonconformity is intentional or not.

I suggest, for these reasons, that the most plausible way to understand the objection is as saying that *following* the principle of consequentialism has bad consequences.

How can following consequentialism have bad consequences? It can if those who follow it fail to implement it correctly—fail, that is, to do precisely what its central principle prescribes. This can happen if those who follow it are weakly committed to it or weak of will. It can also happen if the principle is difficult to apply, as arguably it often is because of the difficulty of projecting the consequences of alternative actions. People with the best of intentions may often fail to do as consequentialism prescribes because what that is often is simply too difficult to ascertain. The temptation, then, is to do whatever one is predisposed to do, for whatever reasons, and to rationalize to oneself that this is what will have the best consequences.

This is a serious objection. The problems in trying to project consequences over an indefinite future were detailed by Moore,¹⁰ himself usually considered a utilitarian, and they have never been convincingly resolved. Those problems certainly make it tempting to convince oneself that the course one prefers for other reasons is the course called for by consequentialism.

Still, consequentialists might well argue that just as the consequences of the actions of those who do not accept consequentialism cannot reasonably be calculated among consequentialism's consequences, so the consequences of the acts of those who carelessly apply it—or worse yet, willfully misapply it—cannot fairly be reckoned among those consequences either. Advocates of any principle are entitled to presume, in advancing it for everyone alike, which arguably is a condition of a principle's qualifying as moral, that any reckoning of the consequences of its adoption will include only the consequences of its both being adopted and being correctly applied. Absent the satisfaction of this condition, the consequentialist can turn back against his critics the charge that, in alleging the bad consequences of consequentialism, they are tacitly presupposing consequentialism itself. Bentham had anticipated this response when he said:

When a man attempts to combat the principle of utility, it is with reasons drawn, without his being aware of it, from that very principle itself. His arguments, if they prove any thing, prove not that the

principle is *wrong*, but that, according to the applications he supposes to be made of it, it is *misapplied*.¹¹

Seen in one light this reply seems decisive. For if consequentialism directs us—all of us, not just some of us—always to do what is best, and if conformity with that principle entails doing what is best, then the consequences of such conformity can never be bad, at least by comparison with available alternatives. If so, consequentialism seems impervious to any other than a frontal deontological challenge to the very idea that it can never be right to prefer a worse state of affairs to a better, even all things considered. (The deontologist, of course, can always say that producing the best consequences may fail to bring about what is *morally* best because it disregards other morally relevant considerations.)

But seen in a different light, this reply comes up short. True, one wants to say, if consequentialism tells us always to do what is best, then if we in fact do that, those consequences cannot be improved upon. But that simply reaffirms the original intuition with which we began. Consequentialists understandably want to disqualify from consequentialism's assessment the consequences of misapplying it. Nonetheless if there is something about the principle that systematically prevents it from being applied correctly, or makes it impossible to state precisely what a correct application would be, then that weighs against the principle, whether or not the failure is counted among consequentialism's consequences. Consequentialism, we should remember, represents advocacy of a particular principle, no less than does, say, Kantianism. We can all see the difficulties in applying a principle like the categorical imperative and properly consider it a failure of Kantianism if solutions to those problems cannot be found. So we should be prepared to acknowledge that if there are equally formidable difficulties in applying consequentialism's principle, that should be accounted a failure of consequentialism as well—not necessarily of the principle per se, but of consequentialism as a theory of ethics that presumes to guide moral conduct.

It is just such a difficulty that I now want to explore. It is concerned with consequentialism's consequences, now not in the sense of the consequences of following the principle, but in the sense of the consequences to which the theory would have us appeal in attempting to determine what is right and wrong. I shall henceforth take consequentialism in the broader sense characterized earlier, as standing for any theory which affirms C, not merely for axiological theories, and not merely for utilitarian theories.

III

Consider the following passage from Elizabeth Anscombe's essay "War and Murder":

The distinction between the intended, and the merely foreseen, effects of a voluntary action is indeed absolutely essential to Christian ethics. For Christianity forbids a number of things as being bad in themselves. But if I am answerable for the foreseen consequences of an action or refusal, as much as for the action itself, then these prohibitions will break down. If someone innocent will die unless I do a wicked thing, then on this view I am his murderer in refusing. So all that is left to me is to weigh up evils.¹²

My interest here is not with the principle of double effect, which one can see lying just beneath the surface of these remarks. It is with the implicit conception of what is properly includable among an act's consequences. In disavowing answerability for unintended though foreseen consequences in the sort of case she cites, it is clear that Anscombe is allowing that the actions of others are sometimes among the consequences of our own actions (the death of the innocent of which she speaks is a death at the hands of someone else). It is just that when they are, and when they are unintended, those consequences are not, in Anscombe's view, morally accountable to us, hence do not bear upon the rightness of what we do.

Two issues emerge here. The first is whether the free and deliberate actions of others are ever properly reckoned among the consequences of our own actions. The second is whether, if they are, they should be placed on the scales when it comes to moral calculations, and if so, with what weight.

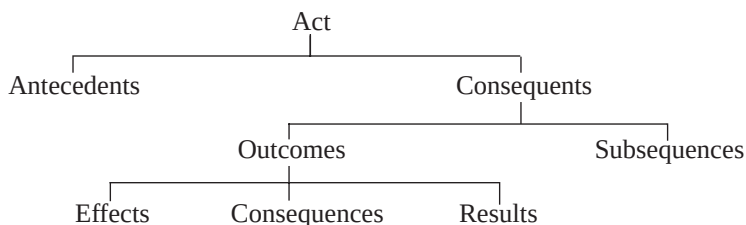
Why is this a problem? It will be helpful to answer this by placing the issue in a broader perspective by looking more closely at the nature of consequences.

IV

Everything that happens prior to a given action is a temporal antecedent of that action; everything that happens after, a temporal consequent (things that happen concurrently may be called temporal concomitants, but I shall,

for the sake of simplicity ignore them for present purposes). In these terms an action's consequences are among its consequents. But not just any consequent counts as a consequence; indeed, most do not. If I move my arm and an apple falls from a tree in Shanghai, that happening (along with every other event throughout the remaining history of the universe) is a consequent of my action. But it surely is not a consequence.¹³ On the other hand, if in moving my arm I knock over a glass of water, the water's spilling clearly *is* a consequence of my action.

Let us call any consequent of an act that would not have occurred but for the act an *outcome* of that act. Any consequent that is not an outcome let us call a *subsequence* of that act. Consequences clearly are among an action's outcomes. Along with effects and results, I suspect that they exhaust those outcomes. The relations among these notions are represented as in table 1.



The differences among effects, consequences, and results resist precise delineation; ordinary language often uses them loosely and sometimes interchangeably. But arguably effects are the narrowest of the three, depending, as they do (by definition), upon a causal connection between the action and its outcome; and results the broadest, encompassing both effects and consequences and including some outcomes that do not fit so readily into either of the other categories (we might post the results of an exam, for example, but we could not post its effects or consequences; just as we announce the results of an election but, again, not its effects or consequences). Consequences include effects, and sometimes the two notions are used interchangeably (as in the Anscombe quotation above, and as in Moore's analysis in chapter 5 of *Principia Ethica*). But there are some candidates for consequences to which the notion of effects is inappropriate; notably, for our purpose, those outcomes that involve the actions of others.

Suppose that, rather than the apple falling after I move my arm, someone picks the apple. Barring an extraordinary configuration of circum-

stances, that action would not be an outcome of my own. If not, it is a mere subsequence of my action, not a consequence. But suppose terrorists threaten to execute hostages if their demands aren't met and carry out the threat when the deadline passes. Are the hostages' deaths a consequence of the refusal? Or, to take a concrete case, when President Bush threatened military action against Iraq unless President Hussein withdrew Iraqi troops from Kuwait, was the subsequent U.S. attack, and the casualties it entailed, a consequence of President Hussein's refusal to withdraw; that is, was Bush's order to attack and its execution by thousands of military personnel a consequence of Hussein's act?

Why are these cases troublesome? It is because some decision or choice of someone else has entered the picture.

If I shout "Hello!" from Echo Mountain, no cognitive response by anyone need intervene before my call echoes back to me. I shout, and the laws of physics do the rest. But if I cry "Fire!" in the proverbial crowded theater, cognitive responses on the part of others must intervene before panic ensues. Consequences, it seems clear, fall into one or the other of two kinds: those that are mediated by some thought, choice, decision—some mental operation, if you like—before they come about, and those that are not.¹⁴ The echo is an unmediated consequence of my shouting; the panic is a mediated consequence.

The picking of the Shanghai apple, of course, presupposes some mental act on the part of the picker. So, many subsequences as well as consequences are mediated (every deliberate act by anyone throughout the future is, in fact, a mediated subsequence of the act of which it is a subsequence). But this raises no distinctive problems. The problems arise when consequents of a particular act are both mediated and outcomes of that act, that is, such that they would not have come about but for the act. Not all such cases give rise to the problem I want to focus upon, however. Every deliberate act by everyone living today is an outcome of many acts performed by their ancestors thousands of years ago (procreating just when they did, if nothing else), and the outcomes of every such act are mediated outcomes of those acts. Likewise, the outcomes of every free and deliberate act I perform today are mediated outcomes of many acts I performed in the past (choosing to shop at a particular market is a mediated outcome of having chosen years before to reside in the particular city containing that market). There are, of courses, questions of responsibility that arise in at least some of these cases. If I give five dollars to a street person who promptly retires to the nearest bar with it, his subse-

quent consumption of alcohol is a mediated consequence of what I did. But it is mediated not only by his decision to go to the bar, but also by the bartender's serving him. This raises the question of where and how moral responsibility for his inebriation should be distributed (whether principally to him for asking for money then heading for the bar, or to me for granting his request, or to the bartender for serving him, or to all three in varying proportions).

The problem that interests me here concerns those cases in which mediated outcomes are *responses* to what someone else has done. For often these are considered consequences of the original act. The door's closing is a consequence of my having asked you to close it; the platoon's lurching forward is a consequence of the sergeant's order to march; and, of course, the calamitous panic in the crowded theater is consequence of someone's shouting "Fire!". Children often hear that punishment will be a consequence of misbehavior; governments often threaten that adversaries will suffer the consequences if they do or fail to do some particular thing.¹⁵

Are all of an act's consequences relevant to its moral assessment? Or only some? And if only some, which ones? And why?

By "all" here I do not mean the problem Moore dealt with, of where to draw the line in the unfolding of consequences into an indefinite future beyond which you may assume that the appraisal of their value won't be reversed. I mean, rather, are all *kinds* of an act's consequences morally relevant—specifically, both mediated and unmediated?

There is no problem (theoretically, at least), with unmediated consequences, particularly when they are effects of the acts in question. They unquestionably belong on the scales for any consequentialist trying to determine an act's rightness. But what of mediated consequences? Particularly those mediated by responses to the original action? Are they to be weighed as well?¹⁶

Clearly not everything others do in response to an act of mine—much less everything they would not have done but for my act—can plausibly be considered a consequence of that act, at least for purposes of moral assessment. That would lead to a proliferation of consequences so great as to compound immeasurably the already seemingly intractable problem of projecting and assessing consequences. Equally importantly, it would enable persons to determine many of the consequences of others' actions simply by the way they respond to them. From a consequentialist standpoint (as well as from nonconsequentialist standpoints that allow a

significant role to consequences), this would enable those appropriately situated, and possessing enough power, to make what others do morally right or wrong by seeing to it that the consequences of what they do are good or bad in the appropriate proportions. This is implied by Moore, with regard to judicial punishment:

One of the chief reasons why an action should not be done in any particular state of society is that it will be punished; since the punishment is in general itself a greater evil than would have been caused by the omission of the action punished. Thus the existence of a punishment may be an adequate reason for regarding an action as generally wrong, even though it has no other bad effects but even slightly good ones. The fact that an action will be punished is a condition of exactly the same kind as others of more or less permanence, which must be taken into account in discussing the general utility or disutility of an action in a particular state of society.¹⁷

Although Moore does not expressly recognize the distinction between what I am calling mediated and unmediated consequences, his reasoning presupposes it. And the extension of such reasoning provides a foundation for a kind of moral totalitarianism. Any government with the power to determine the “consequences” of failure to do as it says—or of doing what it forbids—would not only be able to make disobedience imprudent, as governments regularly try to do anyway, but would make it morally wrong as well. The same would apply to any person or persons with sufficient power. They could render the actions of others right or wrong, after the performance of those actions, and even without the foreknowledge of those performing them. Nudged only slightly to the extreme, this way of thinking would enable one to determine the moral character of the acts of persons long deceased.

At the same time, some mediated consequences seem incontrovertibly to be properly regarded as consequences of prior actions of others. If Kasparov places Karpov’s king in check, and Karpov subsequently moves the king, it is natural to say that he did so as a consequence of the check. The same holds when one says that FDR’s call for a declaration of war was a consequence of the Japanese attack upon Pearl Harbor, or President Bush’s mobilization of U.S. troops was a consequence of Iraq’s invasion of Kuwait. Moreover, the notion of coercion is sometimes understood as threatened intervention in the consequences of an action (or omission) by the performance of an act that the actor would find undesir-

able with a view to dissuading him from performing that action. And the very idea of deterrence (nuclear or otherwise) is that of leading others to refrain from acts they otherwise would perform by making it known to them that the consequences of those acts would be unacceptable to them—unacceptable because those issuing the threats will make them unacceptable. More philosophically, the Prisoner's Dilemma gets off the ground only because the responses of others (those controlling the hypothetical situation) are factored among the outcomes of the options available to the prisoners.

So how does one determine which mediated consequences should be counted for purposes of moral assessment, and which not?

There is no apparent way to answer this in morally neutral terms. The flow of events of which our doings are a part does not admit of a sharp and indisputable individuation of consequences—and perhaps not even of acts either—when the conduct of others is sluiced into them, as invariably it is in interpersonal, social, and international affairs of greatest moral interest. This, indeed, may be what most significantly distinguishes our assessment of human affairs from that of the physical world by science, and casts doubt upon the possibility of anything like a science of ethics, or perhaps even a value-neutral social science.

This opens the way to bias in the judgment of controversial cases in the guise of making neutral consequentialist calculations. This is a problem that would remain even if Koestler's concern about quantificational calculations could be allayed. If a woman is assaulted walking down an inner city street at night, her rape is a mediated consequence of what she did. But should it be considered a consequence for purposes of moral assessment? If it should, it would, *ceteris paribus*, make her walking down that street at night—supposedly something anyone has a right to do—morally wrong from a consequentialist standpoint, with all that implies about further judgments of responsibility. Those who contend that women in such circumstances are asking for trouble tend to count such consequences, to the point of saying in such cases that women act wrongly. Those who contend that to say that is just another way of blaming the victim insist that her victimization was an act of the rapist, pure and simple. If two gays walk down the street holding hands in an area known for attacks by homophobic gangs, is their subsequent victimization a consequence of their action for purposes of moral assessment? Again, predispositions reflecting underlying biases are likely to determine the answers to this question.

The problem is not simply that people often give way to temptation to tailor their estimates of consequences to support what they are predisposed to believe—where, with enough care and conscientiousness, they could get matters straight. Here the problem is that there is no objective nonnormative fact of the matter to be ascertained. One must make the determination in some sense a matter of decision. The question is: which among an action's consequences in situations of this sort *ought* to be counted for purposes of moral assessment? This is a moral question itself, nested inside what to appearances is a neutral determination of facts. And the answers to it often presuppose moral judgments. We are judging more than just an act in these cases; we are judging broader contextual wholes containing acts and various candidates for their consequences.

If this is correct, it means that a rather simple model of what is involved in the making of moral judgments is inadequate. According to that model, there are "facts" and there are "values" (taking this broadly to include norms and principles), and to judge the latter you must first ascertain the former, which is an objective, quasi-scientific undertaking. Having established the facts, you then apply the appropriate principle and arrive at a moral judgment. But this will not do if the preceding account is correct. For on that account one is involved in the making of moral judgments in the very determination of what the facts are (where what the consequences of an act will be are among the preeminent "facts").

It is beyond the scope of my present concerns to try to spell out what enters into the making of these judgments. But these considerations highlight the importance of recognizing the non-neutrality of many so-called "facts" in social affairs and suggest that, to a considerable extent, our understanding of the social world may be constituted by evaluations. Whereas early epistemological idealists recognized an egocentric predicament that led them to conceptualize the world as constituted by perceptions (strictly, the ideas or data of perception), and Kant argued that nature cannot be understood apart from the order and coherence our mind brings to it, perhaps we need to recognize a kind of moralcentric predicament, in which we cannot fully determine our world independently of moral evaluations—a world in which, as Plato thought, the notion of the good (or perhaps some other normative concept) is as essential to our basic understanding of human affairs and institutions as the sun is to our understanding of the processes of the natural order. Having conceptualized the natural order in a way which from the start excludes values, much of ethical theory has puzzled over how to derive value from fact, or

“ought” from “is.” Perhaps the problem, rather—at least where the consequences of acts are concerned—is how to derive facts from values.

V

What implications does this have for consequentialism? They are, I believe, that it fails as an ethical theory.

The problem is more serious than just one of overcoming difficulties in applying correctly whatever consequentialist principle one might take to be authoritative. It is that there is no correct way to apply it in the sense intended. There is no morally neutral determination of consequences that enables us to say what the principle prescribes in particular situations. Put another way, insofar as consequences are regarded as properties of actions, and insofar as morally neutral determination of the properties of alternative actions is a precondition of being able to apply a consequentialist principle correctly, that condition cannot be met when at least some of the consequences are mediated.

For all of this, I believe we should cling to the sentiment that it can never be right to prefer a worse state of affairs to a better; or, at least, we should cling to it until forced to relinquish it by compelling reasons. It is just that I do not think this sentiment supports consequentialism. But it does support a nonconsequentialist axiological position, as, I believe, does the earlier quotation from William James. The particulars of such an outlook I am not able to detail here; but I would suggest that adumbrations of them may be found, not so much in Western philosophy (though, in different form, they may be found in Stoicism and Kantianism) as in Eastern thought, particularly in Taoism and some of the classics of Hinduism and early Buddhism, and more recently in the philosophy of Gandhi. This is an outlook which stresses that, in important respects, it simply is not given to us to know the consequences of actions. Hence we must look to other grounds for our moral judgments (or leave off making them altogether, as Zen Buddhism would sometimes have it). This does not mean that we should not have ends and pursue them with dedication. It means, rather, that our primary emphasis should be upon the means to adopt rather than upon ends, in the faith that if they are good, whatever they lead to in the way of consequences will stand the best chance of also being good.

So, in a sense, we might in the end agree with William James that the best we can do is to try to realize the greatest total universe of good that

we can. It is just that projecting, measuring, and comparing consequences may not be the best way to do that.¹⁸

Notes

1. William James, "The Moral Philosopher and the Moral Life," in *William James: The Essential Writings*, ed. Bruce Wilshire (New York: Harper Torchbooks, 1971), 297.

2. John Dewey, "Moral Theory and Practice," in *John Dewey: The Early Works, 1882–1898* (Carbondale and Edwardsville: Southern Illinois University Press, 1969), 96, 102 f.

3. Henry Sidgwick, *The Methods of Ethics*, 7th ed. (Chicago: University of Chicago Press, 1962), 456 f. Bentham had, of course, in 1789 made the even stronger claim on behalf of the principle of utility that "[b]y the natural constitution of the human frame, on most occasions of their lives men in general embrace this principle, without thinking of it: if not for the ordering of their own actions, yet for the trying of their own actions, as well as of those of other men"; *An Introduction to the Principles of Morals and Legislation* (New York: Hafner Publishing Company, 1948), 4. An opposing appraisal was made by William E. H. Lecky in 1869 when he wrote that "our first question is naturally how far this theory agrees with the feelings and with the language of mankind. But if tested by this criterion, there never was a doctrine more emphatically condemned than utilitarianism. In all its stages, and in all its assertions, it is in direct opposition to common language and to common sentiments"; *History of European Morals from Augustus to Charlemagne*, vol. 1 (London: Longmans, Green, and Co., 1882), 34.

4. Stuart Hampshire, *Morality and Pessimism* (London: Cambridge University Press, 1972), 7.

5. Arthur Koestler, *The Yogi and the Commissar and Other Essays* (New York: Macmillan, 1945), 242 f.

6. I shall, for the sake of simplicity, ignore the many possible permutations of utilitarianism, such as those which appeal to preference satisfaction.

7. I say "on the face of it," mindful of the controversy over internalism and externalism on the question of the relationship between obligation—or recognition of obligation—and motivation, and of the complexities in spelling-out the relationship between motivation and conduct.

8. I say "normally" because one could contrive cases in which one is forced, say, at gunpoint, to follow a particular principle that he does not accept.

9. Deontologists will maintain, of course, that it may nonetheless have bad consequences *morally*. Maximizing value, on their view, does not ensure that all other morally relevant considerations are given their due.

10. G.E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), chapter 5.

11. Jeremy Bentham, *op. cit.*, 5. Robert Paul Wolff made a similar observation regarding the Hampshire-Koestler objections in discussion following presentation of this paper at the College of the Holy Cross.

12. Quoted from Richard Wasserstrom, ed., *War and Morality* (Belmont, CA: Wadsworth Publishing Company, 1970), 50.

13. Not that with enough philosophical ingenuity one could not contrive hypothetical cases in which there were sufficient interconnections between the two events to warrant calling the one a consequence of the other; only that, in the normal course of affairs, this would be an inaccurate characterization. And by no plausible stretch of even the philosophical imagination could all of the subsequent happenings throughout the remaining history of the universe be called the consequences of my action.

14. More broadly, consequents in general—whether outcomes or subsequences—break down into mediated and unmediated. I shall focus upon outcomes, and among outcomes, upon those that are plausibly regarded as consequences. This is partly to keep the focus upon the concept that is central to consequentialism as an ethical theory (we do not yet have Outcomism as such a theory), but also partly because the notion of consequences seems more at home in our everyday discussion of issues than is that of outcomes.

15. The mirror image of many of these issues is found in tort law in the attempt to clarify causation, where, most generally, an act or omission is said to be the cause in fact of an injury if but for the act or omission the injury would not have occurred (thus making the injury an outcome of the act, as I have defined these notions). This leaves the proximate cause to be determined by further refinement of the contributing and determining factors in the injury (rendering it, in the preceding terms, an effect, consequence, or result of the act).

16. One could of course expand the notion of mediated consequences to include responses by nonhumans as well—such as the outcomes of throwing rocks at a hornet's nest or not letting a sleeping dog lie. While these raise important conceptual and ethical issues on their own, I shall, for the sake of simplicity, limit mediated consequences to those involving cognitive acts of humans only.

17. G.E. Moore, *op. cit.*, 159.

18. I want to thank Jill Frank, David O'Connor, and David Tabachnick for helpful suggestions during the writing of this paper. I have also benefited from discussions of the paper as presented at Florida State University, Iowa State University, Notre Dame University, the University of South Carolina, and the University of Virginia.

Another Look at Maxims

Rüdiger Bubner

I

In discussions concerning Kant's moral philosophy, the concept of law usually overshadows the concept of maxim. This is not at all surprising. For systematic reasons, Kant himself wanted to establish within moral philosophy an understanding of law that, in terms of its rigor, would not be inferior to the laws of nature (see *Groundwork*, 4:52). Thanks to the transcendental revolution, practical philosophy was once again to be raised to the level of theoretical knowledge. The strategy informing the entire undertaking is closely connected to Kant's departure from the traditional doctrine of *eudaimonia* and the attending notion that practical reason is a form of practical knowledge or prudence. For the sake of reason's unity, Kant wanted to bridge the ancient gulf separating theory from practice that opened up during the transition from Plato to Aristotle.

It is obvious, however, without recourse to maxims considered as subjective rules for guiding human action, the constructive step to objectively valid laws cannot succeed and, consequently, laws will be unable to influence how the agent is to determine his or her will. A sphere of application must be first secured so that the formal and explicitly moral law can be at all articulated at the level of the will. The act of willing is always directed toward something. Even in the case where I include pure reason itself in the grounds for determining my will, any individual act of willing must nevertheless have a basis in something concrete. This is precisely what maxims have to offer, while the categorical imperative will insure their universality.

Conversely, to avoid reducing our understanding of action to isolated individual acts, the sphere of application must be structured in advance so that it will lend itself to laws. Of course, this cannot first take place on the level of objectively valid norms. For it is just the initial anchoring of

maxims in the subject that paves the way for the will to be determined by laws in the first place. Faced with such complex conditions for proof, Kant simply presupposed maxims as given. In all of this, we can detect a legacy handed down from the European *prudentia* tradition, a last trace of Aristotelianism that critical philosophy never developed or made the subject of further reflection.¹ That this serves as the cornerstone for Kant's moral philosophy, despite the clearly rigorous nature of its underlying intentions, did not really seem to trouble Kant studies for a long time.²

To shed light on the fact that Kant's ideas about maxims were informed by the tradition, it will help to look at a text that lies somewhat off the beaten track. Kant's well-known thoughts concerning maxims are obviously to be found in his major works on moral philosophy and the philosophy of right. But in the authorized edition of lectures, *Über Pädagogik*, which figured among the Königsberg professor's required duties and, along with his other lectures, were based on teaching manuals widely used at the time,³ there is the following observation: "Moral culture must be founded on maxims, not discipline. The former prevent bad habits, the latter forms mental attitudes. One must see to it that a child becomes accustomed to acting in accordance with maxims and not according to certain drives. . . . A child should learn to act according to maxims whose appropriateness it will understand for itself." And further, "Character consists in the skill of acting according to maxims."⁴

These brief remarks are instructive. Rather than relying on specialized philosophical terminology, they are worded in the everyday language of generally accepted opinions.⁵ To be sure, these reflections by no means fit into the specific framework of Kant's project of providing moral philosophy with a systematic foundation. Rather, they mirror the thinking developed within traditional theories concerning the art of living transposed to the perspective of socialization and integrated into the beliefs of the enlightened times. What Kant is calling for is that children's unfettered physical inclinations be reshaped through training into behavior capable of conforming to rules. As we cannot expect children to create such rules on their own, the rules must be established in advance within the living context of social situations. The ability, however, to discern the appropriateness of individual maxims, i.e., their appropriate application within praxis' ever-changing coordinates, must be engendered in the pupil. In terms of Aristotle's understanding of *hexis*, the molding of character will ultimately stabilize action so that it will become second nature for

the individual to act consistently and conform to rules in the face of unpredictable, changing circumstances.

Moral culture, important to the Enlightenment, grew from a fundamen-
tally of subjectively accepted maxims that supported the social structure
like a net. This retrospective glance at elements of the tradition, which
Kant borrowed and incorporated into his formulation of the categorical
imperative under the rubric of maxims, should suffice for introducing the
matter at hand. My intention here is not to break new ground, but rather
to locate the point where Kant drew upon conventional ideas that did not
explicitly originate from his own moral theory. Of course, other well-
known aspects of his teachings, such as the categorical imperative, the
fact of reason, respect for the law, and so on, also build upon previous
knowledge. They are, however, conceptual tools Kant especially crafted
to meet the demands of his proofs.

No one who is familiar with all the relevant historical and cultural
dimensions would want to maintain Kant was particularly original in his
treatment of maxims. By the same token, it would be an exaggeration to
claim Kant gave or meant to give a deeper, more specific account of max-
ims than what was common knowledge at the time. To say Kant intended
to write (or, in fact, was the originator of) a rigorous and coherent theory
of maxims flies in the face of any modest survey of the scattered passages
that must be collected together before they begin to yield any insight at
all. And if we want to reconstruct the real intentions motivating Kant's
mode of proof, it would be risky to maintain the thesis that maxims play
a central role in understanding his ethics. But at least this thesis brings a
little fresh air into otherwise pedantic exegeses.

II

Although, for a long time, Kant studies did not show much interest in
maxims, this rather neglected concept nevertheless became the object of
increased attention.⁶ Two reasons can be given to account for the renewed
interest. First, the boom in popularity enjoyed by the rule concept in the
wake of Wittgenstein's later work inspired a large number of studies ad-
vancing along similar lines within linguistic philosophy and social phi-
losophy. That the basic structure of our social behavior could now be
explained in terms of rule competence also affected Kant exegesis. For
the concept of maxim is always implied in the concept of rule. The sec-

and reason to keep in mind is connected to how ideas inspired by Aristotle's ethics began to reemerge in more recent discussions. But more about this later. For now, let us look at the first point.

Wittgenstein's turn to linguistic pragmatics freed logical analysis from being restricted to constative sentences (assertions, propositions) and, stopping short of all normatively legitimated postulates for action, he discovered within language use itself a complex structure of assumptions and commitments without which consistent participation in language games as such would not be possible. The crucial intuition consists in adequately understanding what it means to follow rules that would not be understood as such if they were only written down or simply stated and did not have certain consequences for the actions in question. If the ability to follow rules is to serve as the basis for mastering a language, then, in the final analysis, this ability must consist in our knowing how to act in a given situation.

Wittgenstein aptly describes acting according to rules as the ability to go on acting in the same way. This description gives rise to the following chain of thought. Without linguistic competence we would never be able to take part in any form of discourse whatsoever. In mastering a language, we simultaneously move directly from knowing to acting by being prepared to recognize publically accepted rules that organize our linguistic transactions. To participate in a language community in the sense of recognizing rules, however, implies we are, in fact, fulfilling expectations other speakers automatically have of us when we engage in any form of language game with them. Corresponding to these expectations are also commitments to future linguistic behavior that limit any arbitrary use of language and provide the necessary continuity for our discourse.

These expectations and commitments, however, are not in any way "morally" qualified. We must obviously look beyond the field marked out by linguistic pragmatism, if we are to determine what makes rules moral. The act of promising furnishes the classical example of where the line is to be drawn separating linguistic from moral considerations. When we make a promise, we must be aware that at the very moment of promising, we have committed ourselves to a definite mode of future behavior that we are now prepared to make good on. Instead, if we lie or dissemble, a state of affairs arises that clearly cannot be explained by analyzing the language involved. Although when we deceive others we often rely on linguistic means, deception itself can hardly be said to be an abuse of language in the same way as faulty grammar, incorrect English, or poor pronunciation.

In itself, it is trivial and certainly does not constitute moral conduct when we establish a connection between what we say in the present and future deeds, which is exactly what we do in making a promise. As is obvious, a promise is subject to many factors whose precise influence is sometimes difficult to anticipate. No one has control over the future. Between the moment a promise is made and its realization, many things can intervene along the way that make former assurances appear in a different light. Realistically, we should allow for a certain latitude in ways of playing the game of making promises. Cultural difference must also be taken into account. In some cultures, the act of promising has less meaning than in others and can be understood as something akin to a heartfelt expression of sympathy rather than a hard and fast commitment. Finally, there is clearly not any way to impose sanctions that would insure promises will, in fact, be carried out once they are made.

The widely discussed example of promising demonstrates, as it were, a value-neutral basis for a linguistic pragmatics as carried out by Wittgenstein and his followers. Thus, if rules are to be brought into proximity to those maxims able to satisfy the demands of Kantian procedure, they must be morally qualified. In his much acclaimed book, *Generalization in Ethics*, M.G. Singer gives evidence of just such a trend.⁷ With regard to his main reference to Kant's ethics, Singer discusses the problem of rules, maxims, and moral principles in such a way that the pertinent maxims have already assumed moral qualities.⁸ Thus, all that remains is to test for the maxims' capacity for generalization according to the well-known consideration: "If everyone were to act in the same way . . ." Obviously, due to the goal Singer sets for his argument, he is forced to weigh maxims in such a way that they yield the necessary moral content⁹ whose ultimate guarantee, "usually right or wrong," can then be entrusted to the logical procedure of generalization.

In contrast, Kant himself assumes rules are made with the "assistance of reason"¹⁰ and, thus, do not merely represent subjective inclinations. It remains unclear, however, whether we are to understand reason's influence to be at work in establishing regularity per se, and in stabilizing a uniform orientation for our actions even when we are at the mercy of changing circumstances. For not all rules will be recognized as candidates for governing moral behavior.

On the other hand, if reason's contribution is not at the structural level of rules, but rather at the moment maxims are first invested with moral content, then the acid test for generalization gives its final stamp of ap-

proval by according to the whole process only the force of transsubjective, valid laws. In this case, maxims would always be regarded as test candidates, while their ability to function as rules would appear neutralized or would simply be assumed.

On this account, it remains difficult to explain how those maxims would fare that exhibit recognizable, moral qualifications before being tested for their adaptability for universal legislation, but after the test must be abandoned. What aspect of maxims would serve initially as the basis for regarding them as "moral"? No matter how we decide to settle this question, it must be argued against Singer that morality is not to be found in a rough estimate of general consequences ("If everyone were to act in the same way"), or in commonly held ideas about norms ("What is usually right"). In the final analysis, both considerations turn out to be empirical. Kant's ethics, however, clearly stand or fall with bringing the will, purged of all empirical motives and considerations, under the control of reason as such.¹¹

Lewis White Beck wanted to incorporate maxims with primary moral qualities into a type of practical syllogism that has its roots in Aristotle's ethics. In connection to a passage from the *Critique of Practical Reason* (5:162; 184), Beck, in his masterly commentary, followed this course of explanation,¹² though Kant in this same passage only discusses the "moral principle," instead of directly addressing maxims. The moral principle appears in the major premise. The minor premise subsumes a possible action under the major premise, and the conclusion, according to Kant in this passage, consists in the "subjective determination of the will." In parentheses Kant explains this as "an interest in the possible good, and in the maxim founded on it."

Surprisingly, Beck treats the "moral principle" in the major premise as equivalent to "maxim." Although this may be plausible when considering the "explanation [*Erklärung*] of practical principles" given in section 1 of the second *Critique*,¹³ it raises difficulties for the syllogism's conclusion which Beck avoids by summarily inserting "decision" where Kant, in fact, speaks of the determination of the will as an interest in the good and the maxim founded upon it. According to Kant's wording, the maxim as subjective orientation is a consequence of the syllogism. It cannot, therefore, have already appeared in the major premise. Not to mention that for Aristotle, to the continual astonishment of logicians, practical syllogisms terminate in concrete action itself and not merely in dispositions to act.¹⁴ Even less do practical syllogisms result in propositions, rules, maxims, and the like.

This attempt to model maxims after practical syllogisms has been repeated numerous times, for example, by Aune,¹⁵ and Atwell.¹⁶ These and other subsequent attempts, however, have not proved to be especially convincing. The reason for this can be traced back to how maxims were treated within the moralist literature, which most likely served as Kant's source of inspiration. Here, rules were not derived from any form of logical deduction, but rather distilled from densely woven aphoristic sayings rooted in worldly wisdom acquired through observation and reflection. This is mirrored in the title, "Maxims and Reflections," which headed most of the standard texts where maxims were collected together, divided, and categorized according to various areas of life.

Recent accounts, however, have stressed the role subjective intention plays in the consideration of rules. The trend seems to date our initial remarks concerning the renewed popularity of Wittgenstein's approach to rules. Instead, an interpretation of rules in terms of hierarchy has taken center stage. For example, Onora O'Neill has written that "maxims are those underlying principles or intentions by which we guide and control our more specific intentions."¹⁷ Christine Korsgaard is even more explicit: "A maxim of action will . . . usually have the form 'I will do Action-A in order to achieve Purpose P. You will only act on that maxim if you also make it your maxim to achieve Purpose P'. . . . Although Kant does not emphasize this, it is perhaps easiest to think of maxims as hierarchically organized."¹⁸ These efforts, and others like them, lead us far afield from Kant's original intentions.¹⁹

In his astute study, M. Willaschek gives maxims more room than Kant intended. Thus, for Willaschek, maxims tend to take the place of hypothetical imperatives that, in turn, are relegated to a short appendix. It is interesting to note, however, that Kant gives more constructive attention to hypothetical imperatives than to maxims. According to Willaschek, the representations of ends constitute the causes of actions, because they establish ends "while maxims determine the course (the choice of means)."²⁰ However much this formulation may have been taken over from Aristotle's notion of *prohairesis*, without a great deal of qualification it makes little sense in terms of Kant's project, as can be shown by some (but not all) of his genuine examples.

Let us take a look at two of these examples. Those who make it their "maxim to let no injury pass unrevenged" are fundamentally oriented toward addressing issues of honor (*Practical Reason*, 5:36; 105) and stand in contrast to anyone who always turns the other cheek when offended.

The maxim, however, has absolutely nothing to say about the means (dual, slander, sabotaging someone's career, etc.) to be employed for exacting revenge. Without considering the matter on a case-by-case basis, there is not any way to determine in advance the means to the end of seeking revenge. Similarly, whoever follows the maxim, "neglect your natural gifts" (*Groundwork*, 5:55; 40), and devotes his life "to idleness, amusement, and . . . enjoyment," is not employing, as a result of his self-indulgent disposition, specific means to put this maxim into practice. Rather, anything that suits him plays into the neglect of his talents. To endow maxims with means-ends calculations is to underestimate the rule-character of maxims.

To be sure, maxims involve intentions, as Willaschek rightly emphasizes. But no "double function" arises whereby a maxim provides the content for universalization, while a second, higher level maxim has as its object the act of willing that corresponds only to those first-level maxims capable of becoming universal laws.²¹ Rather, maxims capable of universalization coincide identically with what all rational beings would will for themselves and, therefore, have the reinforced, binding character of universal laws.²² In short, Korsgaard's problem of a hierarchical ordering of metamaxims underpinning morality does not exist, and an analysis of the role intentionality plays in the morally good will should not obscure the role of rules. By dint of pure reason, rules are elevated directly from the givenness of maxims to the status of laws and, thus, acquire objective, binding character.

III

No one would want to maintain Kant's thoughts about maxims are clear cut or can be easily summarized. When, every now and then, one of the relevant texts seems to give such an impression, it is due to the reader's overwhelming need to simplify matters. One school of thought, obviously at odds with the interpretations already mentioned, sees in maxims miniature blueprints for how we want to conduct our lives. This is another example where recent discussions converge on an aspect of the tradition reaching back to Aristotle. R. Bittner envisions maxims to be expressions of a "subjective notion of the good life."²³ According to Kant, a maxim has "under it several practical rules" (*Practical Reason*, 5:35; 105). In

comparison to the other remarks about maxims Kant makes throughout his works, the meaning of this statement is not particularly clear. For Bittner, the picture of maxims comes together in this way: "As the projection of an entire life, maxims comprise many different actions," which are then obviously guided by rules subordinate to them. To my way of thinking, in reviving the notion of *eudaimonia* and presenting it in the distilled form of maxims, Bittner is headed in the right direction, but he overshoots the mark in one essential way.²⁴

In a book cited at the beginning of this essay, I also argue for considering the structural core of maxims as a worn-down remnant of the *prudential* doctrine that applied to our everyday dealings with the world and was handed down by way of the moralist tradition. But it is hardly plausible that a complete blueprint for our lives could find adequate expression in any one maxim.²⁵ Maxims are always directed toward particular life situations and not to life in general. Similar to aphorisms, the way in which maxims are formulated does not lend itself to fixed interpretation. How they are applied and what consequences are to be drawn from them are dependent on arbitrary circumstances and differ from individual to individual. Besides, considering how many maxims we are acquainted with from the literature and are also able to formulate from our own on-going life experiences, there certainly cannot be a corresponding number of life projects. And furthermore, the genre of literature specifically dealing with maxims is situated at a historical juncture between the rise of enlightened bourgeois society, and the decline of the aristocracy, and brings with it the memory of humanist, aphoristic wisdom. This historical moment is far too complex to provide us only with a few plausible ways of life.

Let us again look at some of the familiar, if scarce, concrete examples of maxims Kant gives. In the famous case of the deposit,²⁶ where the guiding principle is to increase one's "fortune by every safe means," it is possible to make out the image of the upright, honest businessman. In the case of someone who does not let any offence go unrevengeed (*Practical Reason*, 5:35; 105), we can decipher a paragon of a honor-conscious aristocracy. Whoever, in pandering to his pleasure-loving nature, neglects his natural gifts would reveal himself to be a hedonist "like the inhabitants of the South Sea Islands" (*Groundwork*, 4:55; 40–1). To a certain degree, even the most banal maxims lend themselves to this sort of reconstruction. Once I enlisted the fictitious example of a maxim which said, "I always get up early in the morning."²⁷ Even this modest attempt to high-

light and preserve a particular detail of my daily existence could be also interpreted in terms of a more general notion of a healthy and productive daily routine emblematic of a whole range of bourgeois virtues.

O. Höffe writes: "Maxims encapsulate the way life as a whole is to be led regarding certain fundamental aspects and general types of situations encountered in life."²⁸ Clearly, here Höffe is endorsing Bittner, and generally those involved in the "rehabilitation of practical philosophy," as this long, drawn-out project was once called. The thrust of the attempts to breathe new life into practical philosophy was to reconcile the abstract universalism implicit in Kantian thinking about the moral law with ideas about how the concretion of practical reason is tied to the particularity of individual situations. This ongoing project can only be wholeheartedly applauded. That the files have not been closed on this subject is demonstrated by the presence of communitarians, such as MacIntyre, Walzer, Taylor, etc., and the influence they and others have exerted in recent years on the international scene. The communitarians argue against the way in which Kantian thought has been reduced to a procedural formalism and which, under the label of liberalism, Rawls has made common currency in those principal lands most rooted in utilitarian beliefs.

Nonetheless, a fundamental objection still needs to be raised. If we understand the good life as a complete picture of an individual's life, a life carried out coherently and at every turn guided by reason, and if the individual acts and at every new juncture continues to act in similar ways, then we can say that this particular life is reflected as a macrocosm in the microcosm of individual maxims. But inasmuch as every individual maxim incorporates references to explicitly disparate situations (aspects of character and life situations, for example, property, social standing, reputation, pleasures, honesty, and so forth), then, of course, when compared to life taken as a whole, one maxim's sphere of reference must necessarily describe a narrow area of life. If this were not so, then maxims would forfeit their illustrative character that makes them applicable to our particular actions and easy to adopt as our own. As the French moralist writers knew full well, maxims must be formulated in such a way as to deliver their message at one stroke without provoking long-winded arguments. To see a universal in the particular, to infer character from a rule, implies for every maxim the possibility of hermeneutical application. This is where judgment is called for. At the level of maxims, there is not any way to avoid being involved with particularities that everyone knows and has clearly in view. Otherwise, maxims threaten to degenerate

into empty platitudes such as, “Do good things!” or “Prove your worth as a human being!”

Put another way, the bridge connecting the large-scale plans we make for our lives to the individual maxims we adopt to guide our everyday actions is an erratic arc reflecting gaps, uncertainties, and vacillations. It would probably overtax any form of logical deduction to derive individual maxims from the long-term goals we set for ourselves. As has often been pointed out, with regard to Aristotle’s classical typology of life forms (personal hedonism, political ambition, philosophical desire for knowledge), the ancient philosopher was content to provide only rough sketches.²⁹ Filling in what these sketches leave out would be a Herculean task demanding from us all the effort necessary to come up with new interpretations on a day-to-day basis. As Höffe is right to stress, unfortunately Kant “neither presented or sufficiently clarified” what a fully articulated ethics based solely on maxims would be, one which would have eliminated the necessity of interpretation in the first place.³⁰

IV

In conclusion, I would like to put forward a thesis concerning the results of my argumentation. The relevant aspects of the tradition play an undervalued role in the secondary literature. The discussions surrounding maxims usually give the impression that those evoking maxims are also the ones who conceive them. According to this understanding, it is within our power as free, self-aware subjects, who are not slaves to custom but are able to negotiate the permanent tension between custom and the moral law, to establish, cultivate, and answer for our own maxims.³¹ If this were only so. In fact, to a large extent, socialization and culture have taken over this task for us. This is not a blemish on, or insult to, our rational capacities. It is simply how things have evolved.

The French writers’ “Maxims and Reflections,” later echoed under the same title in Goethe’s works, were thought-provoking distillations of worldly wisdom.³² Intended for immediate reception, they were frequently formulated in an anecdotal or gnomic style and were much read and widely imitated at the time. Gathered together in collections, the maxims mirror everyday situations. And by offering advice, shedding light on familiar dilemmas, voicing doubts about the worth of generally accepted conventions, these collections were intended to widen the horizons of the read-

ers' usual practices by augmenting their life experiences and deepening their knowledge of the world. The confidence of these classical writers and teachers that their contemporary as well as their future readers would identify with the collected observations, or at least feel encouraged to contradict them or come up with their own variations, presupposed both a traditional society, as well as entrenched popular beliefs about human nature. Thus, although later writers, such as Chamfort and Rivavol, fell under the sway of the revolution, taken together the body of work comprising the maxim literature remains a legacy of prerevolutionary times when belief in change, accelerated progress, the overthrow of the old order, equality, did not yet constitute the coordinates for the new concept of humanity, which lay in the not-so-distant future. Maxims offered society one way to preserve and cultivate specific forms of interactions down to the finest nuances and, through the medium of literature, they were communicated to young people for the sake of their individual orientations. Remember Kant's remarks, cited at the outset, where he reflects about the role of maxims in establishing modes of acceptable behavior in children. Since we are unable to know everything or live through every situation and experience every passion, the balance of our maxims cannot be a product of our own intellectual labor. Rather, we acquire them through knowledge and experience, and then hone them down to fit our individual circumstances. This does not at all alter the status of maxims. They remain rules I specifically chose for myself and for which I alone (and nobody else) am responsible. Nevertheless, my horizon is extended beyond everything idiosyncratic and is opened up to include the wealth of understanding shored up in the worlds of human action I share with others.

Nicholas Chamfort opens his *Maxims and Reflections* (1795) with the following explanation:

Maxims and axioms, like compendiums, are the work of brilliant people [*gens d'esprit*], who, so it seems, have done the work for mediocre and idle minds. The idle person adopts a maxim in order to spare himself the observations which led the author to his result. Idle and mediocre people do not dare go beyond the maxim, and they accord the maxim a universality which the author, unless he himself was mediocre, did not want at all to give it. A superior intellect grasps all at once the similarities and differences which make a maxim more or less, or not at all, applicable to this or that case.

It need hardly be stressed that to reestablish the connection between the concept of maxim and classical moral thought is not meant to be a painstaking proof of the “influence” other authors may have exerted on Kant. By the same token, to focus in on historical filiations, will not revolutionize interpretations of the handful of statements Kant made with regard to maxims. It is important, however, to target that form of narrow-minded dogmatism that tends to portray Kant as an oracle who has an answer for every question, who is the most definitive interpreter of his own work, and who is a self-enclosed continent of theory building that tolerates no contact with other sources of insight.³³

Notes

1. For a sketch of the history of the maxim concept, which found its way from its original, logical context (*propositiones maximae*) to modern moralist thought (Gracian, La Rochefoucauld, La Bruyère, and others), see my book *Handlung, Sprache und Vernunft* (Frankfurt/M.: Suhrkamp, 1976, 2nd ed. 1982), 196 ff.

2. I also deal with this in the above-cited book, 185 ff.

3. See the Editor's Preface (Rink) of 1803, Kant, *Werke* (ed. Weischedel), vol. 6, 695.

4. *Ibid.*, 740 f.

5. For information about those “maxims” Kant himself followed with regard to his own life conduct, see L.E. Borowski and R.B. Jachmann, in *I. Kant: Sein Leben in Darstellungen von Zeitgenossen*, ed. F. Gross (Darmstadt: Wissensch. BG., 1968), 51 ff., 149 ff.

6. See U. Thurnherr's informative dissertation, *Die Ästhetik der Existenz: Über den Begriff der Maxime und die Bildung von Maximen bei Kant* (Tübingen: Francke, 1994).

7. *Generalization in Ethics: An Essay in the Logic of Ethics with the Rudiments of a System of Moral Philosophy* (New York: Alfred A. Knopf, 1961, 2nd ed., with new Foreword 1971).

8. See the examples on pp. 10, 98 ff., and 292.

9. To name one source, R. Wimmer is right to criticize this. See *Universalisierung in der Ethik* (Frankfurt/M.: Suhrkamp, 1980), 234.

10. See *Groundwork*, 4:62; 44–5, also 4:51; 38; *Practical Reason*, 5:36; 105. The last number gives the references to T.K. Abbott's translation of the *Groundwork* and the second *Critique*, published as *Kant's Critique of Practical Reason and Other Works on the Theory of Ethics*, 6th ed. (London: Longmans Green, 1909).

11. See *Practical Reason*, 5:38; 107: "But that reason may give laws it is necessary that it should only need to presuppose itself, because rules are objectively and universally valid only when they hold without any contingent subjective conditions, which distinguish one rational being from another."

12. L.W. Beck, *A Commentary on Kant's "Critique of Practical Reason"* (Chicago: University of Chicago Press, 1960), 81.

13. The use of the word "explanation" is interesting. It seems it was consciously chosen over "definition," which is what the reader would expect here. Definitions impose arbitrary boundaries, while explanations clarify or elucidate previous knowledge.

14. The few passages where this is directly addressed are scattered. See, for example, *Nicomachean Ethics*, 1147a1 ff., *De motu animalium*, 701a12, and *Metaphysics*, 1032b6 ff.

15. B. Aune, *Kant's Theory of Morals* (Princeton: Princeton University Press, 1979), 12 f.

16. J.E. Atwell, *Ends and Principles in Kant's Moral Thought* (Dordrecht: Martinus Nijhoff, 1986), 45.

17. O. O'Neill, *Constructions of Reason: Explorations of Kant's Practical Philosophy* (Cambridge: Cambridge University Press, 1989), 84. In a later footnote, O'Neill revises the first use of "intentions."

18. C.M. Korsgaard, *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996), 57 f.

19. In a recent review, "Maximen, Identität und praktische Deliberation" (*Philosophische Rundschau* 45:1998, 74), H. Pauer-Stader praises Korsgaard, who teaches at Harvard, for attempting to take the sting out of Kant's "ought" with the help of a "maxim ethic." It is curious that, in the age of globalization, academic mobility, and ready translations, we still know so little about each other that what has been discussed for a long time in this country appears to be a praiseworthy innovation when seen from abroad.

20. M. Willaschek, *Praktische Vernunft, Handlungstheorie und Moralbegründung bei Kant* (Stuttgart: Metzler, 1992), 76.

21. *Ibid.*, 68.

22. For example, *Groundwork*, 4:76; 52: "that the will could at the same time regard itself as giving in its maxims universal laws." See also *Groundwork*, 4:81, 102; 55, 67, and *passim*.

23. R. Bittner, "Maximen," in *Akten des IV. Internationalen Kant-Kongresses*, ed. G. Funke (Berlin: Walter de Gruyter, 1975), 485–98, here, 489 f.

24. For pertinent objections, see H. Köhl, *Kants Gesinnungsethik* (Berlin: Walter de Gruyter, 1990), 52 f.

25. Nevertheless, in his above mentioned *Commentary*, 71, Beck did open up the way for interpreting a person's maxims as "expressing his own policy in life."

26. *Practical Reason*, 5:49; 115. See also: Cicero, *De finibus* III; Hume, *Treatise* III, 2,1 (ed. Selby-Bigge), 479 f.
27. R. Bubner, *op. cit.*, 189.
28. O. Höffe, "Kants kategorischer Imperativ als Kriterium des Sittlichen," in Höffe, *Ethik und Politik* (Frankfurt/M.: Suhrkamp Verlag, 1979), 84–119, here, 90 ff.
29. See *Nicomachean Ethics*, 1095b17 ff., 1094b20.
30. O. Höffe, *Kant* (Munich: Verlag C.H. Beck, 1983), 187.
31. With regard to recent developments in the above-mentioned status of German research, see M. Albrecht, "Kants Maximenethik and ihre Begründung," *Kant-Studien* 85:1994, 129–46.
32. *Französische Moralisten* (La Rochefoucauld, Vauvenargues, Montesquieu, Chamfort), ed. and intro. F. Schalk (Zurich: Diogenes, 1995).
33. Translated by Cara Gendel Ryan and Michael McGettigan.

Kant versus Eudaimonism

Allen W. Wood

Kant was among the first to break decisively with the eudaimonistic tradition of classical ethics by declaring that the moral principle is entirely distinct and divergent from the principle of happiness (*Groundwork*, 4:393; *Practical Reason*, 5:21–7).¹ I am going to argue that what is at issue in Kant's rejection of eudaimonism is not fundamentally any question of ethical value or the priority among values. On the contrary, as I shall argue, on these matters Kant shares the views that led classical ethical theory from Socrates onward to embrace eudaimonism. Instead, where Kant breaks with classical ethics is in the conception of *human nature*. Kant's conception of human nature so altered the application of moral principles that it forced a change in the way happiness was conceived, leading to a reversal of what had earlier been thought about the relation of the principle morality to the pursuit of happiness.

I. Classical Eudaimonism

The classical ethical theories of Plato, Aristotle, and the Stoics took it for granted that happiness (*eudaimonia*) is the encompassing human good, the final end of human life. As such, happiness (or at least its dominant component) is identified by the classical theories with either the possession or the exercise of moral virtue. A separate strand in the Western tradition, deriving from Eudoxus and Epicurus, identified happiness with pleasure; but it, too, took happiness to be the encompassing human good, and sought to find a place for the moral good within eudaimonism. It therefore identified moral virtue as the indispensable *means* of achieving happiness. Christian philosophers tended on the whole to follow the classical theories, especially those of Aristotle and the Stoics, but the Epicurean tradition made a strong comeback in the early modern period. The identification of happiness with pleasure (either in general or in its highest forms) is found in many moderns, among them Malebranche, Arnauld, Locke, Shaftesbury, Bentham, and the

French *lumières*. Since Kant also identifies happiness with subjective satisfaction, and sometimes even with pleasure, he too belongs to this modern Epicurean tradition concerning what happiness consists in.

The classical tradition in ethics assumes that the human good (under the name of “happiness”) can be represented as a single object of desire. Behind this assumption is an idea Kant shares, and would express by saying that reason seeks maximal unity under principles. The idea is that whatever desires, interests, and aims we may have, we can sort out our priorities among them rationally, and the result will be a single end or good that we may represent as our final aim. This rationale is quite explicit, for example, in the opening sentences of Aristotle’s *Nicomachean Ethics*, where it is argued that all arts and goal-directed activities can be comprehended in a single system directed at a single final end.²

Like Kant, classical ethics regards morality as having priority over other goods. Hence classical theories typically identify either happiness itself or its dominant component with either the possession or exercise of moral virtue. Suppose I find myself in a situation where I can increase my wealth, but only by performing a base action, of which I would be ashamed. Since the classical theories hold that avoiding shame takes priority over gaining wealth, they recommend that I abstain from the base action. They express the point by saying that to gain wealth in this way will not, all things considered, serve my interests; the honorable course of conduct is the one most conducive to my happiness. Because classical ethics supposes moral self-worth can be treated as part of the same system of evaluation as other goods, they treat it as an object that can be weighed against them—and prevail against them—in calculating what is best for us on the whole. This is what Socrates does, for example, when he decides that you are better off suffering injustice than doing it.³

II. Kant versus Socrates

Kant parts company with this tradition because he denies that we can integrate moral self-worth into our picture of the good in this straightforward way. He distinguishes the worth of one’s *state or condition* (*Zustand*) from the worth of one’s *person* (*Practical Reason*, 5:60). Our state includes not only pleasure or pain, but also everything external to us that can cause them, and even those aspects of ourselves that are independent of our actions, such as our temperament or natural constitution (*Ground-*

work, 4:398; *Anthropology*, 7:285–92). Our person, by contrast, is the object of moral imputation (*Morals*, 6:223); it includes our actions and our character, both of which depend on and display our good or evil will (*Groundwork*, 4:392–4; *Anthropology*, 7:285, 291–5). In a parallel way, Kant distinguishes the “moral” good sharply from the “natural” good as two entirely heterogeneous goods (*Practical Reason*, 5:59–60; *Anthropology*, 7:276–8). Kant wants to consider the “contentment” people feel when they are conscious of having done their duty as entirely distinct from their happiness, forming not the smallest part of it (*Morals*, 6:387; *Practical Reason*, 5:38, 88, 117, 156). He criticizes the ancients for identifying the two kinds of goods, and the moderns for confusing them (*Practical Reason*, 5:64–5).

The principal example of such a confusion is the rationalist conception of “moral happiness,” which is “satisfaction with one’s person and one’s own moral conduct,” or the contentment you feel when you are aware of having done your duty (*Morals*, 6:387–8). Kant appears to countenance this notion only in one context: the moral-theological context in which the issue is whether well-disposed moral agents might require divine aid in order to insure their continued moral progress (*Religion*, 6:67–8). But he regularly attacks the entire conception as a “misuse of the term ‘happiness’” (*Morals*, 6:387–8) and even as self-contradictory, since it implies that the motive for doing one’s duty is to achieve this sort of happiness, whereas self-contentment with one’s person depends on doing our duty for its own sake, and not in order to achieve any happiness by means of it (*Morals*, 6:377–8).

It is not that Kant considers the moral and natural goods incommensurable in the sense that there can be no ranking among them. In fact, he insists on a very definite ranking between them: the moral good is prior to the natural and constitutes the condition for its worth (*Groundwork*, 4:393). This is how they are combined in the idea of the “highest good” (*summum bonum*), the highest end of all rational and moral striving (*Practical Reason*, 5:110–13). The fact that they can be so combined indicates that there is nothing inherently incommensurable about the moral and natural goods from the standpoint of reason. So in this sense, Kant appears to agree with classical eudaimonism in holding that reason can combine all goods into a single final end.

It is in another sense that moral and natural good are incommensurable for Kant. In relation to our nature, they belong to entirely different (and incommensurable) systems of evaluation. They fall under different prin-

ciples because we value them from two entirely different standpoints—both of which, however, we inevitably adopt. For classical ethical theory, by contrast, rational deliberation ultimately takes place from a single standpoint, because however many “parts” the soul may have (three in Plato’s theory, for instance), there is only one *real* self whose interests concern us, and that is the rational self.

Kant agrees, of course, that our real self is the rational self. But his theory reflects a distinctively modern conception of humanity, because he recognizes that the human self is inevitably disunited, conflicted, self-alienated, in a deeper sense than any of the ancient theories could admit.

Kant sometimes expresses this difference metaphysically, by contrasting the standpoint of nature or the empirical self with that of the free or noumenal self (*Groundwork*, 4:457–8). But this formulation is misleading insofar as it gives the impression that Kant’s conception of human nature is grounded on his metaphysics. Nor can the difference in standpoints for Kant be merely a difference between natural desires and rational principles. For that difference was already part of classical eudaimonism. Least of all is there any disagreement over the question whether the standpoint of reason is *superior* to that of nature. On the contrary, that is precisely the reason why the classical theories say that, since moral virtue is unconditionally prior in value to natural pleasures, we should always give it first preference in calculating our happiness.

When Socrates assures us that we are better off (or happier) suffering injustice than doing it, we cannot help feeling that something is wrong with this. Socrates’ contemporaries obviously also found his thesis incredible. It was doubtless intended as a paradox—a philosopher’s rejection of the standpoint of common sense. The “commonsense” view here seems to be that suffering injustice is *morally preferable* to doing it, but that what is morally preferable might not be the thing that is best for me to do from the standpoint of my self-interest or happiness. This is the view of Polus, for instance, who thinks it is “finer” or “more admirable” to suffer injustice than to do it, but that the doer of injustice is nevertheless happier or better off than the sufferer.⁴ More recent moral philosophers may express Polus’ opinion by saying that *suffering* injustice might be better from “the moral point of view”; but from *my* point of view (the point of view of my happiness or self-interest), it would be better to *do* injustice (if I can do it with impunity).

The commonsense position naively presupposes that the “moral point of view” must be different from my own point of view. But this naturally in-

vites the question: "If it isn't in my *interest* to take the moral point of view, why should I take it at all?" It naturally leaves us with the suspicion that the moral point of view may be nothing but a matter of conventional approval or disapproval, which, if I am sensible, I will ignore whenever it interferes with my self-interest. That, of course, is the position of Thrasymachus, who holds that such conventions are made by the strong in their best interests; a sensible man will do injustice whenever he can get away with it.⁵

Thrasymachus' version of moral skepticism represents an essential advance over moral common sense, because it faces up to the consequences of the commonsense thought that my point of view is different from the moral point of view. Socrates' view represents an even more essential advance over common sense, however, by recognizing that there can be good reasons for taking the "moral point of view" as *my* point of view—even more truly my point of view than the standpoint which either Polus' commonsense or Thrasymachus' moral skepticism represents as mine. Socrates, like Kant, takes the standpoint of moral reason to be more truly my point of view than the standpoint of maximizing wealth or pleasure. Thus, Socrates decides that the shamefulness involved in doing injustice is a greater threat to my self-interest than the disadvantages incurred in suffering injustice.

Since Kant agrees with Socrates rather than Polus here, his reason for feeling uneasy with Socrates' position is different from the reasons of common sense. Kant holds that although Socrates is right that the moral point of view is also most truly my point of view, the human self also necessarily has a divided standpoint, the standpoint of reason and also a standpoint that is opposed to reason. But Kant's new point of departure cannot be adequately represented in terms of the soul's different "parts," tugging in different directions (which was already included in Socrates' position).⁶ Even worse, perhaps, would be the image of a tug of war between two "selves" (especially if they are thought of as denizens of entirely different worlds). For all conflicts of that kind could easily be reconciled with classical eudaimonism, as Socrates does, by identifying my *real* self-interest with the interest of my *true* self (which, as Kant agrees, is the self that takes the moral point of view).

Kant's view is rather that the self is in conflict with itself because the *whole* self *simultaneously* takes two conflicting standpoints. One standpoint is the "natural-social" standpoint, expressed through my natural inclinations as they appear under conditions of social life. The other is the moral or rational standpoint, expressed through my lawgiving reason. These two standpoints cannot be resolved by preferring self-worth over

wealth or pleasure because they do not represent merely two objects of desire between which we might choose. Instead, they represent two *rival conceptions* of what is valuable for me, and especially two rival conceptions of my *self-worth*. The natural-social standpoint conceives self-worth in terms of the worth of one's state, and at the same time comparatively, as ambition or superiority over others. The pure rational or moral standpoint conceives self-worth absolutely and non-comparatively, as self-respect for the worth of humanity, as an end in itself and the dignity of self-legislating personality. To understand Kant's rejection of eudaimonism, we must see why he sees human nature as torn between these two standpoints and why he identifies the principle of one's own happiness with the natural-social standpoint rather than the pure rational standpoint.

III. Happiness as Pleasure

Kant identifies happiness with our well-being (*Wohl*) as natural creatures, or with imagined maximal satisfaction of our empirical inclinations. Of course it is an empirical, anthropological thesis that this constitutes a distinct kind of good that is in tension with our self-worth as rational beings. But Kant has more than one way of presenting the tension, and this leads him to formulate the notion of happiness in different ways which are not obviously reconcilable, and some of which obscure his real views.

At times, Kant wants to contrast happiness with morality by saying that it represents our "animal nature" or our "lower faculty of desire." Sometimes Kant's statements of this view are *purely* hedonistic, even strikingly Benthamite: Happiness consists of pleasure or agreeableness, and is measurable in terms of magnitude, duration, costliness, and fecundity (*Practical Reason*, 5:23–4). Closely akin (in Kant's mind) to the idea that happiness is pleasure is a second idea, that happiness consists in a state of *contentment*, together with the assurance that this state will last for the future (*Groundwork*, 4:393; *Practical Reason*, 5:25; *Morals*, 6:387). This seems to identify happiness with pleasure only in the negative sense that it excludes discontent; beyond this it adds an element of future assurance. A third formulation of Kant's hedonism is the identification of happiness with *desire-satisfaction*: "Happiness is the condition of a rational being in the world when everything goes according to its wish and will" (*Practical Reason*, 5:124).

Yet pleasure, contentment, and desire-satisfaction are not the same. We can be pleased without being contented, and contented without being

pleased. And getting what we wish and will may turn out neither to content us nor to please us. Still, there need be no great problem here as long as we are thinking about *what it would be like to be happy*. For then pleasure, contentment, and desire-satisfaction may be regarded as three different conditions that are jointly required for happiness. Desire-satisfaction without contentment, or pleasures that leave us extremely discontented, or a life that is contented (in a hopeless, resigned sort of way) but lacking in either pleasure or the satisfaction of important desires, do not make for happiness. From this point of view, we might consider Kant's different formulations of a fundamentally hedonistic conception of happiness as merely complementing one another, and as forming a single unified picture.

The picture changes significantly, however, if we consider the different descriptions of happiness in light of a thesis Kant often insists on: "All human beings naturally have the strongest and deepest inclination to happiness" (*Groundwork*, 4:399). Kant even maintains that all empirical desires "fall under the principle of self-love or one's own happiness" (*Practical Reason*, 5:22). When we think of happiness as something to be pursued, the question is not: What does it take to *be happy*? but rather: What object is it that we all have the deepest inclination for? What is it that we desire when we desire happiness?

These questions split apart the elements of the unified picture, because desiring one element of the picture is not the same as desiring another. The desire to be in a contented state is not the same as the desire for a particular object. A person can be content even if (or just because) she does not have any desires at all at the moment. For many desires, it is not at all obvious (and probably false) that our aim in satisfying the desire is merely to achieve the state of contentment we expect to feel when the desire is satisfied. Of course, when we do pursue happiness, we try to desire only what we think will content us if we are lucky enough to get it; and we think of happiness as consisting of pleasures whose enjoyment tends to content us. But the differences between pleasure, contentment, and desire-satisfaction raise questions about why we would want to constrain our ends in these ways by seeking happiness rather than simply the satisfaction of particular desires, including those for pleasure and contentment.

Perhaps it is tempting to think that if you are not content, then you must have some determinate desire that could be satisfied but is still unsatisfied. But that thought is mistaken, as we can see from Kant's own discussion of what he calls "moody wishes." A moody wish is one that

cannot be satisfied because one does not know exactly what it is that one wants (*Anthropology*, §73, 7:251). Kant may think that there is no great problem here, because his account of the empirical faculty of desire says that natural desires involve pleasure in the representation of their object, and an expected agreeableness if the object is achieved. Thus, when he identifies happiness with pleasure, Kant may be thinking that all desires aim at happiness merely in the sense that there is a natural connection between desire and pleasure in the representation of the object. Taken in that sense, the proposition that all desires aim at happiness would reduce to the triviality that our set of empirical desires (or, more accurately, pleasure-possibilities) is not empty. Every desire, it could be alleged, aims at its own satisfaction, which may be thought of as a state of pleasure or contentment. Thus, the desire for happiness is nothing but the generalized form of a desire involved necessarily in every inclination: desire that one should always be in such a state with the prospect of its lasting indefinitely. This line of thought appears to cohere with Kant's stated reason for holding that "all human beings have the strongest and deepest inclination to happiness." For this, reason is given as the following: "For in this idea all inclinations are summed up" (*Groundwork*, 4: 399).

Perhaps Kant thinks that since every desire aims at its own satisfaction, desire in general aims at satisfaction in general, or maximal satisfaction. But of course it is not obvious what "desire in general" could mean, unless it refers to some further desire over and above all my particular desires, a desire with a unique object constituted by a combination of their objects. But it is not self-evident that every human being must have such a desire. From the fact that I have three different desires, for *x*, *y*, and *z*, respectively, it certainly does not follow that I must have yet a fourth desire whose object is *x*, *y*, and *z*, taken together collectively. This is especially so when my desires threaten to conflict: For instance, if I have a desire to take a nap and also a desire to attend a philosophy lecture, then it does not follow that I have a desire to sleep through the philosophy lecture. Even if my desires do not conflict, it does not follow that I must have the further desire whose object is the sum of their various objects.

IV. Happiness as a Sum of Satisfaction

Kant acknowledges this when he speaks of happiness as an "idea" of a "sum" or "maximum" of pleasure or desire-satisfaction, which is fash-

ioned by our reason or imagination (*Groundwork*, 4:399, 418–19). For the fashioning of such an idea involves a positive act of mine, often the setting aside of some desires when they conflict with others. In order to come by such a desire, I must calculate which set of compossible satisfactions practically achievable by me constitutes the maximal sum that it is rational for me to pursue. Further, I must find some reason to pursue this new and artificial object, which is something over and above the objects of my original and basic desires (some of which have even been excluded from it).

Thus the apparently innocent step of introducing an “idea” of happiness actually involves Kant in a fundamentally new conception of happiness that is not at all the same as pleasure, contentment, or desire-satisfaction, and is quite incompatible with regarding the desire for happiness as arising merely from our animal nature. For the desire for happiness turns out to be a second-order desire, namely that a certain combination or subset of our first-order desires should be satisfied.

This makes it even clearer that it cannot be inferred simply from the fact that we have first order desires that we have a desire for happiness. It is one thing to say that each desire aims at a certain state, which counts as its own satisfaction. It would be quite another thing—which Kant would not want to say—that for every desire *d*, *d* must be accompanied by the second-order desire *d'* whose object is that *d* should be satisfied. Kant recognizes the existence of desires we positively do not want to satisfy, desires such that we have the second-order desire *not* to satisfy them.

Kant does want to claim, however, that all human beings have a desire for happiness, an object represented by “an idea in which all inclinations are summed up” (*Groundwork*, 4:399; cf. *Judgment*, 5:208). Happiness is described as an “ideal of imagination” (*Groundwork*, 4:418), or a “system of inclinations brought into tolerable harmony” (*Practical Reason*, 5:73–6), and the pursuit of happiness is said to involve a “unity of maxims” imparted by *reason*, which “inclinations otherwise cannot have” (*Religion*, 6:36–7; 32). In forming our idea of happiness, we consider not only our various desires, but also the means at our disposal (*Practical Reason*, 5:63). The resulting idea is our image of how we would like life to go for us.

The possibility of acting on second-order desires, however, depends on the negative freedom necessary for practical reason. Brute animals, as Kant conceives of them, seek the satisfaction of their desires one by one. Because the brute will is “coerced by impulses,” such a will has no ability

to suspend acting on an impulse so as to engage a larger conception of its general well-being. Only a being with negative freedom can choose against satisfying its desires (*Groundwork*, 4: 446; *Practical Reason*, 5:28–30), and this is necessary for any effective pursuit of a maximum of satisfaction derived from a set of partly conflicting desires. A being lacking negative freedom might very well be pleased, satisfied or contented with its state, but it could not *pursue happiness*.

Kant sometimes explicitly recognizes that the concept of happiness is available only to rational beings, not to brutes: “The concept of happiness is not one that the human being perhaps abstracts from his instincts and hence gets from his animality; instead, it is a mere *idea* of a state, to which he tries to make that state adequate under empirical conditions” (*Judgment*, 5:430). Thus, from the fact that all human beings (along with other animals) tend to seek pleasure, satisfaction, or contentment, it by no means follows that human beings have a desire for happiness.

This means that Kant needs some account of why human beings do sum up their desires into the idea of a whole of satisfaction. It would be viciously circular to explain the desire for happiness by anything like a desire to *maximize* pleasure or satisfaction, because that desire already presupposes the desire to make one’s state adequate to the idea of a *comprehensive whole* of satisfaction, which was the very desire to be explained. Nor would it work to argue that even without presupposing the desirability of maximizing satisfaction, framing an idea of happiness is instrumentally valuable to a rational being by enabling: it enables to avoid unnecessary conflicts between its desires and to use its resources for satisfying them more efficiently. For it is not clear what “efficiency” would mean here. Pursuing happiness no doubt results in the satisfaction of some desires, but also the frustration of others; without an idea of happiness, there could be no measure of overall efficiency in desire-satisfaction.

We can ask, of course, whether beings who form the idea of happiness and pursue it are generally more contented with their state than beings that do not. There are good reasons to doubt that they would be. Forming a concept of happiness and disciplining oneself to pursue it expend scarce resources that might have been used more effectively in satisfying particular desires. This is what lies behind Kant’s remark that “a single inclination, definite as to what it promises and as to the time at which it can be satisfied, can easily outweigh a wavering idea” (*Groundwork*, 4:399). Further, the process of deliberating about happiness itself might have bad side effects. It might, for example, create new and artificial desires (in-

cluding, of course, the desire to bring our state into harmony with our image of happiness itself) that would make it harder for the animal to reach a condition of contentment. As a matter of fact, Kant has a firm opinion on the question whether pursuing happiness is advantageous to us in terms of pleasure, desire-satisfaction, and contentment. Along with Rousseau, he is quite certain that it is *not*. This is why he thinks the welfare of an animal would be better provided for by instinct than by rational faculties, and therefore that if reason has a natural purpose in human beings, then it cannot consist in promoting their pursuit of happiness (*Groundwork*, 4:395). Kant is not the first to notice that rationally pursuing happiness is not an especially efficient way of achieving it.

V. A Wavering Idea

The most urgent question for us at this point is: Why do human beings pursue happiness? No satisfactory answer to this question can be given by referring to human finitude or to our animal nature (as Kant does in the *Analytic of the Critique of Practical Reason*). Kant's most explicit answer is given in the *Religion*, where he asserts that the desire for happiness, falling under our predisposition to humanity, calls for "a physical but *comparative* self-love (for which reason is required); namely, judging ourselves happy or unhappy only in comparison with others" (*Religion*, 6:27).

This answer to the question, however, is surprising and even counterintuitive. For it locates the original motivation in human nature for the idea of happiness not in our natural but in our *social* condition, and the social meaning of natural desires. We consider ourselves happy as a way of thinking of ourselves as better off than other people, and we think of ourselves as badly off only to the extent that we think of our condition as one that might cause others to despise us (*Religion*, 6:94). Unlike many of the animal desires which go to make up the idea of happiness, the desire for happiness is neither so natural nor so innocent as eudaimonists would like to pretend.⁷

This is the meaning of Kant's words in the opening paragraph of the *Groundwork* (which people usually read right over without appreciating their significance): "Power, riches, honor, even health, general well-being, and the contentment with one's condition called happiness, make for pride and even arrogance if there is not a good will to correct their influ-

ence on the mind" (*Groundwork*, 4:393). Perhaps we think that this means only that arrogance is merely one possible side-effect of thinking oneself happy. But Kant's real view is not nearly so innocuous. He thinks that arrogance is originally the entire rationale for our natural desire for happiness. In other words, it is our overriding need to confirm our self-conceit that makes us lay claim to happiness as a mark of superiority over others. We form the idea of a comprehensive whole of satisfaction in order to compare ourselves to others, always with the hope that the comparison will make us look good. The real point of being happy is to feed our insatiable *amour propre*.

If this view about the motivation of our desire for happiness strikes us as arbitrary and unmotivated, then we need to consider it in relation to Kant's theory of human nature as a whole. Kant regards our rational capacities as arising through a natural process, stimulated by our entry into social relations with others. Natural teleology has arranged it that our species-powers develop through social competition or antagonism, and has arranged our inclinations so that they bring about this development (*Idea*, 8:18–23). Kant maintains that our powerful natural inclination for happiness is only the most fundamental expression of the "unsociable sociability" that is fundamental to our nature as rational and social beings (*Idea*, 8:20–2). It is only to be expected, then, that the motivation behind the most basic and comprehensive of our *rational* desires should be this competitive impulse. Kant's theory that we seek happiness in order to confirm our self-conceit may be a surprising one, but it is far from being arbitrary or unmotivated within Kantian anthropology. On the contrary, it is virtually necessitated by his naturalistic theory of human nature as it develops in history.⁸

As Kant sees it, therefore, the eudaimonistic standpoint is originally the standpoint from which each human being seeks superiority over others. By contrast, the standpoint of morality or reason is one from which the worth of all rational beings is *equal*. From this standpoint, human beings are required to respect every rational person as having the dignity of an end in itself. They are to adopt only those maxims that they can rationally will all others to adopt, and they give themselves moral laws enjoining that they overcome all competition by uniting their purposes systematically into an ideal realm of ends. Human nature is therefore caught in a fundamental struggle between the natural-social standpoint, whose principle is one's own happiness, and the pure rational standpoint, whose principle is the moral law.

We might think that Kant would allow some inequality between agents on moral grounds, since he holds that the worth of my person is greater if I have a good will than if I have an evil will. But Kant consistently maintains that moral self-worth is an “inner” worth, involving a comparison only between myself and the moral law, never a comparison between different people. When I am conscious of failing to do my duty, I regard myself with “self-contempt and inner abhorrence” (*Groundwork*, 4:426), but I do not regard myself as lower than other people. And when I do my duty, my self-contentment is never a ground for feeling superior to them: “All human beings are equal to one another, and only he who is morally good has a superior inner worth” (*Moralphilosophie Collins*, 27:462). If this statement is not to be understood as some sort of Orwellian paradox, then the morally good person’s “superior inner worth” must be interpreted as involving comparison only with the moral law, never with other people. As Kant also says: “Moral self-esteem, which is grounded in the worth of humanity, should not be derived from comparison with others, but from comparison with the moral law” (*Moralphilosophie Collins*, 27:349).

Kant does allow that I must feel respect for a humble plain man in whom I am aware of greater virtue than I find in myself, but he insists that this is really only a respect for the moral law, of which the humble man’s conduct provides me with an example (*Practical Reason*, 5:76–7). At the same time, he makes it clear that respect for the moral merits of others is really no different from respect for their talents or learning, since (he claims) the real object of the latter is not the learned or talented person, but our own duty to develop our talents or improve our mind (*Practical Reason*, 5:76–7). Genuine moral humility (the awareness of our moral shortcomings) is never a matter of comparing ourselves with others, but is always grounded on our awareness of the dignity (absolute worth) of rational nature in our person, which is an inner and incomparable worth (*Morals*, 6:435).⁹ Kant believes, of course, that justice requires the punishment of external wrongdoing, in order to protect the right to external freedom. But he insists that forms of punishment must always respect the criminal’s dignity as a human being, which can never be lost or forfeited (*Morals*, 6:463). And punishment is regarded by Kant as a form of coercion, which is therefore entirely out of place in moral (as opposed to legal) contexts. There is no place in Kantian ethics for looking down on others as morally inferior, just as there is no place in Kantian ethics for employing moral blame or disapproval as social sanctions for moral rules.

Nature’s purpose in giving us a desire for happiness, however, is to

employ our competitive impulses in order to develop all the predispositions of our species, including even the rational predisposition that reveals to us the irrationality and moral wrongness of all such competitiveness, and commands us to conduct ourselves according to the laws of a realm of ends in which all rational beings are respected as having equal worth. This natural origin of our desire for happiness sheds light on the tension between Kant's conception of what happiness *consists in when (or if) we get it* (pleasure, desire-satisfaction or contentment) and what counts as happiness considered as an *object of pursuit* (an idea whose content depends on our comparison of our condition with that of others). The tension opens up the possibility that even if we achieved what is represented in the idea, then we would still not be happy (because we would still be displeased, unsatisfied, or discontented). And in fact, as we have seen, Kant thinks it is not a mere possibility, but a normal part of the human condition, that people are generally unhappy even when they get what they have represented as their happiness.

Nor is this an accident, in Kant's view; it is the way nature has arranged things. The natural purpose of unsociable sociability is to give people an incentive to develop the faculties of their species through competition. It is essential to this purpose that human beings should be prodded out of their natural laziness by being placed in a condition of *discontent* (*Idea*, 8:21). Looked at from the standpoint of natural teleology, therefore, we could say that *our desire for happiness serves its natural purpose by making us unhappy*. This is why Kant says that happiness is "the idea of a state to which we try to make our state adequate under empirical conditions (which is *impossible*)" (*Judgment*, 5:430, emphasis added). The desire for happiness, far from aiding us in reaching a condition of lasting contentment, involves the formulation of an idea to which our state can never in fact correspond, and it is part of the natural function of this idea that it should be so.

Kant's views here are far stronger than anything that could be derived merely from our finitude or dependence on nature. It clearly does not follow from the concept of a finite natural being in general that its natural needs are *incapable* of satisfaction in the system of nature as it exists, still less that its natural needs could not conceivably be satisfied in any system of nature. But precisely this is what Kant believes about human beings. The real barrier separating us from our happiness does not lie in a nature that is hostile to our desires, but lies entirely within our own idea of happiness itself. Kant thinks that even if we were omnipotent in rela-

tion to nature, we would still not be happy: "Nature, even if it were subjected completely to human choice, still could not adopt a definite and fixed universal law which would harmonize it with that wavering concept and so with the purpose that each person chooses to set" (*Judgment*, 5:430).

Kant maintains that although there are many "technical" imperatives ("imperatives of skill") telling people how to achieve arbitrary ends (in geometrical construction, for example, or the medical art, or other practical techniques), there are no general principles (or "pragmatic imperatives," or "imperatives of prudence") that one might follow with confidence in seeking happiness. "Counsels of prudence" (enjoining such policies as economy, courtesy, and restraint) are only general rules, which never hold without exception. They tell us what has been found to promote an individual's welfare on the average (*Groundwork*, 4:418). The reason Kant thinks we cannot do better than this is not only that life is too complicated for there to be any unexceptionable rules telling us how to reach a fixed goal. The problem is rather that the goal itself is too indeterminate and variable for us even to begin setting out to achieve it in a systematic way. Not only do different individuals have different ideas of happiness, but the same person is constantly changing his idea of happiness and even at a single moment no individual ever manages a self-consistent idea of happiness (*Groundwork*, 4:417–18).

The basic incoherence in everyone's idea of happiness is that what everyone "wishes and wills" includes both being pleased, satisfied, or contented, and also achieving the objects represented to us delusively by "the deceiver" inclination (*Anthropology*, 7:151). Nature has arranged it so that the things people want are such that when they get them, they will still be discontented enough to keep on exercising and developing the faculties of the species (*Groundwork*, 4:418).¹⁰ In short, it belongs to our nature always to be striving for happiness but never to be happy, and even to have an idea of happiness in terms of which we could not possibly be happy.¹¹

VI. Morality versus Happiness

From the above it is clear that although our desire for happiness proceeds from our rational predisposition, its origin in human nature also makes it a fundamentally irrational desire. Kant asserts that the idea of happiness is an ideal of imagination rather than of reason (*Groundwork*, 4:418). He explains this more fully in his anthropology lectures (25:1081):

Happiness is the kind of ideal of which we can form no concept, in which we could posit it, if we understand by "happiness" the greatest sum of joys, i.e., the complete satisfaction of all our inclinations. We cannot even represent the possibility of enjoying a life composed of sheer enjoyments. We can never produce a complete whole with which we could be wholly satisfied; hence this is an imagining to which no concept corresponds.

The idea of happiness is therefore always a 'delusion' (*Wahn*), in Kant's technical sense of that term: it is a pure figment of our imagination that we treat as something real (a real possibility, toward which we can direct our efforts) (*Anthropology*, 7:274–5). The desire for happiness is also a delusion in the sense that it is premised on self-conceit—on the false and radically corrupt notion that I am worth more than others and that confirming my self-worth involves gaining some superiority over them (in matters such as honor, power, or wealth). In fact, however, the worth of every rational being is absolute, and there is no real superiority over others to be had even if I could realize my idea of happiness.

"When one's own happiness is made the determining ground of the will, the result is the direct opposite of the principle of morality" (*Practical Reason*, 5:35). No such statement as this could ever follow merely from the fact that we are finite beings with natural needs, since that fact leaves completely open the content of our needs, and with it the whole question of their harmony or disharmony with morality. But this statement does follow from the fact that the principle of morality is based on our noncomparative worth as rational beings, while the principle of happiness is based on the comparative conception of self-worth and the delusive ambition to achieve superiority over others.

Now we are in a position to understand the Kantian explanation of what is wrong with Socrates' thesis that we are happier suffering injustice than doing it. Part of what is wrong is that Socrates wants to look at the difference between doing injustice and suffering it from the eudaimonistic standpoint. For from that standpoint, as Callicles points out to Socrates, the power to do injustice with impunity is seen as a mark of superiority, while suffering injustice is regarded as a humiliation.¹² It would take an entirely different standpoint, based on an entirely different notion of self-worth, to appreciate the loss of self-worth involved in doing injustice. And this standpoint would not be eudaimonistic.

The rest of what is wrong with Socrates' claim is that, true to its eudaimonism, it represents the loss of self-worth involved in doing injus-

tice *comparatively*, as if the judgment common sense makes from the standpoint of happiness were merely the reverse of the truth—as if it were the just man, even in being wronged and mistreated, who has reason to lord it over another. Once we correctly grasp the standpoint from which we can appreciate the loss of self-worth involved in doing injustice, we see that it involves no comparison between different individuals and no possibility of gain on one person's part at the expense of the other. For the moral standpoint is one from which every rational being, strong or weak, rich or poor, honored or despised, good or bad, has absolute worth. The only comparison this standpoint involves is one between myself and the principles I give myself as a rational being. The loss of self-worth involved in the commission of injustice is simply the inner shame at failing to live up to one's dignity as a rational being. And according to Kantian ethics, moral virtue gives me nothing to feed my *amour propre* in comparison with others.

Given Kant's view of what our desire for happiness consists in and how it relates to what is really valuable, he would be quite justified in scoffing at the utilitarian idea that morality would be well served by a rational pursuit of the general happiness. For even a single individual's pursuit of happiness is riddled with irrationality. If we think of calculating the general happiness as a process of typing the utility functions of all individuals into a computer and then deriving a collective result through the right algorithm, then Kant's view is that our project cannot get off the ground, because none of the utility functions we feed into our database will meet even the most minimal conditions of self-consistency. Even if people's individual conceptions of happiness made sense, then there would not be much room for maximizing happiness among them because the pursuit of happiness is basically a zero-sum game, in which any gain at all for one party will be at the expense of another. For the most part, we can speak of *agreement* between what will make individuals happy only in the way that Francis I reported that there was agreement between himself and Emperor Charles V: "What my brother wants (Milan), that I want too" (*Practical Reason*, 5:29).

VII. The Rational Pursuit of Happiness

We must be careful not to exaggerate the anti-eudaimonistic strain in Kant's theory. Kantian ethics is not *against* the pursuit of happiness, and

Kant even condemns the opposition to happiness as a morbid and monkish attitude, which he thinks no one can adopt without hypocrisy and even involves a hidden hatred of the moral law (*Religion*, 6:24–5).

“This difference of the principle of happiness from that of morality is not therefore an opposition between them, and pure practical reason does not require that we should renounce the claims on happiness; it requires only that we should take no account of them whenever duty is in question” (*Practical Reason*, 5:93). This remark is entirely consistent with Kant’s claim, quoted earlier, that “when one’s own happiness is made the determining ground of the will, the result is the direct opposite of the principle of morality” (*Practical Reason*, 5:35). For the opposition refers only to the will’s *motivation* (its “determining ground”). To give the principle of morality motivational priority to the principle of one’s own happiness is not to exclude the possibility of acting in ways that are largely conducive to one’s own happiness. Kant’s view, in other words, is that we should not let the principle of our own happiness motivate us in any case where moral principles are at stake. But morality has nothing to say against our pursuing happiness as long as a dutiful disposition governs this pursuit, limiting our self-love by striking down our ambitious self-conceit (*Practical Reason*, 5:71–5). It is not pursuing happiness as such that contradicts the standpoint of morality, but only the principle of pursuing one’s own happiness unconditionally, irrespective of the demands made on us by our own autonomy and the dignity of others.

Morality even gives us positive grounds for valuing the permissible parts of happiness, both our own happiness and that of others. When rational beings make the permissible parts of their happiness their end, they thereby confer objective value on those ends, because the dignity of humanity is honored when human beings achieve the ends they set, and a human being’s happiness is the sum of her nonmoral ends. Morality therefore even provides a reason for framing and pursuing our idea of happiness, which is that when we use reason to fashion an idea of happiness and set happiness as our end, we exercise our rational capacities. When we further such ends, whether our own ends or those of others, we honor the rational nature that is exhibited in them. To fail to value happiness, therefore, would be to show contempt for ourselves as rational beings.¹³

The moral law, in other words, permits the pursuit of happiness whenever it is not specifically contrary to duty. We even have an indirect duty to promote our own happiness insofar as this promotion contributes to our perfection as rational beings and honors our rational nature as an end

in itself (*Morals*, 6:388; *Groundwork*, 4:399). But if we are to honor all others as our equals, as duty requires, we cannot pursue happiness in the spirit of arrogance that originally made the prospect of happiness so delicious to us. The point of morality is to honor human dignity, and when we see what our desire for happiness originally meant, we must acknowledge that this means that morality aims not at maximizing human happiness but constraining people to forgo enough of their happiness that their various ends, originally antagonistic to one another, can be brought into harmony under the laws of a realm of ends.

Given Kant's view of the role of the desire for happiness in human psychology, prudential considerations themselves also militate against eudaimonism as a principle. Because nature did not put us on earth to enjoy happiness, an ethical theory that places the maximization of happiness before everything else is likely to be self-defeating even in its own terms. It is in danger of encouraging us to indulge the dreams of sweet contentment that float before our eyes, always just beyond our reach, whose natural function is only to fuel our discontent (*Conjectures*, 8:122–3). Even achievable parts of happiness (or the satisfaction of particular desires and ends falling under the idea of happiness) are difficult to attain, and will be especially difficult of attainment if they are brought under the explicit pursuit of happiness.

Prudence itself, therefore, as well as moral wisdom, would sooner advise us to abandon the delusion of perfect happiness, and turn our attention instead to the serious business of life, which is to regulate our lives by rational principles. These principles will limit our pursuit of happiness, but because their ground is honoring our worth as rational beings, they will allow for what Kant sometimes calls a “reasonable self-love” (*Practical Reason*, 5:73; *Religion*, 6:45n), that is, a pursuit of one's own happiness that is moderated enough to accord with duty and to allow for the moral claims of others, who are our equals as ends in themselves.¹⁴ The limited happiness achieved through this moderated self-love may, in any case, be as much as our paradoxical condition will ever afford us.¹⁵

Notes

1. For the claim that Kant breaks with the previous tradition, see Allen Wood, *Hegel's Ethical Thought* (New York: Cambridge University Press, 1990), 56–8. But the Cambridge Platonists, and Ralph Cudworth in particular, anticipated Kant

by holding that it “would destroy all morality” to ground moral obligation on the agent’s own happiness, as Hobbes and Locke had done. See Stephen Darwall, *The British Moralists and the Internal ‘Ought’* (New York: Cambridge University Press, 1995), 145.

2. Aristotle, *Nicomachean Ethics*, 1:1, 1094a.

3. Plato, *Gorgias*, 472; *Republic*, 354.

4. Plato, *Gorgias*, 474.

5. Plato, *Republic*, 343–7.

6. There may be more than an inkling of Kant’s views concerning the irrationality of desire in Plato’s discussions of desire and reason in the *Phaedrus*, *Republic*, and *Symposium*. But neither the social dimension of these views, nor their specific application to the idea of happiness, are present in Plato.

7. This is made more explicit in Kant’s anthropology lectures: “Nature cannot guarantee us anything but the convenience and sustenance that can be used only among other human beings; and all poor circumstances in regard to conveniences and sustenance are unbearable to us only insofar as we cannot have them in common with other human beings. We do not complain about nature itself in regard to our niggardly circumstances, but only because other human beings have it better than we do. If my every meal is bread and water, then this grieves me because I know that other human beings have it better” (25:470; cf. 25:1103, 1324); and: “A bad meal, which nevertheless satisfies our needs, does not make us unhappy. But to have torn clothing and to appear in society with this, this is unhappiness. One feels cut off from society and loses the respect of others. Such a person is not offended by the lack but from the low esteem he expects from others. Human beings are always ready to despise him who is unhappy or destitute according to his estate in society. Hence clothing and external goods belong very much to the happiness of life” (25:1504).

In these passages, Kant sometimes cites either (Henry) Home (Lord Kames) or David Hume, but the only plausible source for what he says is from the writings of a different Scotsman: “The Greeks and Romans lived, I suppose, very comfortably, though they had no linnen. But in present times, through the greater part of Europe, a creditable day-laborer would be ashamed to appear in publick without a linnen shirt, the want of which would be supposed to denote that disgraceful degree of poverty, which, it is presumed, no body can well fall into without extreme bad conduct. Custom, in the same manner, has rendered leather shoes a necessary of life in England. The poorest person of either sex would be ashamed to appear in publick without them”; Adam Smith, *The Nature and Causes of the Wealth of Nations*, The Glasgow Edition of the *Works and Correspondences of Adam Smith*, eds. R. H. Campbell, A. S. Skinner, and W. B. Todd (Oxford: Oxford University Press, 1976). 5.2.4, 869–70.

8. For a fuller discussion, see my paper “Unsociable Sociability: The Anthropological Basis of Kantian Ethics,” *Philosophical Topics* 19:1991, No. 1.

9. See my discussion of the cited passages in "Self-Love, Self-Benevolence and Self-Conceit," in *Aristotle, Kant and the Stoics: Rethinking Duty and Happiness*, eds. S. Engstrom and J. Whiting (New York: Cambridge University Press, 1996).

10. Kant does not, however, seem to hold the absurdly pessimistic view of Schopenhauer, that all desire-satisfaction is a delusion, and that nothing of our happiness can ever be attained. When he discusses the way in which the pursuit of happiness inevitably leads to discontent, the idea seems rather to be that although we may satisfy our desires (for wealth, knowledge, health, or long life), when we do so, the result is likely to be that there will arise further sources of discontent (wealth leads to care and suspicion, knowledge to our awareness of evils in life previously hidden from us, and so on) so that we continue to remain at a distance from the final contentment we represent in our idea of happiness (*Groundwork*, 4:418–19).

11. The inherent discontent of the human condition is emphasized by Victoria S. Wike, *Kant on Happiness in Ethics* (Albany: SUNY Press, 1994). Her metaphysical explanation of this, however—that it rests on Kant's distinction between the phenomenal and the noumenal, and destines us for happiness only in a future life—seems to me to obscure the important anthropological theses (especially Kant's Rousseauian conception of human nature as it develops in the social condition) that lie behind Kantian doctrine on this point.

12. Plato, *Gorgias*, 483–6.

13. From this point of view, in fact, what makes a person's happiness objectively valuable is *only* the fact that a person has set it as an end. Hence one corollary of Kant's position is that there can be no moral ground at all for paternalistic conduct that foists some happiness on another that the other has not chosen to pursue.

14. This important concept in Kantian ethics, especially in relation to the conception of the highest good, is explored by Stephen Engstrom, "The Concept of the Highest Good in Kant's Moral Theory," *Philosophy and Phenomenological Research* 52:1992.

15. Earlier versions of this paper were read at the meeting of the U. K. Kant Society in St. Andrews in September 1996, and at a conference honoring the memory of Lewis White Beck at the College of the Holy Cross in Worcester, Massachusetts, in May 1998. For helpful discussions on those occasions I wish especially to thank Henry Allison, Michael Friedman, Otfried Höffe, Onora O'Neill, Adrian Piper, Stanley Rosen, and John Skorupski.

Kant and the History of the Will

Yirmiyahu Yovel

Lewis White Beck's work was exemplary both in analyzing Kant's systematic ideas and in exploring main threads in his historical background. In this vein—if in a different style—I dedicate a two-tiered essay to his memory. The text is drawn from a blueprint for a study I am preparing on modernity and the will. Part I sketches a bare-bones “history of the will” until Kant, and Part II elaborates on Kant's place in that story, namely, the various aspects of his theory of the will.

Part One: Modernity and the Will

Modern philosophy is usually presented as the history of epistemology and metaphysics, that is, of the intellect in relation to the world. Yet not only modern philosophy but also, I think, modernity at large (including its genesis) can be better described in terms of a history of the *will*.

Of course, there is no single “modern will” in the essentialist sense. Yet several characteristic features of the will can be spelled out, some usually working together, others competing within a common family denominator. The modern will is typically self-assertive, expansive, and individualistically shaped; in most of its varieties it strives to self-realization and self-enhancement, to personality and individual autonomy; it stresses initiative and the attribution of responsibility, demands to be recognized as a normative source and the bearer of rights, and seeks the embodiment of that recognition in external institutions. The modern will questions tradition, values novelty and innovation, unsettles rigidly received rules, and strives to penetrate beyond any given horizon, whether of knowledge or life. In all that the will is typically restless, stretching beyond itself and its present situation in constant search of the new, the distant, the different. And, perhaps most characteristically: the modern will refuses to conform to the constraints of the world as given. It strives to change the world in its own image and interests: to reshape and control

outer nature (through technology, planning, resource, exploitation, etc.) and to rationalize the social and political world—be it for the purpose of maximizing gain and pleasure, or for realizing a moral utopia, or (as in Kant) for both. In that quest, the human will tends to posit for itself the goals it pursues, and to set itself as the ultimate ground of moral precepts, political legitimation, and the standards by which historical progress is gauged.

Metaphysically speaking (using Platonic terms which, on this issue, I consider pertinent), the modern will reorients the mind's fundamental Eros, its quest for transcendent infinity, back into immanent, this-worldly affairs. That crucial shift is manifest, for example, in the circular infinity of insatiable desires rekindled by their satisfaction (which capitalist "market" economy has turned into a world system), or in chasing an ever distant historical utopia—universal liberation, justice, love for the other, even secular salvation.

Philosophy, however, is a late-comer. As "Minerva's owl" (Hegel) it often conceptualizes major trends in life long after they had been active in practice. Descartes, the canonical founder of modern philosophy, was still largely attached to a basically medieval, Augustinian concept of will. In practice, however, the modern European will had started taking shape a full century before Descartes, with the rise of the new, proto-capitalist urban classes; with the great naval discoveries that opened new horizons for endeavor, vision, and exploitation; with the formation of a semi-centralized monarchy that could exploit (and initiate) new opportunities; and, of course, with the struggles around the Reformation, which transformed people's ideas and aspirations concerning God, salvation, the next world and the present one.

The colonial ventures of the sixteenth and seventeenth centuries shattered the geographical and anthropological confines of the medieval world, just as the image of the infinite cosmos, which was then emerging in science and metaphysics, broke the boundaries of the closed Aristotelian cosmos. What was to fill the vast, unlimited horizons that were opening up? Whereas the cosmic space was to be inhabited and mastered by mathematical formulae (set by God but, through knowledge, co-mastered by man), the new geographic and ethnic horizons were to be filled and dominated by an expansive human (i.e., European) *will*, also stretching itself to infinity.

Vehicles of this expansive will were trade and religion, the sword and the cross, missionaries in search of alien souls to convert, and venture

capitalists coveting a yield in gold, silver, spices, sugar, and slaves. The royal banner, the savior's icon, the banker's credit, the cartographer's science, and the soldier's prowess were all embodiments of the new, proto-modern will, and ingredients in its construction. However, we should not understand this movement as narrowly utilitarian. People were turning to unknown lands in search of gold and rare commodities, but also in search of power, pride, adventure, the glory of God, and the gratification of their own assertive selves. Their venturing will no longer waited for the End of Days to see the world transformed; it looked out for the shape of a new world—physically and metaphorically—to emerge within the existing world. Thus the Spanish *conquistadores* were searching for Eldorado in its triple sense of gold, adventure, and a dim spiritual utopia; and their notorious insatiability, although driven mostly by greed, also expresses a will that wills its own infinite use and exertion.

At the same time, the rising urban classes turned more and more toward this-worldly affairs as a way of giving meaning to their lives. This enhanced the formation of the *individual* will that saw its main territory in the affairs of this world and claimed a measure of freedom—*libertas*—from the normative constraints of collective identity as dictated by tradition, community, tribal, and religious affiliation. Religion was not neglected, but was experienced in a more personalized, often this-worldly oriented way, or at least as a decentered, less totalistic dimension of a person's life.

Parallel to that arose the sense of a personal *career*—cultivating one's individual path in the world as a valid life goal. A career is not just any achievement or endeavor. A career is an individual itinerary traced by a singular individual constructing her sense of self by seeking personal and familiar self-fulfillment in this world. Pursuing a career does not mean that the person is more egoistic, but that she is more individualistic—that she reclaims and recaptures her singular life from its immersion in a traditional society and in the extra-mundane goals that this society (and the religion overseeing it) have determined for her. People anchored in traditional societies, even when they strive for success, are not precisely career-driven: this concept is more strictly applicable to cases in which terrestrial interests are working under a dominant personal perspective, significantly severed from a person's communal ties and social and denominational loyalties.

The early sense of individuality seems (*pace* Weber) to have preceded the Reformation and may even have helped make it possible. For it was,

above all, the rising, self-conscious individual will that initiated the revolt against the Catholic Church's impersonal formalism, a church whose hierarchy was qualified to dispense of salvation regardless of the clergy's personal disposition and state of mind. In rejecting this metaphysical objectivism—in "subjectivizing" religious experience and value—the Lutheran Reformation implied the primacy of the individual's disposition and mental state, although, for reasons of anti-Catholic polemics, it also adopted some version of the doctrine of predestination. The milder version of predestination did not, however, exclude free will; it only denied that ultimate salvation depends on it. And even the strictest Calvinist version left—indeed, opened up—a large domain for a person's will to assert and extend itself in *this* world, in compensation for the will's failure to produce an outcome in the next world.¹ The strictness of Calvinism actually toughened the religious will, because it required a person to live a true Christian's life regardless of what had been predestined for her. Thus Calvinism demanded—and, where successful, cultivated—a strong will-power in ordinary members of the community, more intense and powerful than anything medieval Catholicism could hope for except in an elite minority of monks retired from the world. Also, predestination in all its forms placed a religious value on the non-utilitarian ethos of *doing things for their own sake*, a mental disposition that eventually spilled over into worldly affairs, including humanist morality and education (such as Kant's), and was thus "secularized."

So much for the theology of the Reformation. And what about philosophy? The late-coming owl of Minerva, once it started conceptualizing the will's evolution, did so with keen interest. The will—either under its own name or under different names—became a central theme in modern philosophy whose changing conceptions manifest certain characteristic patterns and trails. If worked out into a "history of the will," these transforming conceptions can provide an alternative narrative of modern philosophy, encompassing (and partly superseding) such concepts as "secularization" and "the Subject" in providing a genealogy of modernity.²

A focal point in any such study must be Kant's Copernican Revolution of the concept of the will. Kant's revolution, and its further transformation by his followers, provide a crossroads from which to look back on the evolution of the concept of the will since Descartes and the early Enlightenment, and from which the era of high modernity has taken its departure.

Kant and His Predecessors

Kant's predecessors can be divided into ancient, medieval, and early modern. The Greeks hardly used the concept of the will. Aristotle does not even have a special term for it. For him the focal term of practical reasoning was *Proairesis*, a form of action-oriented deliberation that weighs the various aspects of a proposed course of action and follows its natural outcome. It was primarily with St. Augustine and later Christian theology that the will entered the center stage of human action, and was conceived as a free power of choice, and also as a basically negative principle, the source of evil, sin, and moral weakness. This view exempted God from direct responsibility for those negative features and transferred their origin to the irrational depths of the human will. Even when exercising its positive role, which mostly consisted in restraining and molding itself in accordance with the divine will, the human will was free to rebel and break the divine law, because it was endowed with *liberum arbitrium*, the power of quasi-arbitrary choice for which no previous reason, however strong, could supply a fully sufficient ground.

Thus, free choice introduced an irreducible element of irrationality into the will's very constitution. Also, by making the will the source of evil and negativity in the world, this irrationality was further expressed in the will's typical works, because its most distinctive product—that which no other agent in the world, but the human will could generate—was not Good, but Evil.

The Early Moderns

The early moderns—Descartes in particular—remained by and large within the Augustinian tradition. The free will plays a crucial role for Descartes in both epistemology and ethics, and its true role in these domains is to efface itself. The will is, first, the origin of error, which Descartes conceives as a cognitive kind of sin. In order to overcome the “sin” of error, the rational philosopher must disable his will, by (a) acquiring compelling rational ideas that leave the will no freedom of maneuver, and (b) by having the will paralyze *itself* (abstain from judgment) on all issues lacking clear and distinct (rational) ideas that compel the will to give its assent. In both ways, the human will is best treated by self-negation, since unlike the human intellect, it is fundamentally irrational.

In ethics, too, the will's role was to restrain itself so as to suit the limited possibilities offered by the world and agree with it without much conflict and suffering. In this neo-Stoic approach, it is no longer the divine will but the ways of the world to which the human will has to adapt. As for Descartes' creator-God, He could have created eternal truths that differ from those governing the actual world, which implies that God's intellect, unlike ours, is not compelled by rational ideas. If God did not equip the world and our minds with laws and ideas that break the known rules of logic and are irrational from our viewpoint, it is because His veracity—a *moral* perfection, pertaining to His will—barred Him from cheating on His human creatures.

The view that rational truths are contingent was severely criticized by Spinoza and Leibniz, the more rigorous rationalists. Yet Leibniz, anxious to salvage the idea of creation from Spinoza's onslaught, in the final analysis also pictured God as acting from choice rather than rational necessity, which in Leibniz is binding only on God's intellect, but not on His will. Since the world's factual features are the product of God's will, they are all contingent. This does not make them totally arbitrary, because the divine will's choice of a factual world is guided by the principle of the good, again (like veracity in Descartes) a moral principle. Still, the supposed goodness of the world leaves its fundamental ontological contingency unchanged. There is nothing in the world *itself* that requires its existence (no Spinozistic *causa sui*). The world's being there, and its being what it is, are due to the logically impenetrable external principle of the divine choice and will.

Nevertheless, Leibniz's theory implicitly draws a distinction between two forms of nonrationality, contingency, and arbitrariness; and the divine will plays an opposite role with respect to each of these. Matters of fact are logically contingent but not quite arbitrary, because there is *some* ground—a moral ground—for their existence. While the will as *choice* is the source of the world's factual contingency, the will as *good* responds in part to that same contingency, by providing the assurance that the world nevertheless has some rational sense: it has a *moral* sense, which does not emanate from the world's own being and cannot abolish its basic contingency—but is *endowed*, *bestowed* upon the contingent world by the goodness of the same divine will whose act of choice had produced that contingency in the first place.

Leibniz thus thematized an experience that was to dominate the sense of life in modernity—the experience of the world's inherent contingency,

which makes it appear “senseless” and existentially frightening. Leibniz responded to his own challenge by renovating the idea of teleology and making it compatible with modern science; yet I think what makes Leibniz modern is rather the problem than his offered solution.

Rousseau and Kant

Kant, following Rousseau, responds by a new theory of the will. He construes the will as fully rational *in potentia* and as reshaping the world rather than conforming to it. In the end, the contingency of existence is redeemed in Kant by the *human* (not the divine) will, which indeed *bestows* a moral meaning upon the world through the will’s moral-historical action. (See the end of Part II.)

Bridging the way to Kant was Rousseau’s notion of a “general will,” which indicated to Kant (1) that the will can be construed as a universal, and thereby as a *rational* power; (2) that the freedom (i.e., self-determination) of the will is to be understood not as choice merely but primarily as *autonomy* (i.e., shaping itself in accordance with its inherent universal structure); and (3) that the rational will must also reshape the *outer* social world by endowing its own rationality to it.

Kant integrated these ideas into his “Copernican Revolution” in which there is no longer a creator God to account for the rationality of the physical world, nor for the binding force of morality. The a priori structure of the human intellect fulfills the first task, and the a priori structure of the human will fulfills the second. Being remains contingent, but acquires a rational-moral meaning through the human will and its action in society and history.³ This idea was so fundamental to Kant that he even defined “the final end of creation itself” in terms of the will’s moral-historical ideal.

The Kantian concept of will differs from all previous conceptions in that, for Kant, the will defines the Good as its own immanent structure rather than conforming to a good given beforehand. Correspondingly, the will is conceived as rational, universal, and autonomous, independent of God, free of tradition, overcoming nature, reforming human society, and reshaping the external world, rather than molding itself in consonance with it. As practical reason, the human will is the source of the Good (*das Gute*) in personal morality, and of the “Highest Good” (*das Höchste Gut*) in the social and historical domain.

This involves a major shift over and above the Augustinian and Leibnizian traditions. The will becomes rational and, as such, is identi-

fied with practical reason. As rational, the will's genuine and primary product is no longer evil (as Augustine and Descartes imply), but rather the Good. Indeed, the human will—now identified with practical reason—is the source that generates rationality in the social world understood in the substantive sense of “reason” (*Vernunft*) rather than as mere instrumental intelligence.

Also, as “practical reason,” the Kantian will asserts its primacy over the cognitive domain—including science, theology, and speculative philosophy. This is dramatized by the moral will endowing the world's existence with its only possible meaning and goal—a function that no other agent, including God, can fulfill in Kant's humanistic scheme of things.

At the same time, opposing the theology of predestination as well as Spinoza's rational determinism, Kant accepted free choice as the will's underlying faculty, and thus, on certain readings, admitted a residual “irrationality” in it (as e.g., in the form of “radical evil”). The Augustinian will—recast as *Willkür*—could neither be fully rationalized nor eliminated from a moral theory that stresses responsibility, and makes the free personal choice (or adoption) of rationality an indispensable condition of being rational. This is where Kant differs radically from Rousseau, in whom the concept of oppressive or “totalitarian” reason is conceivable, whereas in the Kantian system of ideas it is self-contradictory and therefore absolutely impossible.

Fichte and Hegel

In an earlier work I developed Kant's theory of the “Highest Good” into the regulative idea of history.⁴ This implies a “second stage” in Kant's practical philosophy, revolving around an “expanded” view of the will that wills its own morality embodied in the outer world and reflected back from it. This view becomes explicit in Fichte and Hegel. Fichte's moral will is as world-shaping as his “I”; it represents the Ought to which the world must conform through human action. Kant's and Fichte's vision opened the way to post-Kantian utopian projects, both liberal, nationalistic, and utopian-Socialist, which Marx later criticized. Marx used Hegel's historical dialectic to undermine Fichtean utopianism, yet Hegel, unlike Marx, did ascribe to the will an essential role in historical progress. Although he rejected the utopian will, Hegel construed human society and history as the embodiment of the will, which is the rational power (or Idea) at work in the whole realm of Right, including ethics, politics, cus-

tom, social and juridical institutions, and their evolution. This was no longer Fichte's a priori will that supposedly molds the world according to a moralistic blueprint of the Ought, but a dialectical will embodied in the real historical process and working in accordance with its dynamics. As such—as objectified and dialectical—the will plays a crucial role for Hegel in constituting the ethical world (*Sittlichkeit*) generally—including the family, civil society, work, law, and the state; and the ethical world, as the will's product and embodiment, is the necessary medium in which Hegel's absolute spirit—God Himself—is being realized through human life and activity.

Anti-Rational Reactions

The Kantian revolution and the reactions to it—from Fichte, Hegel and Marx, to Schopenhauer, Nietzsche, and Heidegger—set the conceptual path of high modernity with its grand designs and great disillusionings, its moral utopias and technological disenchantment, and the struggle going on within it between the universal, world-shaping “rational will” and anti-rational voluntary principles like Nietzsche's “will to power,” Freud's *thanatos*, Kierkegaard's existential leap, Bergson's *elan vital*, or Heidegger's death-oriented “resolution.” These reactions to the rational and autonomous Kantian will are set within the universe of modernity and could not have been possible in their actual form outside it; also, they all (like American pragmatism, at least in its beginning in Peirce) maintain the Kantian insistence on the primacy of the practical domain over the theoretical. Thus, rather than simply opposing modernity, I think they ought to be viewed as a genuine branching off of modernity itself, comprehensible only within its context and thus dialectically belonging to it.

Part Two: Kant's Will

Having sketched the historical context, I shall in Part Two offer a systematic reconstruction of Kant's theory of the will. Because I wish to take all its aspects into consideration, I shall approach the matter in the form of a conceptual synopsis rather than by detailed textual analysis.⁵

Kant takes morality, by its very concept, to be a disposition to do the good for its own sake. Such disposition, and the action that translates it in terms of a given empirical situation, cannot be natural, but must be *Wille*-

d; hence, the will is the key concept in morality. Following Rousseau, Kant views the will as having a universality-capable structure; in acting morally, the will actualizes its inherent universal structure, i.e., actualizes itself. And since this is not a natural process but a *Wille*-d one, it follows that in order to be rational (moral) the will has to will its self-realization—it has to will itself, as realized. Thus, morality implies the (rational) will willing itself.

This crucial point—that in morality the will wills itself—has several important implications in Kant. Perhaps the best known is the anti-utilitarian implication: that the goal of doing good is doing good; and nothing can be good if it has a different ulterior goal. Another implication is that a moral will is directed, as its final and ultimate goal, not at the state of affairs the action is to produce in the world, but at the state of the will to be attained by the action and by the decision preceding it: the will wills to realize itself as rational and moral in and through any morally relevant situation in which it finds or places itself. A further, less obvious implication is that being moral is to will to will. The will's task is to activate itself, to reach out and actualize its rationality. It must not sit back and merely react to circumstance. Still, the will's ultimate goal is its own morality (i.e., pure practical rationality), even if in a subsequent stage this goal requires the will to work for the "moralization" of its social environment and, eventually, of the whole historical world.

Perhaps the most uniquely Kantian implications of the principle that the moral will wills itself concern the object and ultimate justification of morality. Kant conceives of this justification in terms of the *causa sui* model traditionally reserved for God: the moral will, by willing itself, grounds itself in a (benign) circular manner. It requires no external authority nor ground but is, in that sense, self-sufficient. By willing, the rational human will gives itself both the law and the exclusive authority to enact that law: just like the traditional God, the human will asserts this authority from within its being. Kant's implicit use of the *causa sui* model here is not surprising, because, in effect, he set out to replace God with the rational (if finite) human will as the unique source of all normativity, be it moral, political, or even religious; and because, furthermore, Kant typically views his humanist morality as equally—indeed, as inheriting the unconditional character and sublimity of religion—and therefore as similarly based on a self-grounding principle that provides absolute justification. Herein lies one of Kant's great attractions for many moderns, who wished their humanistic culture to inherit the absolute characteris-

tics of religion; yet others, myself included, will see this as evidence that Kant's thought had not been fully secularized, but remains essentially informed by the religious mind, and, in the end, fails to take seriously its own modernizing idea of human finitude.⁶

The Good Derived from the Will

Kant's revolution of the will (for no other word aptly describes it) is technically expressed in a somewhat dry text entitled "On the Concept of an Object of Pure Practical Reason" in the *Critique of Practical Reason* (5:57–71). Kant here reiterates the idea that the moral will wills itself, by claiming that the Good (*das Gute*) does not precede the will but is derived from the (pure) will as its inherent object. In other words, the will does not recognize a Good given in advance, but projects and constitutes the Good as the expression and of its own (universal) form and mode of operation. It follows that, to be moral (or to become practically rational, which amounts to the same), the will must not simply obey or accept a good passively given to it, but must constitute the good that it obeys, as the projection of its own rational self; at the same time, the good itself becomes good only in so far as it is derived from the rational, self-universalizing will as its product.

This is a more precise and better spelled-out way of saying that the will wills itself. The object here is nothing but the will's own mode of operation, its inherent rationality, which is externalized into a binding objective norm. Hence ethics recognizes no ready-made law, nor good, to which the human will must conform in order that it too be good. Ethics is based on the pure human will's self-legislation—its projection of itself as universally good—rather than on some antecedent "moral value" or material virtue prescribed by an external authority, be it God's will, social convention, or an ontological "law of nature" read off the face of the world. Just as in Kant's Copernican Revolution of the philosophy of nature all objective states of affairs are constituted by the human intellect that endows them with its own inherent universal structure, rather than read off the face of a pregiven system of things in themselves, so in ethics, too, the good does not independently precede the will but is constituted by the pure will and therefore dependent on it.

This view opposes all classic moral theories, both secular (especially Platonic) and religious (Judeo-Christian). It raises the human will to a

status unknown in the history of philosophy. First, the human will replaces God's will in founding moral commands, legitimizing political power, even establishing true religion; and second, the rational will takes precedence (has a "primacy") over the intellect in breaking the confines of phenomena and operating within the noumenal world as its own. Also, the moral will alone can establish a critically valid teleology, and even project a "final end for creation itself," thus providing the only possible Kantian answer to the query, "why is there a world at all." (Kant's answer places the ground for the world's existence in what the world has yet to become, which implies the primacy of the yet-absent rationality over the existing one.) Finally, the will's normativity allows Kant to postulate—hold as rationally valid—certain beliefs that cannot be decided on pure cognitive grounds, in particular the beliefs that the will is free, and that the world can adequately respond to moral reshaping. The postulates of practical philosophy are insights into the nature of what there is drawn from the nature of what ought to be *Wille-d*, of how the will has to shape itself and act in the world.⁷

Pure Practical Reason Interpreted as Will

Underlying these implications is Kant's interpretation of pure practical reason as will. This move changes the nature of both reason and the will. Reason (or rationality) is conceived not as calculus only, but as interest, a motivating power; and the power of the will is understood as inherently structured by the form of law, and as striving for universality in both its inner operation and the way it ought to shape the outside world. Such a will is self-willing: seeking rationality for its own sake, and striving thereby to realize itself. It also means being autonomous, because the will, in acting rationally, needs nothing external either to guide or to motivate its action. Not only does it "obey its own law"; it also motivates itself and sets the ultimate goal to itself.

This triple self-sufficiency—of guidance, motivation, and ultimate teleology—is also intended by Kant's positive answer to the question "whether pure reason can be practical." The famous question has three meanings: (1) Can human rationality provide guidance in moral affairs, that is, give itself the law by which to act morally? (2) Can it also provide the motivation by which to follow its own law? (3) Is human rationality an instrument only, or can it work as end in itself? In other words, is "being practical" a matter of being simply pragmatic—adjusting means

to ends, solving strategic problems, and, in general, promoting goals we receive from external sources—or can human rationality project of itself the goals toward the fulfillment of which it mobilizes its strategic and calculating capabilities?

In answering that pure reason can indeed be practical, Kant has all three aspects of the question in mind. Human rationality can (1) provide moral guidance, (2) supply the necessary motivation to follow its law, and (3) set its own goals (in Kant's term: its "essential ends"⁸) as the projection of its immanent nature. This implies a different meaning of rationality: *rationality as end in itself*. Instrumental rationality is not thereby abrogated, far from it, but becomes subordinate to a higher form of rationality, which Kant describes as goal-setting activity and, indeed, as a self-sufficient teleological system. Reason's calculating and instrumental powers (which we may call "intelligence," in order to distinguish them from reason's substantive meaning) are the problem-solving capabilities used by strategists, technical advisers, and various advocates who borrow their goals from their clients' interests. However, human reason in the proper sense sets its own goals in Kant and follows its inherent interests.⁹ When acting in a genuinely rational manner, we need not—and ought not—borrow our goals from external interests, but ought to serve and advance reason's inherent "interests" and "essential ends" (these are both Kant's terms).

Reason as Interest

The concept of an "interest of reason" recurs constantly in Kant's writings. It is crucial to his metaphilosophy and concept of rationality, and also, specifically, to his theory of the will. Human rationality, in all its modes, is a self-directed interest. It is not a calculus only, a mere static form, but a rational drive, seeking to realize itself—that is, to realize rationality as end in itself—in every relevant domain and situation, as required by that situation. Rationality as interest strives to realize itself in particular acts and instances; and the rational will represents that interest in the practical domain: morality, law, politics, education, etc. It is this interested, self-driving character of rationality that allows Kant to (specifically) identify practical rationality with the will (*Wille*).

All human acts are performed from interest; and moral acts are performed from the interest of pure practical reason to assert and realize itself. More precisely, because in Kant rationality is always related to the

activity of a subject and performed in the first person, moral acts are generated when an agent identifies the interest of reason as his own, and acts from it on the basis of that identification.

This is an act of self-universalization: recognizing and retrieving my particular interest in the interest of universal reason. At the same time it is an act in which universal reason individuates itself by becoming the active motivation of some particular will. The two aspects—the universal being individuated by a particular act of will (choice), and the particular will universalizing itself by adopting the interest of reason as its own, and thereby actually *constituting* itself (to be explained later)—are of equal importance to Kant. Each is a necessary but insufficient condition for realizing morality, that is, practical rationality. Just as behaving “according to the law” but not “from the law” is neither moral nor practically rational (because it lacks, as active “intention,” i.e., there is no self-universalization), so an intention that fails to universalize itself is equally failing. The adequate “intention” universalizes the particular will and individuates universal reason in a single act. (Later we shall construe this act as *Willkür* choosing itself in *Wille*.)

To elaborate: a moral rational act is performed when someone particular adopts the universal interest of reason as her own interest. In such a case, the agent chooses her own will’s universal structure (expressed in the form of the moral law) as the determining ground and motivating interest of her action. Thereby she overcomes her particularity in view of a universal principle that is derived from her own rational will and therefore is not coercive but free. The universal element is not experienced as externally imposed but as self-structuring and self-imposed, and thereby an act of self-determination (i.e., autonomy or freedom). My rational self realizes itself through my particular will, which in adopting the rational self’s universal form and interest as its own allows me, the single subject of all those actions, to become a self-realized, self-asserting, free individual.

One cannot overstate the importance to Kant’s model of pure practical rationality that self-universalization be always performed by a particular will that adopts the universal element and commits itself to it in a special intentional act (called resolution, intention, disposition, respect [for the law], etc.). All these names indicate that morality is not a mere logical calculus or operation. It requires a first-person act involving the dynamic of commitment and resolution. This, in Kant’s famous formula, amounts to acting “from respect to the moral law.” Respect is not an ordinary feel-

ing but a stance, taking a stand; it is based on recognition and puts this recognition into effect. What is thereby recognized? On a first level, the binding power of the law and its implicit origin in the agent's own rational self (*Wille*). Yet looming behind the law is the other person—all other persons—as the truly fundamental object of recognition. Only this fact endows the universal law with specifically *moral* value. For we might ask: Why must a person wishing to be good express a universal feature in her action? What makes universality the gauge of morality? The Kantian answer, I think, lies in the kind of recognition implied in the concept of respect. The object of my recognition is not the moral law alone but, underlying it, all other humans ("finite rational beings") who are my equal partners in enacting and constituting that law. We saw that in Kant there is no "command-in-itself" preceding the will; the moral law is constituted by the subject's rational will as it places itself in the standpoint of any other rational will (i.e., universalizes itself). And the reason why the will should so place itself is because it recognizes the potential finite rationality (i.e., humanity) of all these others, a status equal to its own that makes them necessary partners in producing a moral ought. Therefore, my will must place itself in their standpoint and make sure there is no contradiction between our mutual respective individual wills; namely, it must be able to universalize its maxims.

The Two Versions of the Categorical Imperative

Herein lies, I think, the answer to an old puzzle in Kant scholarship: how is the "end-in-itself" formula of the Categorical Imperative derived from the "universalization" formula? The text of the *Groundwork* seems to suggest this order of derivation. Yet the logic of Kant's position calls for a different reading. It is no accident that no one—Kant included—has been able to adequately derive the end-in-itself formula from the universalization formula. However, things will start making sense if we reverse the order—more precisely, if we realize that "universalization" is not the primary moral step for Kant, and is not a formal device merely, but is based upon a first-person commitment and resolution of the will: to recognize every other rational being as "end in itself" and an equal moral partner. The self-universalization of the will involved in this recognition is also the basis for demanding it to universalize its maxims.

As long as one takes the test of universalization to be a calculus only, a logical operator, and to be the primary moral factor, no "derivation" of

the end-in-itself formula is conceivable. By what miracle is a logical flaw (contradiction) transformed into moral offense? Or, to put it differently: What is the moral value of universalization, which makes it the cornerstone of Kantian ethics? Why ought we to put our maxims to the test of universalization, rather than any other test?¹⁰ These queries will dissipate if we give the concept of “universalization” its specific Kantian meaning, not as a neutral logical operator but as a first-person act of the will that implies from the outset the recognition of all human persons as ends in themselves, and the commitment to treat them as such. This commitment of the will is the basic moral underlying universalization. The reason why a moral person ought to be able to universalize her maxim is that, thereby, she articulates her recognition of all human persons as equal moral subjects and her disposition to act from that recognition in all morally relevant situations. It is this recognition and disposition that endow moral value on universalization and make it necessary. And, given this reading, the two main formulae of the Categorical Imperative become, as Kant claims, equivalent and derivable from each other.¹¹

Choosing Bondage Cannot Be Autonomous

Because the universal must be realized through a particular will freely adopting it, no “coercive reason” is logically conceivable in Kant’s theory of rationality (unlike Rousseau, or Marx). An individual following the dictates of reason by merely conforming to an external norm behaves irrationally. By the same token, it is logically impossible in Kant to renounce the autonomy of one’s will while claiming (as some autocratic thinkers, religious and political, do) that this abnegation is an autonomous act, because it is performed of my free choice. Acting “from the law” does not only mean “acting from conviction or with a conscious intention.” It means acting from the recognition that the law expresses one’s own rational self (which is why it is binding) and realizes one’s personal autonomy. Where this recognition is lacking, or one’s autonomy is renounced or forfeited, the act is neither rational nor autonomous even when based on free choice. For example, if one chooses to submit oneself to the commands of God the choice is, by definition, heterogamous from the outset. This highlights the Kantian distinction between freedom and autonomy. Kantian autonomy does not mean free choice, but self-choice—the will as choice choosing itself as rational, and thereby constituting its autonomy. It cannot forfeit an autonomy that has never been constituted.

Kantian autonomy (unlike free choice) does not exist from the outset. It is gained at the end, when free choice has been used in a certain way (when *Willkür* chooses *Wille* as its genuine self). In all other cases, especially when free choice is used to abnegate itself, no autonomy has existed (or even remotely approached), and the act, although free, is fundamentally heteronomous.

The Dual Structure of Willing

Let me now sketch the structure of willing in Kant, as it emerges in his mature work.¹² As finite rational beings, we participate in two radically heterogeneous worlds, phenomenal and noumenal, the world of nature and the world of moral-practical reason. Consequently, our willing is split in two heterogeneous principles. The natural inclinations strive for a goal called “happiness,” a vague global objective that Kant construes as the maximization (in quantity, intensity, and duration) of pleasure and utility. Parallel to that, our rational will (*Wille*) wishes to realize morality (practical rationality) for its own sake; its totalizing objective is called “virtue” and construed as a lifetime of continual moral conduct. These two parallel interests or modes of striving are part of the human ontological makeup and cannot be eliminated. Every human will (Kant optimistically believes) aspires to be both happy and moral; the many who opt for evil are actually failing to actualize the good they latently admit. In Kantian idiom, no human will is “demonic” (pursuing evil as an end in itself), and none is “animal” (lacking a moral sense altogether). Nor, as finite creatures, can we possess a “holy will” that always does the good as a matter of course, and knows no conflict or opposition from a rival principle. Human beings do good against the competition of evil, and do evil while knowing they transgress what their own will admits. Even the greatest criminal (again Kant’s optimism) implicitly recognizes the validity of the moral command he chooses to break, and thus, at least dimly, must experience a sense of guilt.

On the other hand, each of us is a creature of nature striving for happiness. This striving cannot cease; it persists in the most moral of persons. When acting morally we do not eliminate the striving for happiness; we only reject it as a ground for decision, and make our move—as we can—from the awareness of moral duty alone. Yet the striving for happiness has not desisted; it has only been neutralized as a basis for action, and thereby transformed into a mute, ineffectual hope. Having (in choosing

morality) disregarded the goal of happiness (and sometimes acted against it), we can only hope that sometime, somehow, the goal of happiness will also be fulfilled.

Later I shall survey the major shift that this persisting hope introduces in Kant's practical philosophy. But first we must examine the precise nature of moral conflict in Kant.

The Actual Moral Conflict

The lack of holy will makes morality a matter of inevitable conflict (articulated by the concept of duty and ought). Who are the partners to that conflict and what is their terrain? It is often believed that the conflict goes on between the free will and the natural inclinations, each trying to suppress the other in a direct battle of forces, until the stronger emerges victorious. But this is not Kant's view at all. First, the competition is not between a natural force (the brute inclinations) and a free force (reason), but between two types of "maxims" and dispositions, by which Kant understands freely designed projects and life-plans; and secondly, the Kantian model allows for no direct or actual clash of forces. It is more like an election campaign in which two candidates are competing over who will be placed in the executive role of determining the state's actions, and who will be excluded from that role.

This presupposes an electing power, which Kant postulates in the form of *Willkür*, the freely choosing will. The two candidates are *Wille*, the moral-rational will, and the will to happiness, which translates the natural striving of the inclinations into a free maxim or life-plan. When a morally relevant situation arises, the striving for happiness and the moral will each suggest a guideline for action (a maxim) and the sufficient motivation to carry it out.¹³ Each offers its own set of maxim-*cum*-motivating as candidate for actually determining the agent's decision and action—in Kant's term, for becoming the "determining ground" (*Bestimmungsgrund*) of her (single, unified) will. The two rival motivations are always mutually exclusive, in the sense that only one of them can be adopted into the will's determining ground, in which case the other is automatically rejected or ejected from that function.¹⁴

The moral conflict in a person's life takes place between these two potential, mutually exclusive sources of action, whose competition concerns the question: which of the two is to enter (be adopted into) the will's *Bestimmungsgrund* and which is to be ejected from it?¹⁵ The arbi-

ter in this competition is the agent's will functioning as *Willkür*. Its role is to face the alternatives and freely choose between them, i.e., choose between Good and Evil.

The foregoing analysis shows the fallacy in two current (and usually joined) beliefs: (1) The belief that moral conflict in Kant takes place between free causes competing against natural causes: the actual conflict is between two equally free causes (reasons). (2) The belief that the natural inclinations are pitted (as perhaps they are in Spinoza, or in Plato) in direct, actual combat against the moral will. Actually, the two compete by proposing (freely designed and) mutually exclusive motivations for the agent to adopt as the determining ground of her decision and action. When a candidate motivation is adopted, its moving power is engaged and it becomes the actual and only motivation; and if excluded from the will's determining ground, then its moving power is neutralized and plays (ideally) no role in the person's following behavior. The dichotomy allows of no "mixed motives" as intermediary, because any purported mixture will revert to the pole of heteronomy and will, in this sense, be homogenous; thus the very concept of mixed motives is, strictly speaking, inapplicable in Kant.¹⁶

Reason and Inclination Can Concur

Our analysis also helps to dispel another common fallacy. It is often argued that one cannot be moral in Kant unless one frustrates the natural inclinations. But this view (started by Schiller) contradicts a major pillar of Kant's theory—his claim that pure reason has sufficient motivational power to generate a moral act regardless of what utility and the inclinations recommend. Thus, the shopkeeper in the *Groundwork* is able to return the child's correct change out of a pure sense of moral duty, even while knowing that this will also help his commercial reputation. Ontologically, he can act from one kind of motive in disregard of the other. In cases where both morality and prudence recommend the same act, one may be uncertain which of the two mutually exclusive motivations was actually operative; but this, to use a Kantian jargon, is a problem of *ratio cognoscendi*, not *ratio essendi*: it concerns the limitations of our self-knowledge, not the ontology of the act itself.

Since no actual battle of forces rages between the inclinations and pure reason, the question "who is stronger?" becomes immaterial. Even supposing (as some "realists" might) that usually the inclinations carry more

motivational power than reason, *Willkür* is free to disregard the inclinations altogether and choose *Wille* as the *sole and exclusive* ground for action. In that case, even reason's "lesser power" (assuming the comparison makes sense) suffices to generate a moral act; and it is spared a war of attrition against the passions, because they were barred access to the potential battlefield. Thus, my construal of Kant's model of the will, because it does not depend on quantitative considerations, makes Kant's claim that pure reason can motivate action sound more plausible than it usually does.

Wille and Willkür

The foregoing allows us to delineate more precisely the roles of *Willkür* and *Wille* in more precise detail. The main background point is that in Kantian morality, the decision is made between two rival sets of reasons, and is performed in both cases within freedom. One set of reasons is embodied in *Wille*, the rational will, which latently recognizes the validity of the moral law, based on the equal status of all human beings as ends in themselves; and the other set is derived from the natural inclinations striving for happiness. (Each maxim, of course, is spelled out in further particular reasons applying to the situation.) The arbiter is *Willkür*, the power of free choice, which Kant postulates as the minimal or "negative" base for freedom. In order for a moral-rational act to be generated, *Willkür* has to choose *Wille*, the rational will, as the sole ground for action. It is not enough to recognize the validity of *Wille*: we do so anyway, latently, all the time; one needs in addition a special act of resolution, in which *Willkür* endorses the proposal of *Wille* and make it the agent's *actual/effective* reason and motivation. This act is not arbitrary but has a good reason: *Willkür* adopts *Wille* as the agent's authentic will and as the only adequate way to realize *Willkür*'s own intent of self-determination (see below). Yet it is essential to realize that *Wille* does not produce a moral act just by providing reasons in its favor: a good reason in Kant's model of the will is not sufficient to lead to action because the agent remains free to reject its rationality (or to reject rationality altogether).

Hence there must be an additional act in which the agent, through *Willkür*, activates the recognition of the moral law latent in *Wille* and makes it the determining ground of her act. This is the crucial moment of "intention" in Kantian morality, without which the concepts of responsibility (acting "from" the law and freedom itself) lose their meaning.

Willkür and *Wille* as Failed and as Adequate Self-Determination

Above I described this act as the particular will's self-universalization, and as making the universal interest of reason our own personal interest. A parallel description, more suited to the present issue, is to say that, in order for a moral act to arise, *Willkür must choose itself in Wille*. This formula highlights the fact that a *Willkür* must still opt for *Wille* and can refuse to do so; which means that, in Kantian morality, even the strongest rational grounds are insufficient: there remains, as Kant says, the "in-scrutable" source of freedom that no rational ground can capture. Also, saying that *Willkür* chooses itself in *Wille* indicates they are two aspects or functions of the same individual self or will. Finally—this is the novel point—the formula emphasizes the moment of self-determination, which underlies both *Willkür* and *Wille*.

Kantian scholarship has wondered how the two are related, and why *Willkür* should opt for *Wille* rather than for maximizing pleasure and utility. I think the answer should be sought in the interoperation of freedom as self-determination. *Willkür* as free choice is essentially an attempt by the will to determine itself rather than be prey to foreign causes; but the attempt fails, because mere *Willkür* is empty, and provides no "self" that can determine itself. The self in Kant—this is his revolutionary critique of the Cartesian subject, both in knowledge and in practice—does not exist at the outset but *constitutes* itself through its universally-oriented operation. Therefore, in order to realize what *Willkür* sets to do—accomplish self-determination—the will as particular must adopt a universally oriented standpoint, the one inscribed in its own structure as *Wille*; in this way, Kantian autonomy is the adequate realization of the notion of freedom implicit in *Willkür*. To be what it is—to accomplish what it seeks by its own intent and definition—*Willkür* must choose itself as *Wille*, must realize itself in it.

In calling *Willkür* "negative freedom" Kant misled certain readers, who interpret *Willkür* as the mere absence of external causes. This reading fails to see *Willkür*'s positive claim to self-determination and relegates it to mere indetermination.¹⁷ Human willing, however, is more than a spontaneous cause; it implies a *first-person self* that intends to determine itself. Thus the essential intent of both *Willkür* and *Wille* is self-determination, but in *Willkür*'s case the intent defeats itself because it assumes an empty (or impoverished) and incoherent form. In exercising mere *Willkür* I trail the external circumstances and causes that happen to cross my way,

and give them my formal assent or dissent, but do not give myself an auto-engendered content or sense of direction. My choice adds the imprint “mine” to diverse occasional contents, but in a deeper sense there is nothing genuinely *mine* in *Willkür*’s structure-less assent and denial, or in the contents it happens to pick. Left to itself, *Willkür* functions as a mere subjective focal point, an abstract “owner” of something external that neither expresses nor constitutes my selfhood, but remains contingent to it. On the other hand, *Wille* as such is not an actual self either, but a potentially genuine self that would become actual if chosen and adopted by *Willkür*. Thus the goal they inherently share cannot be realized by either of them separately: *Willkür* alone frustrates it, and *Wille* alone remains inactive. Their common goal—self-determination, and the very construction of a self capable to determine itself—is realized only when *Willkür* chooses itself in *Wille*. And that means: when it translates its intended self-determination in terms of a rationally-mediated autonomy, obeying a law one legislates oneself for all rational beings, and attaining a sense of individual selfhood through self-universalization.

The need for *Willkür* to choose itself in *Wille* thus transcends morality. It sets the Kantian conditions for selfhood and practical rationality in general. If *Willkür* chooses the inclinations rather than *Wille*, or if a rational-looking act conforming to *Wille* is performed without *Willkür* having actively chosen it, then both acts are, in Kantian terms, *irrational* acts. The former misses reason altogether, and the latter is anti-reason disguised as reason—either a case of conformity or of outright oppression. It is a case of universality without *self-universalization*, a *simulacra* of *Wille* which I cannot recognize as mine and must experience as alien. The varieties of this phenomenon, from mindless conformism to bureaucratic oppression to outright political despotism, are well illustrated in modern political history (since Robespierre), but flatly contradict Kant’s logic of action. “Oppressive reason” is an inner contradiction in Kant’s terms, a conceptual impossibility; although repression in reason’s alienated name has become all too prevalent.

The World-Shaping Will

So far we discussed the individual will, engaged in a moral career possibly leading to “virtue.” But Kant expands the will’s object and terrain: not only the individual person, but the world at large must become morally good. The “best of worlds” is not, as Leibniz thought, the actual

world created as a pre-established harmony by God's will, it is a world to come—a moralized world (a harmony of nature and morality) to be created by the human will in future history. The moral will now wills itself through the mediation of the empirical world it confronts: it wants to see itself embodied in the social and historical world and reflected to itself from it.¹⁸ In other words, the will wants to determine itself by determining the shape of the social world which is its construction and a projection of its own rational self (just as the object of knowledge is an objectivized projection of the finite rational intellect).

How, systematically, does this passage from the personal will to the world-shaping will take place? An important lever is the duty "to further the highest good in the world." Kant first construed the highest good as the conjunction of the two objects of the personal will, morality and natural happiness. The two can be conjoined without contradiction only if the agent continues to work for morality's sake, not for the sake of reward, and the course of the world rewards her. We have seen that the striving for happiness, even after it had been rejected as a basis for action, persists as a passive, ineffectual hope. Kant insists that the moral agent *deserves* happiness—there is a *moral duty* to reward the morally worthy. To whom is this demand addressed? No longer to God, as in the religious tradition, but to the human will. Since the actual world often harms the just and benefits the wicked, the human will has the task and duty to *change and reshape* the world—to create a new social order whose global patterns (in politics, jurisprudence, education, religion, international law, etc.) will reward the morally worthy, and enhance the overall well-being of an expanding moral "kingdom of ends." Thus, solving the old problem of theodicy reverts in Kant from God to the human will, and becomes a historical task leading from personal morality to a world-shaming project.

But this is only the first step. Reward to the worthy is a paradigm case for the broader idea of a moralized world. Reward designates a natural event occurring in accordance with a moral law: it indicates that a certain area of the world has acquired and exhibits a moral shape. Generalizing this pattern, Kant reinterprets the concept of "the highest good" as a synthesis between inner and outer (subjective and objective, personal and institutional) morality, accompanied by universal well-being. Thus, using reward to the worthy as a starting point but going beyond it, the older Kant shifted his emphasis to the more comprehensive issue of embodying inner morality in external institutions.

Another lever to the world-shaping will is independent of the issue of reward.¹⁹ The objectification of morality in the outer world—in Kantian jargon, a synthesis of morality and nature—is actually present in any moral act, however local. A moral agent changes the shape of some local bit of world by making it work according to a moral law. She endows a natural event with moral significance and thus reshapes it. Still in other words: she objectifies morality within some natural object. Thus every act “moralizes” the world in its own miniature way. But reason, in Kant, has a totalizing dynamic: it cannot be satisfied with one-time, fragmentary steps, but drives on toward the creation of ever larger wholes. The moral totality is the “highest good,” construed as a world governed by inner and outer morality and providing general well being. This becomes the regulative idea of history, which is the task of the human will to accomplish. God has no role in this construction, except as an “anti-futility postulate,” assuring the human will that it can plausibly succeed in imprinting moral reason upon an alien nature, although ontologically the two domains are radically heterogeneous.

Moral history has become so important in the latter Kant's work that, in the *Critique of Judgment*, he sees in the will's moral-historical ideal as “the final end of creation.” The will endows being itself with its only possible global meaning and justification. The world cannot be given a global explanation in terms of its creation or beginning, only in terms of its moral goal, which is determined by the human will. The world exists for the sake of the moral system that it has to become through the human will. Being has no value in itself; it is justified by its future moralization; and the agent in making it so become is neither the divine will nor the human intellect, but the human will alone.

Is this new, towering view of the will compatible with the doctrine of the *Groundwork* and the Analytic of the second *Critique*? It is compatible with the core idea, but expands it beyond its original meaning and scope. As mentioned, the will now wills itself through the mediation of its embodiment in the world, and this makes it an active, not a reactive will, a will that does not only respond to circumstances but has to initiate global changes in the world; and one that recognizes the moral import of the empirical world, construed either in the narrow sense of happiness and reward, or in the broader sense of moralizing the external, empirical world at large. The passage from the first to the second stage of the will represents partly an explication of basic Kantian morality, and partly a shift in Kant's personal emphasis and interests. The result, in any case, elevates

the role of the will—the new, revolutionary Kantian will, which is now the source of both the Good and the Highest Good—to historical and even cosmic proportions. It is, prior to Hegel, the climax of the modern, rationalist “history of the will.”

Notes

1. I hereby suggest a modification of Weber’s famous thesis. According to Weber, the believer demanded a sign that he was elected, and found it in worldly success; not very logical but, says Weber, a plausible psychological reaction. My modification bypasses the incoherence involved in the sign-issue. It simply states that the will, knowing its impotence to effect anything in the next world, turns its energies almost exclusively to this world. The motive is not a search for sign of election, but for self-expression, for finding a proper new terrain for the will’s energies.

2. In the last decade there has been some interest in the “History of Ethics” (the better works here are by Jerome Schneewind), yet—as the name indicates—this interest concentrates on moral theory rather than will, and the self-image of the human person (or “practical subject”) they imply.

3. We can only explain how things work in the world, but not why these things—or the world at large—are what they are; even the number and quality of our a priori concepts is inexplicable and, to us, contingent.

4. *Kant and the Philosophy of History* (Princeton: Princeton University Press, 1980, 2nd ed. 1988); henceforth *KPH*.

5. I have offered most of the analysis in the paper: “Kant’s Practical Reason as Will: Interest, Recognition, Judgment, and Choice,” *The Review of Metaphysics* 52:1998, 274–94, and partly (in what concerns the “world-shaping will”) also in *KPH*.

6. Kant allows for human finitude only in that our will does not automatically follow its own absolute law, but has to opt for it and can deviate from it; this distinguishes us from the infinity of the divine will (of a “holy” will). But this, to me, is only partial (and rather conventional) recognition of human finitude. Kant failed to see that with God departing from the founding grounds of reality, knowledge, and normativity, the very absolute nature he ascribes to moral commands is no longer valid; however, if it is the human mind—and will, and interest, and tradition, and its critique—that inevitably constitute moral and political norms, then these are not thereby rendered necessarily senseless or invalid, because human finitude requires a new concept of validity, rejecting the perspective of *sub specie aeternitatis* and admitting temporality as a positive ingredient of the moral domain: moral validity does not require absolute timeless features. Kant’s insistence on the absolute, atemporal nature of man-made morality failed to drive

secularization to its consequent outcome. And it encouraged the visions of absolute secular utopias and salvations that have so far proven to be false Gods. (I mention this as a personal comment, which does not affect the paper itself, in which I try to interpret Kant according to his own standpoint.)

The *causa sui* structure of the pure human will fails to do justice to human finitude in assuming that absolutes can be drawn from it in the same way that they used to be drawn from God's authority. This is a failure to drive secularization to its consequent outcome—namely, that with the passage of authority to the human mind, as both intellect and will, no absolutes, no utopias, no ultimates are available any more, although this does not deprive the human moral norms and products of their value. The view that it does is itself an illicit remnant of the old theological (absolutist) way of thinking. (For a longer critique, see the Epilogue of my *Spinoza and Other Heretics*, vol. 2, *The Adventures of Immanence* [Princeton: Princeton University Press, 1988].)

7. I refrain from saying: "of what ought to be" because this handy ending misses an essential Kantian stress. The postulation is ultimately justified by the desirable state of the will—its fulfilling its inherent structure—rather than by a desired state of the world. The world is, rigorously speaking, the will's medium (and terrain) in actualizing its own inherent form and Ought. (This analytic situation justifies much of the criticisms voiced against Kant, and at the same time highlights the distinctive moral character of, at least, his basic moral stance as given in the *Groundwork* and the Analytic of the second *Critique*. It was, however, modified and even transcended in Kant's theory of the "Highest Good" [as distinguished from the Good *simpliciter*], and the ensuing theory of the world-shaping will [and moral history] implied in it.)

8. See *Critique of Pure Reason*, "The Architectonic of Pure Reason," A839/B867. For a fuller discussion, see *KPH*, chapter 6.

9. The concept of the "interests of reason" is fundamental to Kant's enterprise, and the phrase *Interesse der Vernunft* recurs several hundred times in his works in one variety or another, although most commentators, embarrassed by its apparent incongruity, tend to sweep this concept under the carpet. A notable exception is Gilles Deleuze's small introduction to *Kant* (in the French series *Que Sais-Je*). For a fuller treatment of the question see *KPH*, especially chapter 7. Later books stressing this crucial issue include: Richard L. Velkley, *Freedom and the End of Reason* (Chicago: University of Chicago Press, 1989); Roberto R. Aramayo, *Crítica de la Razón Ucrónica* (Madrid: Tecnos, 1992); and Susan Neiman, *The Unity of Reason* (New York: Oxford University Press, 1994). See also Pauline Kleingeld, *Fortschritt und Vernunft* (Würzburg: Königshausen & Naumann, 1995). A recent book that would have profited from this outlook is Christine Korsgaard's *Creating the Kingdom of Ends* (Cambridge: Harvard University Press 1996), esp. chapter 4. Korsgaard, like many English-language commentators, works primarily with the *Groundwork* and bases herself on localized textual analysis. Even on that

piecemeal evidence, she reached the right conclusion (namely, that Kantian morality does not contradict the setting of an end); but I think she could have drawn that conclusion in a more fundamental way—and also less painstakingly, with less need to overcome the conventional bias, hers and others’—had she started from the overview question of what rationality means in Kant, and how it is related to reason’s “essential ends” and “interests.”

10. For an elaboration of this matter, see my “Kant’s Practical Reason as Will,” from which these queries are quoted.

11. More details are given in “Kant’s Practical Reason as Will.”

12. I mean, beyond the *Groundwork*, and including the book on religion.

13. By using the term “a morally relevant situation,” I mean to imply that not every move in life should be subject to moral judgment. There is a host of trivial actions that normally are morally indifferent—though they may become morally relevant under certain circumstances. The Categorical Imperative is used to test the maxim behind a concrete option we are considering when a situation calls for moral judgment.

14. It is “ejected” if already occupying that function, while the agent undergoes a change of heart; otherwise it is simply rejected.

15. It may seem that a case of “mixed motives” allows both to coexist in the *Bestimmungsgrund*, but analysis will show this to be impossible. There is no middle ground in Kant: all mixed motives belong to the immoral side of the dualism.

16. It can serve in looser contexts, such as education, or in analyzing transitions in someone’s itinerary, but cannot be coherent *per se*.

17. This reading is implied, e.g., in Onora O’Neill’s account of the relation between *Wille*, understood as autonomy, and *Willkür*, characterized basically as freedom from alien causes; *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), 53, and again 76.

18. This is a new variety of the will willing itself through its self-transcendence, except that now it transcends itself not empirical (social) world in whose evolving shape the will wants to see its own laws and ideals (a “kingdom of ends,” autonomy-based jurisprudence and education, free political constitutions, a world confederation ensuring peace, etc.) reflected and being gradually realized in history.

19. I seem to have overlooked this argument in *KPH*.

Moral Mysticism in Kant's Religion of Practical Reason

Joseph P. Lawrence

I

One can hardly imagine a more decisive dismissal of mysticism than the few sentences in Kant's late essay "The End of All Things" (1794), in which he lampoons the Chinese philosopher Lao Tse. He was the author, according to Kant, of that

monstrous system of the *supreme good*, which is supposed to consist of *nothingness*, i.e. in the consciousness derived from annihilating one's personality, of *feeling* oneself flowing into and being swallowed up in the abyss of the divinity. In order to have prior sensations of this state, Chinese philosophers exhaust themselves in dark rooms with their eyes closed trying to think and have sensations of their nothingness. From this tendency comes the *pantheism* of the Tibetans and other Eastern peoples. Its metaphysical sublimation produced, with the same result, *Spinozism*. Both have a close familial relation to the very ancient system according to which all human souls *emanate* from the divinity and are ultimately reabsorbed into it. All of this for the sole reason that men might yet have some place in which finally to enjoy *eternal rest* . . . but it is really a concept with which the light of reason goes out and all thinking itself ends (8:335–6).¹

In light of this critique, it is remarkable that a few years later, Schelling began developing Kantian insights into a broad tapestry that each year appeared progressively more Oriental in outlook. Indeed, Schopenhauer followed by making the move completely explicit, promulgating his own project as a synthesis of Kantian philosophy and Buddhist wisdom.

What, if anything, did Kant himself contribute to these developments? Was Schopenhauer really right in suggesting that Kant's reduction of objective existence to sheer phenomenality was the same as the claim of

the ancient Indian sages that everything is Maya, the illusion of difference that pours forth from the One, but only in order, after being cleansed by the sufferings imposed by time, to be resolved back into it?

The answer, curiously enough, may lie in the project of Kant's old age, that of interpreting Christianity in the light of "reason alone." Kant's attempt clashed with Christianity's reliance upon a highly specific and historically limited revelation, one that culminated in the absolutely unique person of Jesus Christ. The appeal to the universality of reason would seem to have more affinity with the Upanishad appeal to the universality of the "One," than with the common assumption of the Abrahamic religions of the West (Judaism, Islam, and Christianity) that creator and creature are eternally divided. While Kant himself requires such ontological dualism (the possibility of radical evil) in order to secure the reality of freedom, his rationalism nonetheless pushes in another direction: the ultimate victory of moral reason in the establishment of the Kingdom of God.²

The possibility that finite spirit can will itself instead of willing God, thus securing ontological separation, generates, in other words, no more than the illusion of freedom. Genuine freedom is given only where self-will is sufficiently chastised that it learns to acquiesce to the command of morality: "sacrifice yourself and do what is right." Because the command of morality is simultaneously the command of reason, the attainment of genuine freedom is simultaneously the self-realization of rationality. Whether this "reabsorption" into the Supreme Godhead, the underlying unity presupposed by all reason, is itself the final emergence of reason's light or its dark and awful extinction depends upon which reason we are talking about. Does reason refer only to the rational capacity of a human individual—or does it refer to the moral imperative that is addressed (unconditionally) to that individual?

Eastern religion, often dismissed as "irrational," grows out of a profound intuition into the underlying unity of reality. As such, Eastern religion (taken apart from its many superstitious mutations) is rational religion. Kant himself conceded as much by identifying the principle of Eastern mysticism with Spinozism. It is unity, after all, that constitutes the deep premise not only of analytic modes of rationality ($A=A$), but of synthetic modes as well (A is the cause of B). Synthetic rationality is all about uncovering relationships between things that, on first encounter, appear to be very different. Its deep premise is that unity already underlies difference. The notion that salvation is given in the form of a cogni-

tive realization of this underlying unity (Enlightenment) makes of religion something supremely rational. This is even more so the case when, as in Buddhism, the path to such enlightenment is regarded as something that can be explained. This enables people to take up the work of securing their own salvation, instead of waiting passively for some deity to bestow it in an act of grace. In the autonomy afforded by rational clarity of vision, we find a distinctively Kantian moment.³

To explain the irrational frenzy with which many Western thinkers hold on to reason (in its discursive form), we need to understand that (at least in the modern epoch) the characteristic assumption of the rationalist is that it is reality itself that is irrational. Rationality must be "constructed" and "imposed" upon the looming chaos that is nature. Thus Hobbes. Thus the basis of all technological reason (the privileging of the artificial over the natural). And, as he is usually interpreted, thus Kant: reality has to be "tamed" by thrusting the forms and categories upon it.

There is no reason, however, to assume that Kant shared such a vision of underlying noumenal irrationality. The inaccessibility of the noumenon (which means, after all, the intelligible) is an inaccessibility that only affects the finite understanding, which is irredeemably discursive. Reason itself goes beyond the limits of mechanical understanding by positing the idea of the totality. If the totality is all-inclusive, there remains nothing that can condition it mechanically from the outside. Reason does not exhaust itself in the demand for explanations. Ultimately it can only be satisfied with immediate self-evidence (or unconditionality).

Kant reminds us, it is true, that human beings lack the kind of intuitive capacity that, by uniting reason's conception of deep unity with an actual vision of such unity, would yield cognition. But this assertion that we cannot cognitively grasp the ultimate unity of reality is not at all tantamount to the dogmatic assertion that reality itself lacks such unity at its ground. An intuitive intellect, or God, is for Kant a very real possibility, one that is clearly *thinkable* enough to serve as the primary regulative function of judgment. Moreover, while theoretical reason lacks an organ with which to intuit God, the same need not be said for *practical* reason, since the latter requires recourse only to the intelligible idea of perfection and thus can dispense with the question of its objective existence.

It is tempting, of course, to conclude that, not knowing whether reality "in itself" is rational or irrational leaves us with no choice but to commit ourselves fully to the limited sphere of human rationality. Forays beyond that sphere will have to be regarded as departures from reason. It is in this

light that one might interpret the remarkable statement from the preface to the *Critique of Pure Reason*: "I have found it necessary to deny *knowledge*, in order to make room for *faith*" (Bxxx).

Kant, however, was no fideist. Faith for him involves no Kierkegaardian leap beyond and away from reason. In his *Religion Within the Limits of Reason Alone* he often speaks of "rational faith" and "rational religion." Instead of being the antithesis of reason, true faith is nothing but reason itself transformed and purified through the critical expulsion of its speculative hubris. In other words, Kantian faith is just another name for *practical reason*, the facility to generate (or to partake in) ideas that are luminous enough to inspire acts of moral self-sacrifice.⁴

The impossibility of theoretically demonstrating the objective reality of those ideas assures that they are given only in an act of pure interiority. They must arise, that is, solely from "within." What arises from within is not, however, to be regarded as simply a construction of the "self," for the strength and power of the moral idea is demonstrated by its ability to inspire the free abdication of precisely that self. That within the self dwells a power that cannot only dictate to the self, but even demand its annihilation, is the basis for what I am calling Kant's "moral mysticism." Such mysticism, far from representing the antithesis of religious "rationality," instead constitutes its very heart.

II

Kant discerned in Christianity an anticipation of the religion of the future, a religion of pure, practical reason that, fully realized, would secure perfect harmony for humanity as a whole. His guiding insight was that, just as reason itself is one, "there is only one religion" (*Religion*, 6:107; 98). While *faiths* conflict (and insofar as they give rise to actual warfare), religion itself is inherently universal, arising out of the moral imperative to sacrifice the standpoint of the particular subject. Where religious faiths are contaminated with specific doctrines and deities, they are entangled in idolatry and fall short of the genuinely religious principle. True religion for Kant draws not from revelation (which is historically limited), but from reason alone. Its full emergence is the hope of the future.

Unlike Hegel, who believed that reason could discern truth in Christian *doctrine*, Kant was convinced that the living heart of Christianity could only be realized by freeing it from those objectively constituted

conceptions that facilitated the growth of Christianity as an institution. Beliefs that can be dogmatically fixed are beliefs that can be imposed institutionally. Where this takes place (as in Judaism, Islam, and Roman Catholicism) religion is usurped by politics.⁵ According to Kant's understanding, genuine religion must arise, ever afresh, from within the morally sensitive individual. This conviction forms the basis for a profoundly radical critique of Christianity, one that is more akin to what we find in Nietzsche than in Hegel. The Christianity of the churches has to be destroyed—not in defiance of Christ, but in respectful allegiance to his wisdom and divinity.

To assert that religion is inherently rational is not to suggest that its object can be rationally disclosed. When Kant identifies religion with reason, it is practical rather than theoretical reason he has in mind. In other words, Kant's God (as Fichte and the young Schelling persuasively argued) is a possibility and a project, not an objectively given reality. The justification for conceiving God in this way is twofold. Practically speaking, if God were an objective presence, self-sufficiently grounded in a sphere that transcends the human, then the causality deriving from his presence would, as absolute, nullify the possibility of any causality that might separately (and spontaneously) arise from human freedom. The God who is essentially a project rather than a fact is, on the other hand, fully compatible with the free agency of those who are at work on achieving the project. To act freely is to act in accord to the Categorical Imperative that God must be. Divine causality and human causality are united in the sphere of freedom.

If practical reason requires a nonobjective God as the condition of human freedom, theoretical reason comes to its aid through the critical recognition of its own inherent limit. The highest act of theory is the recognition that a theoretically comprehended God is an impossibility. For the comprehension of God from a theoretical point of view would presuppose a subject that stands apart from the God thus comprehended. But a God that has a free subject standing outside it (constituting it as an object of investigation) would be so limited as not to be God at all. Doctrines *about* God, as assertions of theoretical reason, always fall short of God as a living reality. Theoretical reason thus culminates in the recognition of the priority of practical reason. The only way to know God is to act in accord with his inner command. It is a command that arises spontaneously from within, not one that can be imposed in the form of a statute or decree.

Emphasis must be placed here on the spontaneity with which both command and action arise. Any attempt to derive from the moral imperative a set of fully determinative moral judgments (e.g., “never lie” or “never kill”) that can be relegated to a book of statutes subverts this spontaneity and effectively subjugates practical to theoretical reason.

Moral philosophy taken alone tends to err in this direction by focusing attention on actions that accord to self-constituted rules. Obscured thereby is the necessary indeterminacy of the moral imperative, an indeterminacy that can only be grasped from the religious point of view. Hegel was essentially right when he called attention to the *emptiness* of the Categorical Imperative. What he failed to understand, however, were the positive implications of precisely that emptiness. We stand always in the immediacy of the imperative to do what is right only to the degree that no rule or convention can intervene to tell us exactly what we are to do in a given situation. What speaks in the self must be more than what is disposed over by the self. “Religion,” Kant writes, “is the recognition of all duties as divine commands” (*Religion*, 6:153; 142). True autonomy is not the self-assertion of the finite will, but the movement within that will of a divine (and unconditioned) command. Before examining this point more closely, however, the centrality of Kant’s concern with religion should first be established.

III

The *Critique of Judgment* concludes with a series of sections (82–91) that are remarkable for their inner continuity and reach. Kant begins by conceding that if nature consisted only of inanimate matter the question of purpose would scarcely arise. Mechanical explanations of why there are earth, air, and water are fully sufficient, until, that is, they are regarded as conditions for the emergence of living organisms. At that point, their existence takes on a new dimension—they reveal themselves as serving the higher purpose of life.

Living organisms are different. By virtue of their “indescribably wise organization” they awaken already within themselves the question of purpose. Even the simplest organisms display a kind of inner purposiveness that makes mechanical explanations for why they exist insufficient.⁶ Not only does their internal differentiation render their parts into means for the survival of the whole, but in the simple acts of growth and repro-

duction they display a mode of self-organization that serves to distinguish them from anything that has its cause outside itself.

Still, the question of purpose is never fully answered by the recognition of intrinsically given form (e.g., where the purpose of a plant is regarded as its capacity to produce another plant just like it). A satisfactory answer to the question "for what purpose?" always takes the same general form: "for the sake of what stands even higher." Without plants, animals would starve. The hierarchy that emerges must, however, end somewhere. Teleologically perceived, nature exists that man might exist. "Without man, the whole of creation would be a mere wilderness, a thing in vain, and have no final end" (*Judgment*, 5:442).

In humanity the question of purpose is magnified to its greatest possible extent. For if all nature exists for the sake of thinking humanity, then the question why we exist becomes the question of why anything at all exists. We find ourselves at the crux of the most exasperating of all enigmas: that nature itself includes nothing "higher" means that man cannot be construed as serving any natural purpose whatsoever. This apparent purposelessness of humanity is the real origin of the religious need, for it poses a possibility with which the intelligent and sensitive individual simply cannot live: the whole of reality might be fundamentally absurd.

An answer to this enigma is given only to the degree that humanity contains within itself a purpose that is sufficiently clear to answer the question of meaning. A beautifully blossoming rose does not raise the tormented question "why?"—for its beauty always already provides the answer. In the words of Angelus Silesius, the great seventeenth-century German mystic, "a rose has no why—it blossoms because it blossoms." In section 83 of the *Critique of Judgment*, Kant considers the same possibility with regard to man. Nature's purpose might be regarded as sufficiently realized in the happiness of man, which might well be defined as that which has no need for the question "why."

The problem, however, is that man's "nature is not so constituted as to rest or be satisfied in any possession of enjoyment whatever." The proof of this is given in Kant's pessimistic observation that no thinking person would agree to live life over again from the beginning. While we might indeed strive to make earthly happiness our end (the culture of suburban America is proof enough of that), we do so at our own peril, for we then become "incapable of positing a final end for our existence and harmonizing therewith." Kant shows something of Freud's insight here. Full gratification of the pleasure principle comes at the expense of what grows

out of its frustration and denial: the whole realm of culture. Assuming that even an enjoyment-oriented society will always fall short of "satisfaction," Kant suggests that happiness be deferred for a reason. If nature does not achieve its ultimate goal by making human beings happy, then its failure serves a purpose: human beings must be forced freely to constitute meaning for themselves. What unfolds out of this forced freedom is the entire sphere of culture.

Even here, however, we still have not arrived at the goal. Culture is the product of an elite. Its very existence comes at the expense of entire masses that are kept, Kant says, "in a state of oppression." Culture divides. Different cultures clash. Warfare is the awful result.

Once again, however, failure serves a higher *telos*. Anticipating his later essay on *Perpetual Peace*, Kant writes in the *Critique of Judgment* (5:433): "If on the part of men war is a thoughtless undertaking, being stirred up by unbridled passions, it is nevertheless a deep-seated, maybe far-seeing, attempt on the part of supreme wisdom, if not to found, yet to prepare the way for a rule of law governing the freedom of states, and thus bring about their unity in a system established on a moral basis."

The deferral of happiness merely means that humanity has a higher purpose to serve. But it is not a purpose that is given by nature. It is purpose as project and postulate. "The evils visited upon us, now by nature, now by the truculent egoism of man, evoke the energies of the soul, and give it strength and courage to submit to no such force, and at the same time quicken in us a sense that in the depths of our nature there is an aptitude for higher ends" (*Judgment*, 5:433–4).

Teleological judgment culminates in the moral proof of the existence of God, whereby the key to the proof is that God's very absence (the source of those "evils visited upon us"), that is to say, the very impossibility of establishing a theoretical proof of the existence of God, is what delivers us into the necessity of seeking moral perfection. The suffering of humanity is itself what makes the achievement of moral goodness an absolute imperative. We suffer in order that we might feel compassion. We feel compassion in order that we might live for others rather than for ourselves alone. Nature's catastrophic insufficiency is thus artfully constituted: the very inadequacy of what is gives birth to the idea of what ought to be. Kant's philosophy is religious not because it yields an objective conception of God, but, quite the contrary, because it helps reconcile us to a reality in which God is sorely absent. Genuine religiosity is grounded in the deeply felt need for what can be imagined, but never

realized: human beings who live harmoniously with and for one another. In that imagined community of ends, God would be a reality.

IV

It is this idea from the *Critique of Judgment* that puts Kant's entire critical project into proper perspective. That Kant shares with the mystic the notion of God as *Deus Absconditus* lends a kind of religious urgency even to the project of the first *Critique*: the demonstration that syllogistic reason fails in its attempt to secure objective knowledge of God, freedom, and the immortality of the soul. Reason extended beyond experience gives rise to irresolvable paradox. The principle of causality, for instance, can just as easily be used to show the necessity of an infinite regress as its impossibility. Speculative reason must beware of delivering proclamations about whatever lies beyond the pale of experience.

A man of faith will have no problem with this result, concurring that philosophy and religion belong in two completely separate spheres. A Christian dogmatist, on the other hand, will have grounds for complaint. The critical project puts the seal on the dissolution of the old Medieval synthesis, a dissolution that unfolded from the two simple doubts: first, Descartes' doubt whether experience could ever penetrate behind perceptions, understood as objects of a subjective viewing, to yield an actual knowledge of things themselves; second, Bacon's diametrically opposed doubt whether we can rely on anything but experience, given that reason yields not knowledge, but tautological platitude. If the purpose of philosophy is to establish dogmatic certitudes, then the Christian philosopher will be forced to pretend that the epistemological turn never took place. Even the most reasonable doubt should be ignored: thus does dogmatism reveal its intrinsic lack of faith.

There are, then, two rather obvious ways in which a Christian could respond to Kant's critical revolution. The fideist is happy to accept the imposition of limits on philosophy as a way of setting theology upon its proper foundation in faith and revelation. The dogmatist, on the other hand, has no choice but obstinately to deny that a genuine revolution ever took place. One denies philosophy altogether, while the other holds on to it, but blindly, as if the questionable had not been rendered questionable at all.

Another response is possible, however, for those who realize that philosophy's concern lies deeper than the establishment of propositional

certitude. This third response was, roughly speaking, that of German Idealism. Without question it deserves to be taken seriously by anyone who wants to understand Kant. Just as the contemporary neopositivist comes closer to embodying the Puritanical ethos of Kantian sobriety, the idealist, in the very enthusiasm with which he reads his Kant, more successfully accounts for that remarkable "awakening" from dogmatic slumber with which Kant began an unparalleled decade of philosophical productivity.

The genuine philosopher thus asks the sleeping dogmatist: "why take pride in the proof that God exists if the proof generates no real insight into who or what God is?" If God is defined simply as the cause of the world, then he exists as surely as the world exists. But whether he is anything more than an exploding vacuum is left fully undetermined. When the sources of knowledge—understanding, experience, and reason—are simply assumed as givens, as passive mediums through which we are provided a list of true propositions, then we are left with no more than the empty shell of knowledge. The knower stands so decisively apart from the known that it becomes difficult to see how one can speak of knowledge at all. Thus the "slumber" of the dogmatist: lulled asleep by the repetition of syllogistically secured "beliefs," the dogmatist is locked into a dead world where certainty itself blocks access to the face of the living God. Holding desperately to belief, the dogmatist never learns anything of the faith that, beyond belief, first enables us to let go.

The names Schelling and Hegel should suffice to remind us that Christian thinkers, motivated by the serious attempt to get beyond dogmatic assertion to a genuine knowledge of Divinity, might view the critical turn as an opportunity rather than a danger. If we are not equipped to know the Absolute as a transcendent object standing apart from us, perhaps we can understand ourselves as participating directly in the Absolute's own act of self-knowledge. Letting go (through radical questioning) of the means by which we gain knowledge, e.g., experience and syllogistic reasoning, allows us not only to enter into our own elemental ignorance, but it also opens the space in which the Absolute can manifest itself. Released into the open mind of acknowledged ignorance, the field of intuition is torn open. By being restored into its proper element—profound and serious questioning—philosophy has not received a deathblow at all, but has been restored to life. Only one who is actively searching stands in the possibility of an authentic awakening.

The initial result of Kant's critical endeavor is the stark contrast between two forms of knowing. The first unfolds from the standpoint of the

subject. It is objective knowledge, the knowledge of science. Kant regards its objectivity as secure, insofar as the very form of objectivity is contained within subjectivity itself. The knowing subject has constituted its own object and is thoroughly capable of grasping what it itself has made. To do so requires, however, an act of self-denial. The subject must discard its arbitrary interests and concerns until it realizes the pure form of knowing, that mathematically self-identical structure in which all subjects are one, gazing into the open geometry of space, counting the sequential moments of nows that have been cleansed of all color and content. Objective knowledge is the recovery of mathematical form from the objects of experience.

The second form of knowing is philosophical (or transcendental). It is not directed towards objects of experience at all, but towards the act of knowing itself. It arises from the skeptical refusal simply to accept the threefold instrumentarium of cognition: understanding, experience, and reason. Instead of fetishizing objectivity in the mode of scientism, it uncovers the act through which objectivity is constituted. By doing so it reveals the limit of knowing as such, the impossibility of penetrating beyond the subject-object union to a disclosure of what it is that underlies either the subject or the object taken in itself. The philosophical form of knowing completes itself, then, by freeing itself into the cloud of unknowing, supported and sustained by the impenetrable. Following the antinomies of reason, the universe is infinite in terms of the inexhaustibility of possible objects of experience, but radically and decisively finite insofar as it is limited by what eludes the very possibility of objectification: the eternal One, unbounded, though font of the infinite series of all that can be bound. Physically unlimited; metaphysically limited. In other words, to understand Kant's project in the *Critique of Pure Reason* is to realize that Newton has to be reined in just as surely as does Aquinas. Kant remystifies the world even as he demystifies it.

While theoretical reason cannot lift the veil of eternity, practical reason arises out of receptivity to the unconditional command it places upon us. The God who thus commands can never be objectively known. Practical reason unfolds not as the knowledge that he exists, but as the *will* that he exist, because exist he must. From the Kantian perspective, the God who exists is thus the God of the absolute future. His existence is the project of human history, the establishment (at the "end of time") of the perfectly moral world order. What makes this project possible is the moral sense that lies buried in the heart of humanity: the God who is not, but yet

should be. The living God is the God born forth from our very need for him. If Kant rejects the notion of God as all-embracing transcendental thing, then his purpose is not only to set humanity free, but to free God as well. His aim is renewal rather than patricide. Instead of the libertine's "now everything is permitted," Kant offers instead a challenge: "give birth to God, by doing only what is right."

If Kant's critical philosophy issued a death certificate for the God of old, the securely and necessarily existing God, then only in order to make room for the God of life, born eternally out of the innermost divine essence: self-sacrifice. The death of God is for the sake of all who are called upon to give birth to God. For only the God thus born commands fully and completely the attention of those who, having known the pain of his absence, are able to concentrate themselves fully into the will that he might be.

V

I confess that I have departed from the austere and sober language of Kant. Even so, I can cite precedents in Kant's own writings. Well known is the ode to duty in the *Critique of Practical Reason*: "Duty, thou sublime and mighty name, that dost embrace nothing charming or insinuating but requirest submission and yet seekest not to move the will by threatening . . . but only holdest forth a law which of itself finds entrance into the mind and gains reluctant reverence . . . what origin is worthy of thee?" (*Practical Reason*, 5:87) Even more familiar is Kant's assertion at the end of the same work that two things, the starry heavens above him and the moral law within him, fill his mind with ever new and increasing awe.

Is there anything, however, that we can place beside Kant's own definition of mysticism as "the consciousness derived from annihilating one's personality, of *feeling* oneself flowing into and being swallowed up in the abyss of the divinity"?

The question brings us back to the issue of self-sacrifice. To act morally is to leave oneself totally out of the picture. It is to suspend the dictates of prudence and indeed the entire calculating capacity of reason. Kant's point is a radical one. The utilitarian concern with self-interest is not the sign of a merely inadequate moral theory: it is the sign of evil itself.⁷

Thus the proof of the possibility of goodness must be given through the example of one who has committed the ultimate act of self-sacrifice, not

only laying down his life, but doing so after so expanding his compassion that he has effectively taken on the sufferings of all of humanity. The proximity here to the Buddhist conception is palpable. Christ does not redeem us by infusing something upon us, but by opening a path, showing through his example that morality, which appears to be so antithetical to human nature, is indeed a possibility. The greatness of Christ was that he "opened the portals of freedom to all who, like him, choose to become dead to everything that holds them fettered to life on earth" (*Religion*, 6:85; 79).

It is not only the case that a good person does not aspire to any kind of earthly reward, but, given the extent to which human beings are mired in evil, the mere appearance of goodness brings upon itself savage punishment, awakening "all the persecutions by means of which evil men can embitter life" (*Religion*, 6:81; 75). Christ not only had to die, but he had to be hounded, put to shame, and systematically tortured.

In sharp contrast to the human criterion of worth, how well integrated one is socially, how happy one is with one's lot in life, the moral criterion seems to demand that the saint be a pariah. Goodness must go unrecognized insofar as it is society itself that represents the most vicious principle of moral corruption (*Religion*, 6:93 ff.; 85 ff.). The smile demanded of us by society is one that Christ never shared (unless while playing with the young children). Moral goodness is measured by the willingness to sacrifice *everything*, ultimately not only the connection with one's fellow human beings but with the world itself. It is a willingness that can only be forged where one has made complete peace with death. Where can that peace have been made except in "the consciousness derived from annihilating one's personality"?

At the same time, we know that Kant regarded morality as fully compatible with autonomy. Indeed true autonomy can be realized only in morality. This translates into the condition that reason be fully alive and awake in the act of moral self-legislation. The rational self must be realized, not surrendered. The "light of reason" is not to be extinguished.

Here we may ask, however, whether Kant was right in assuming that mysticism, which according to its own self-interpretation is the awakening in us of a light divine, must involve the opposite, a collapse into utter darkness. Kant's understanding of the matter is rooted in the position of the first *Critique*: while we can think the Absolute, we lack the organ with which to intuit it. Pure intellectual intuition must be an attribute of divinity alone. Reason must restrain its speculative impulse, restricting its gaze to the field of sensibility.

On the other hand, it is equally clear that Kant was not entirely happy with this position. Practical reason must unfold from an imperative that operates upon us *unconditionally*. This, the absolute nature of morality, is not easy to reconcile with the merely hypothetical nature of the moral idea. Duty must be perceived directly as a divine command. How can this be the case if it boils down to acting “as if” God were a possibility, acting “as if” human beings might one day do what human beings can hardly even be imagined doing: utterly disregarding their own self-interest?

In the *Critique of Judgment* Kant discovered an answer to this dilemma. There is an experience, he suggests, in which the mystical “third eye” can indeed be opened. Something can awaken within us “the feeling of a supersensible faculty.” This Kant calls the sublime. By encountering it we discover that we do indeed possess “a faculty of mind that transcends every standard of sense” (*Judgment*, 5:250).

Only in the experience of the sublime can the true “endgoal of nature” reveal itself. Our greatest dilemma is that the goal of nature is not given in nature itself. Insofar as nature fails to provide the conditions for our genuine happiness, we are left living absurd lives in an absurd universe.

There is, though, a way out. It is the sublime that takes us beyond nature. No object of nature, Kant insists, can itself be sublime. Even the “starry sky above” is sublime only to the extent that it can awaken *within us* a faculty that extends beyond the heavens. This faculty is the ability to intuit the sublime as that “in comparison with which all else is small” (*Judgment*, 5:250). If the starry sky extends potentially forever, what could that be in relation to which it would be rendered “small”?

Kant gives the answer by distinguishing (in section 26) between the infinity of the number series and the infinite itself as that which goes beyond all possible number. If nothing is greater than everything, then the infinite itself is this very “nothing,” which is why Kant says the sublime has to be given in the form of utter “simplicity.” That which stands above and embraces the endless series of what can be numbered and distinguished cannot be grasped by the understanding or fathomed by the imagination. It awakens these capacities, to be sure, but as an “abyss” into which the imagination “fears to lose itself” (*Judgment*, 5:258). True infinity’s relationship to the spurious infinite of always yet another has its analog in the structure of a human life. Caught in time, even the shortest life could potentially be analyzed into an inexhaustible series of ever more finely distinguished moments. Yet ultimately this spurious infinity of

moments is swallowed up and consumed by the true infinity of death. Zeno's paradox is resolved by the necessity of that death.

Kant does not put it so succinctly, but his examples point in this direction. His account of the mathematically sublime is followed by an account of the "dynamically sublime." Examples include the tempest and the earthquake—whatever has the power not only to overwhelm, but utterly to eradicate us. It is in this light that we are to understand "our special reverence for the soldier" (*Judgment*, 5:262). The power we face is overwhelming, yet in the experience of the sublime we find ourselves standing up to it.⁸

It is no surprise, then, that the feeling for the sublime does not come easily. At our first encounter with it, we find ourselves not in a state of reverence, but in a state of panic. The abyss that looms before us is terrifying (section 29). For this reason, Kant says that the feeling for the sublime has to be cultivated. The mark of the truly cultured person is the ability to stand in love and admiration before that which sends those who are weak into a state of frenzied anxiety. That the appreciation for the sublime has to be cultivated means that it can be achieved only through hard discipline and *practice*.

Thus Kant seems to have understanding enough for why "Chinese philosophers" might "exhaust themselves in dark rooms with their eyes closed trying to think and have sensations of their nothingness." Until we can feel at peace with the abyss, at home in the uncanny, we are not delivered into the full encounter with the sublime. It is for this reason that Kant calls that encounter the basis for *morality* (section 29). We cannot be ready to sacrifice ourselves, to sever our attachment to the world, until we have grown in the realization that the great abyss beyond is a friendly place after all. Indeed, it is our home: what Plato called the Good. A good heart is constituted by the direct experience, in the sublime, of goodness itself.

The ultimate experience of the sublime is given, according to Kant, in the intuition of the moral law. It is universally accessible to all rational beings (though, as we have seen, its discovery still requires discipline and culture). In the sublime, we pass beyond the field of sensibility altogether:

There is no more sublime passage in the Jewish Law than the commandment: Thou shalt not make unto thee any graven image, or any

likeness of anything that is in heaven or on earth. . . . The very same thing holds good of our representation of the moral law and of our native capacity for morality. The fear that, if we divest this representation of everything that can commend it to the senses, it will thereupon be attended only with a cold and lifeless approbation and not with any moving force or emotion, is wholly unwarranted. The very reverse is true. For when nothing any longer meets the eye of sense, and the unmistakable and inefaceable idea of morality is left in possession of this field, there would be need rather of tempering the ardour of an unbounded imagination to prevent it rising to enthusiasm, than of seeking to lend these ideas the aid of images (*Judgment*, 5:274).

The ground of moral judgment is revealed thus as a *feeling*, uncovered through Kant's analytic of the sublime. It is rooted in a nonsensible intuition that gives rise to "the ardour of an unbounded imagination." The moral law calls for the utter silencing of the calculating ego.⁹ True self is absolute self, unafraid to die by virtue of a profound realization of its own inner secret: absolute self, the very source of life and world, dwells only in the element of highest sublimity, in that pure and empty region where divinity is free of all form, even the naked form of sheer existence. The light of reason shines most brightly where harmony is attained, where the freedom of emptiness grants enough room for each and every part of the whole to develop itself fully, but always under the unwavering condition that it find and respect the limit that not only makes it what it is, but also detaches it from and renders it respectful of, what it is not.

VI

The moral law is thus consistent with the old adage: "a place for everything—and everything in its place." The problem this leaves us with, however, is that there is no blueprint that predetermines where the limit lies. To find one's place, one has little choice but to test one's limits—and in turn to be tested by them. The one irrevocable limit is the necessity of death. Evil consists in nothing so much as the willful delusion that death itself can be pushed aside. What motivates it is more than contentiousness and ill will. At bottom, it arises from fear.

As we have seen, fear is also the first step in our encounter with the sublime. The imagination is stirred by what the mind cannot compre-

hend. Among its products are the demons that hold us in terror and give us the irrational resolve never to die, never to melt into that last abyss.

But a second step follows. If we can learn to stay calm in the face of what provokes so much anxiety, then the sublime can step forth in all of its naked simplicity. The hyperactive surge of imagination that sustained our fear is suddenly purged—so, too, the relentless drive to understand.

What is left is the void of naked essence. It is the no-thing that is greater than everything, and indeed, that gives rise to everything through the movement of its own withdrawal. Its very purity is the result of its eternal self-sacrifice. The universe grows infinitely (or as far as the power of perceiving it can ever be pushed). It is awakened and enticed onward by the infinite expanse that opens out of the divine death. God is only insofar as he is not (which is not knowledge for science, but for life).

The command of such a God can only be delivered by the example he sets. In the willingness to die that pours forth from the inner emptiness of the divine, we discover the capacity to let go of the world.

It is this “letting go” that constitutes the secret inner unity of Kant’s three *Critiques*. It is given in the experience of complete *indifference*: indifference in the objectivity of knowing, indifference in morality’s sacrifice of personal interest, and indifference in the experience of the sublime and the beautiful. The calm in the center of the storm is the realization that all things do have their limit. Emptiness is full to overflowing, but only because each and every form that arises from it is the self-same emptiness. As such, form yields to form. Time itself is the gift of the holy law. The promise of death is resurrection: eternal life, but always in the form of the newly awakened.

VII

The analytic of the sublime is not the only evidence that Kant could venture this close to the limit of words. A very different kind of evidence can be adduced from Kant’s remarkable enthusiasm for the *Book of Job*.¹⁰ I call his enthusiasm remarkable, insofar as it undoes the widespread impression that for Kant the moral law (and therewith God) is given in the form of a fully determinate idea of reason.

A God limited in this way could never have summoned the aid of Satan in exploring the dark recesses of Job’s heart—nor would he have had any need to do so. The divine light of reason would have told him all he needed to know.

The Kant honored by the tradition is the Kant who was self-righteously certain that Abraham should have known, by simple application of the Categorical Imperative, that God could never have commanded him to slaughter his son Isaac (*Religion*, 6:87, 187; 6:82,175). Reference to the determinative law of morality would have made it clear to him a priori, that is, even before the test, that the voice of Elohim is not the voice of Jehovah. Thus, when Job so thoroughly gave in to his suffering that he cursed the God who made him, a Kant armed with such self-righteous certainty would have fallen on the side of the three friends of Job. They were the ones who knew the justice of the creator, knew that God was either punishing Job or carrying out some divine trick of alchemy. If God permitted the poor man to suffer, then it was presumably for his benefit. Had Job recognized this, he would not have cursed his maker.

Yet this is not Kant's interpretation at all. In discussing the *Book of Job*, Kant takes the side of moral inscrutability. He suspends his general disapproval of the old covenant and accepts the book into his canon. By doing so he accepts, it would seem, the awful voice that speaks out of the whirlwind—the God who himself is the Lord of famine as well as plenty, pain as well as pleasure, and death as well as life.

Understanding how this could be the case is crucial for the entire argument I am trying to make. While it would be ludicrous to assert that Kant was first and foremost a moral mystic (he was too much the child of his age for that), I do want to suggest that underlying and even sustaining his elaborate instrumentarium of reason was a genuinely mystical impulse. My purpose, in other words, is not artfully to construct a doctrine of moral mysticism around Kant's philosophy of religion, but instead to disclose the spirit of such mysticism as the hidden heart of Kant's entire philosophy. Nor is this a particularly original claim. In many respects, my argument is simply an attempt to clarify what we always already knew about Kant: that while he refused to attend church services, he did so in ways that were fully consistent with the Pietism of his childhood and youth.

What attracts Kant to the *Book of Job* is what separates that book from the statutory (and hence political) tendencies that Kant decries in so much of Judaism. The message he derives from Job is that it is never enough to follow a table of commandments. The moral life has to be lived from within and not from without. The emptiness of the moral law not only situates it beyond edict or decree, but enables it to enter fully into our hearts, in a way that is simple and utterly unmediated. We do not have to tell ourselves what to do. We do not have to laboriously try to decipher

the implications of a moral rule for action. Purity of heart alone is sufficient, for in that purity the heart fuses with the divine essence. Where anything else is the case, heteronomy prevails. To understand the necessary indeterminacy of the moral law is to realize that Kant comes close to Plato's insight in the *Statesman*: the Good towers above any law that could ever be written. It is accessible only to Spirit, the principle of full awakenedness. The true statesman is the one who knows how to act.

In order to become the recipient of an unspeakable wisdom, Job had to suffer, unspeakably so. His true goodness was not reflected in his initial piety. Here Satan's skepticism is well founded. Where is the virtue when one so blessed by God sings God's praises? No, strangely enough, Job's true goodness was not revealed until he seized the courage to curse God to his face. There, where he cursed the day he was born, the womb that gave him birth, and the God who made it all, he attained ultimate purity of soul, full and complete *sincerity*. Setting aside utterly the lies we all live by, Job looked at reality face to face.

And, by so doing, was redeemed.

Condemned alone were the hypocrites, the so-called "friends of Job," who awakened God's scorn by predictably chanting the conventional pieties. What they were condemned to, one presumes, was suffering of their own—suffering that would undo their every pretence at "belief" until finally they too learned the hard lesson that sincerity, while it will never be rewarded by the world, is yet loved by God. For only the heart that has emptied itself even of the idea of God becomes open enough to give birth to his word.

And, Kant would quickly add, to his *deed*.

Notes

1. I am quoting (with minor alterations) from T. Humphrey's translation of "The End of All Things," included in *Perpetual Peace and Other Essays* (Indianapolis: Hackett, 1983), 99–100.

2. Dualism ultimately implies the possibility of eternal damnation. From the viewpoint of practical morality this has its advantage, for one who acts could conceivably fall into passivity if salvation is guaranteed for everyone in the end (though why the promise of ultimate salvation would enable one to surrender to a billion lifetimes of pain is hard to understand). Nonetheless, morality at the same time forbids us from regarding damnation as a real possibility for anyone but ourselves (it is valid only "as far as each man has to judge himself"); *ibid.*, 95.

3. "It must be inculcated painstakingly and repeatedly that true religion is to consist not in the knowing or considering of what God does or has done for our salvation but in what we must do to become worthy of it"; *Religion within the Limits of Reason Alone*, 6:133; 123. [The last number gives the reference to T.M. Greene and H.H. Hudson's translation of *Religion* (New York: Harper Torchbooks, 1960).] An interesting (and I think generally convincing) reading of Kant as a kind of proto-Buddhist can be found in David Appelbaum's *The Vision of Kant* (Rockport, Mass.: Element Books, 1995).

4. This is a central contention of Lewis White Beck's essay on Kant's philosophy of religion, which can be found in a volume entitled *Six Secular Philosophers* (New York: Harper & Brothers, 1960).

5. Kant's dismissal of Judaism is so biting that it is hard to believe that it was not motivated by anti-Semitic prejudice. His central claim is that "Judaism is really not a religion at all but merely a union of a number of people who, since they belonged to a particular stock, formed themselves into a commonwealth under purely political laws, and not into a church" (*Religion*, 6:125; 116). In his discussion of Judaism, Kant displays his most serious limits as a Biblical scholar. While his knowledge of the *New Testament* is exemplary, his knowledge of what Christians call the *Old Testament* is disturbingly weak. In a word, he writes about Judaism as if Moses were its only prophet. Kant's work on religion would have been much stronger if he had realized that the prophet Isaiah was perhaps the most eloquent voice of the form of religiosity he himself wanted to espouse. Even better, of course, would have been recognition of the same principle (the spirit of religion over against the mechanical observance of ritual) in the *Upanishads* and the *Bhagavad Gita*, works that arose in India at about the same time as Isaiah preached. The basic contours of Kant's critique of Judaism is fully developed in Book IV of *Religion Within the Limits of Reason Alone*, which is an attempt to distinguish true religion from *Pfaffentum*. Whether translated as "Papacy" or as "Clericalism," *Pfaffentum* plainly refers to the Roman Catholic "moment" within Christianity, that which gave birth to all of the terror of the Inquisition.

6. That Kant was wrong in anticipating no Darwin (*Critique of Judgment*, section 75)—not to speak of a Crick and Watson—to serve as the "Newton" for blades of grass is not as devastating a criticism as one might think. For Kant nowhere argues that organic nature is not mechanically constituted. He simply points out that one who is sensitive to its beauty and complexity cannot help but to view it—speculatively—as the product of an intelligent arrangement. Conversely, anyone who declares himself fully satisfied with a mechanical explanation is failing to see something essential: the mystery of life renews, as if under a microscopic lens, the mystery of how anything at all might exist, a mystery that inherently excludes even the possibility of a mechanical explanation.

7. The Kantian critique of modern technology would thus be more stringent than the Heideggerian. The issue is not simply that moral reason is different from

instrumental reason. Beyond this, there is an indictment of our whole way of living that can be dramatized in the context of the impending ecological crisis. In light of that crisis, it is doubtful whether the Categorical Imperative would sanction the lifestyle even of the most austere of contemporary philosophy professors. If, for instance, everyone on the planet were to operate a motor vehicle, then, according to the best information we have at our disposal, the environment would collapse. Given that the lifestyle of a middle-class American or European cannot be universalized, we must accept the conclusion that it is morally wrong. What are we to do when an entire age becomes locked in what Kant would call evil?

8. As one might recall, Socrates adopted the same military metaphor in his *Apology*. He cannot cease philosophizing, he tells the Athenians, even if ordered to do so, for a god (something overpowering) has assigned him to his post and there he must take his stand. Indeed, Socrates provides a much purer example of all of this that does Kant himself. When threatened with inquisitorial proceedings for his own free thinking about religious matters, Kant chose to abandon his post. Presumably, his own experience of the sublime had not been carried quite far enough.

9. The effect upon us of the sublime is to "discharge" the very surge of imagination that it awakens. More than that, it ultimately silences the entire cognitive play (until Reason alone remains). Thus does it appear "to contravene the ends of our power of judgment, to be ill-adapted to our faculty of presentation, and to be an outrage on the imagination" (*Judgment*, 5:245).

10. See Kant's letter to J.C. Lavater from April 28, 1775 (10:176–9), that can be found in A. Zweig's translation of Kant's *Philosophical Correspondence, 1759–99* (Chicago: University of Chicago Press, 1967), 79–83. Kant went on to fully expand his interpretation of the *Book of Job* in his 1791 essay, "Über das Mißlingen aller philosophischen Versuche in der Theodicee" (8:253–72), which has been translated as "On the Miscarriage of All Philosophical Trials in Theodicy," by A.W. Wood and G. Di Giovanni and appears in *Religion and Rational Theology* (The Cambridge Edition of the Works of Immanuel Kant, New York: Cambridge University Press, 1996). Kant's primary concern in the essay is to disclose the danger that institutionalized Christianity, particularly when backed by the authority of the state, tends to produce moral hypocrites like the friends of Job, who see the depth of Job's suffering but are prevented by their own pious "theodicies" from responding to him with genuine compassion.

Kant as Educator: Reason and Religion in Part One of the *Conflict of the Faculties*

Susan Meld Shell

Kant's insights into the strengths and necessity of education, and his insights into its insufficiency and inefficacy, were never reconciled.

—Lewis White Beck¹

It is . . . fitting that a human being expend his life . . . teaching others to live; [this is so, however,] only if the number of students is large enough that the sacrifice of his own life to teaching others . . . need not be paid any regard. For that to be the case, schools are necessary. But for this to become possible, one must cultivate *Emile*. It is too bad that Rousseau did not show how from *Emile* schools might originate.

—Immanuel Kant (20:29)

I. The Idea of the University

Lewis White Beck states the problem of Kantian education in its clearest and starkest form when he says that Kant's "strict moral philosophy has, and can have, no place for moral education."² This is so, not only owing to difficulties with which moral philosophy has struggled ever since Socrates famously asked whether virtue can be taught; it is also and especially the case because the foundation of virtue for Kant is based uncompromisingly on freedom understood as accountability, or the responsibility of a moral person for deeds that are, as such, "imputable" to him. Our knowledge that we are free comes to us via "conscience"—that fundamental '*factum*' or 'deed' of reason that commands obedience to the moral law, and in so doing "announces" to us that we *can* do what we *ought* to do.³ Two difficulties for moral education here arise: on the one hand, the very possibility of external helps or hindrances to moral goodness seems to detract from the absolute freedom of the moral person, who must be

able to take full credit or full blame (it is difficult to see how there can be half-measures here) for his own goodness or evil. On the other hand, and on the assumption that we take full responsibility for our own moral improvement (or lack thereof), it is hard to see how such improvements can come about, given that in order to do so, as it seems, we would already have to be the improved beings we aspire to become. Difficulties of this sort led Kant to insist that moral improvement must finally be understood as a kind of noumenal or 'intelligible' conversion, outside of space and time, by which genuine character is formed, and involving something like a "new creation."⁴

There is yet another problem, suggested by the reflection on Rousseau cited above: namely, how—assuming that moral education is possible—one can sacrifice oneself to it (as does the tutor in *Emile*) without committing against oneself an unnatural injustice. Kant's tentative answer, in this early note, relies on sheer numbers to offset the otherwise disturbing consideration to which *Emile* itself gives rise and which presumably explains the fact that Rousseau himself refused to conduct his own life as does his tutor—the consideration, in other words, that makes *Emile* less a prescription for universal reform than a testament to its impossibility.⁵

Despite these difficulties, Kant devoted his life to education—to which instruction in morality was central—first as a house tutor, and then, for more than forty years, as a private docent and professor at the University of Königsberg, where he eventually rose to the position of Rector. It is therefore likely that Kant's treatment of the university may contain important clues, not only to his understanding of education (and its limits) but also to the meaning and importance that he placed upon his own activity as educator.

Kant's only extended discussion of the university begins, compellingly enough, under the aegis of a factory: "it was no bad whim of him who first seized the thought and proposed its public carrying out, to handle the entire content of learning . . . so to speak by *mass-production*" [*fabrikmäßig*] (*Conflict*, 7:17). Organized, in the manner of factories, by a "division of labor" corresponding to the various branches of the sciences, the university is a "kind of commonwealth" made up of "public teachers" or "trustees." That "only scholars can judge scholars as such" confers upon the university a certain autonomy to perform, through its faculties, certain (still unnamed) tasks, to admit new students, and to confer by its own power [*Macht*] the rank of doctor as a status "universally recognized." The university may thus be thought of as a workshop whose

main business is the production of new workers: a body of doctors for the begetting of doctors. The university, so understood, is a self-replicating (or living) body, albeit one that also creates “unincorporated” scholars, who either labor in workplaces (called academies) or busy themselves as amateurs (7:18).

That the public standing of the scholar continues to express itself in terms of the Latin “doctor” (which literally means “teacher”) signals both the (hidden) wisdom harbored within traditional institutions and their potential for abuse: the title of “dean” [*Decan*] descends (as Kant, with a pedantic flourish, notes) from astrology by way of warfare. We should be especially wary, in the university as elsewhere, of the ever present desire of men to adorn themselves with titles (7:18 n.).

But Latin usage also points to the university’s ecclesiastical roots, which elevates (in descending order) theology, jurisprudence, and medicine above the “lower” faculty of philosophy. The conflict between theology and philosophy, as we are soon to learn, is the “exemplar” of the conflict of the faculties as such. Here, Kant only roughly sketches out his central strategy: to play upon the government’s desire to rule, in order to bolster the independence of the scholarly community *vis-à-vis* church censorship. The status of the people (whom the government wishes to influence) as “incompetent” [*idiotes*] dictates the subservience of practicing priests (along with the lawyers and physicians) to the publicly authorized faculties. Precisely as the producer of such university graduates, whom Kant refers to as governmental “instruments” and “tools,” the university can appeal directly, and on its own behalf, to the government’s preponderant interest in achieving the strongest and most enduring (legal) influence upon the people (7:18, 19). If the government, as user of these tools, retains the right to sanction the higher faculties, it does so only by renouncing any authority of its own to teach, which would impugn the dignity of government by opening it, of necessity, to legitimate scholarly criticism. The lower faculty, on the other hand, which “has no commands to give,” must be left “free to judge everything,” both because reason admits no “order to believe”⁶ and because otherwise truth (which government, too, requires) will not emerge. Government’s interest in influencing the people by the most powerful and enduring means itself dictates (as Kant hopes convincingly to show) that philosophy remain at liberty (7:20).

To be sure, that liberty confines itself within a scholarly commonwealth that is itself, at least in part, an established tool (or toolmaker) of govern-

ment. If Kant had once counted on greater latitude for public criticism ("argue as much as you like, only obey!"), the more recent events evidently convinced him of the utility (or need) of excluding from the relevant discussion those "men of affairs" (i.e., priests, lawyers, and physicians) whose "public" utterances his earlier essay "What is Enlightenment?" had famously and flagrantly included. By Kant's current lights, such "technicians" may not publicly object (whatever their private thoughts) to the official doctrines they have contracted to uphold. Either the excesses of the French Revolution, or his unhappy experiences with King Frederick William II (whose "peculiarities" of temperament are duly noted in the Preface [*Conflict*, 7:5]), or both, seem to have convinced Kant of the need to draw back from any attempt to influence the people directly in those matters which were for him most crucial. The (relative) benevolence and competence of Frederick the Great had led Kant, in that earlier essay, to declare his age one of "enlightenment." The current work permits itself, instead, to describe the current regime as an "enlightened government." That government, which, through the freedom of thought that it allows, is "releasing the human spirit from its chains," now permits to be published a work on a subject for which he was earlier censured and, more importantly, censored. It is this censorship—the very palpable threat of public silence—that, to Kant's great and evident relief, has been lifted. Preventive censorship was an established fact, even under the reign of Frederick the Great, with an exception made for work executed in and published by the universities, which were submitted to the relevant faculties. What was at stake in the affair—both in the writing that led to Kant's promise to desist from further public utterance on religion, and in the promise itself—was nothing less, as we shall see, than prospects for enlightenment as such, and with it, hope. Under Frederick the Great, the members of the Berlin Censorship Commission were themselves men of learning who could be counted on to exercise their authority with prudence and discretion. Frederick William's peculiar susceptibility to the wiles of his ministers and mistresses threatened to replace such men with crude schemers and half-educated *Schwärmer*, drawn (in many cases) from the ranks of local pastors. Under these new circumstances, the university seems to have assumed special importance for Kant as a bulwark against the twin dangers of despotism and popular fanaticism.

The key to the university's success lies in its (uneasy) synthesis of freedom and power. Kant makes no effort to hide the fact that historically the university is a creature of the state: "without attributing premature

wisdom and learning to the government, (we can say) that by its own *felt need* (to “affect” the people by certain teachings), the division that appears to have a merely empirical origin could have come from a principle, *a priori*” (7:21). Like the government, the university is (or may be thought of as) a kind of manmade organism, an historically embodied idea, however crudely it may initially have been grasped and articulated. The force propelling this historical development (in the biological sense) is the state’s “felt need” to exercise control over the populace—control which, precisely because man is free, cannot be accomplished by brute force alone. The division of the higher faculties corresponds, accordingly, to the fundamental incentives by which the government can achieve its end of “influencing” the people: appeal to their eternal well-being, their civil well-being (as holders of property) and their physical well-being (as embodied rational agents) (7:22).⁷

Kant takes advantage here of the peculiar history of the German universities and their primary association with state, rather than ecclesiastical, authority. The “traditional” primacy of theology is objectively correct, from the state’s own point of view, because it allows it to uncover and “steer” the “inmost thoughts” and “most secret directives” of its subjects. (That natural instinct reverses the order of importance allows Kant a wry jab at clergymen, whose hypocrisy, in this respect, is especially damning.) All three higher faculties base their teachings on statutory books (of which the Bible is preeminent), which derive their authority wholly from the government’s ability to command (absolute) obedience without regard to whether or not its statutes are reasonable (7:22–3). A written Bible is needed (as in the case of other statutory books) to supply a constant and accessible norm for the direction of the people. (As for claims as to its peculiar character as the language of *God*, Kant appropriates to himself the uttering of the Word [7:21].) If the higher faculties (and Biblical theology especially) presume to any authority of their own, then they encroach on the province of the lower faculty, where the free play of reason prevails over the former’s borrowed dignity; hence the importance, says Kant, for the sake of theology itself, of avoiding a mis-marriage with philosophy.

II. Peculiarity of the Faculty of Theology

What, then, is the distinctive characteristic of the theological faculty? Biblical theology, Kant says, demonstrates the existence of God on the

grounds that he spoke in the Bible (7:23). A further showing that the latter actually happened not only is unnecessary to religion rightly understood (as Kant will later prove at length), but also an encroachment on the faculty of philosophy, to which the study of history properly belongs (7:23, 28). The theological scholar as such (Kant plays here on the Latin *purus*, and *putas*—terms connoting both chastity and immaturity) grounds himself, rather, on “a certain (indemonstrable and unclarifiable) *feeling*,” on a “supernatural opening of his understanding,” and on “grace,” abjuring all independent appeals to morality, to reason, or to virtue. The theologian’s adherence to pure faith, linked to an indemonstrable and unclarifiable feeling, dictates (if he is wise) a circumspective public silence as to the question of the Bible’s origin, lest he raise doubts among the people that he cannot answer (7:24). Dogmatism, unchastened by criticism, is here, as elsewhere in Kant’s thought, an open invitation to skepticism. Kant harbors few illusions that theologians in practice will remain so “(un)contaminated”; his point is, rather, to indicate what a theological faculty that was thoroughly consistent with itself could claim on its own behalf. The answer is a special feeling for the miraculous [*Wunder*], which Kant will later relegate to the realm of the “aesthetic.” Yet (as we shall see) this feeling is not as far as one might first suspect from the “wonder” [*Bewunderung*] from which pure reason draws its own (subjective) strength.⁸

III. Concept of the Faculty of Philosophy

The mark of the lower faculty of philosophy, on the other hand, is freedom, or the power to judge autonomously, on the basis of reason alone—power that cannot, by its nature, be externally commanded (7:27). Such a faculty is necessary to the others (whom it aids by controlling), because truth is the first and necessary condition of learning as such. In addition, since philosophy, by its nature, seeks not to dominate but only to be heard, this faculty is, like a trustworthy minister of state, above suspicion.

The university so understood gives working effect to the truth through several levels of engagement: the faculty of philosophy’s free criticism gradually corrects the historically established teachings of the higher faculties, which, in turn, produce increasingly enlightened popular practitioners who are prohibited from injecting, in their dealings with the people, their own (less enlightened) opinions and objections. And, in any case,

such objections are likely to decrease, inasmuch as such officials, “becoming more enlightened about their duty, are not repelled by altering their discourse” [*Vortrag*] (7:29). Change can thus come about “without polemics or attacks that only stir up unrest.” Precisely because people at large are uninterested in intellectual subtleties, and “feel themselves bound” to keep up with what state officials announce to them, learned discussion can be carried on without the people being thereby incited to disorder and rebellion. Thanks to the engine of government, duly co-opted, the people’s very tractability (which might otherwise seem to constitute an absolute stumbling block to enlightenment) becomes the vehicle of their liberation.

Rousseau’s *Emile*, Kant once complained, furnishes the model for a retail education; whereas the real task is to devise one that is “wholesale.” Kant’s idea or image of the university as human factory is a proposed solution to that problem.

IV. Philosophic Instruction in the *Critique of Pure Reason*

Reason as both lawgiver and judge calls to mind Kant’s discussion, in the *Critique of Pure Reason*, of philosophic instruction. On its account, one cannot be taught to be a philosopher (which is an unachievable ideal) but only “to philosophize”:

Philosophy is the system of all philosophical knowledge. It must be taken objectively, if by it one understands the archetype [*Urbild*] for the estimation of all attempts to philosophize, an archetype that ought to serve for the estimation of each subjective philosophy, whose structures are frequently so manifold and alterable. In this way, philosophy is a mere idea of a possible science, that nowhere is given *in concreto*, but which one seeks by many paths to approach, until the one footpath, very overgrown through sensibility, is discovered, and the till now so abortive image, arrives at being like the archetype, so far as this is granted to man. Until then . . . man can only learn to philosophize, i.e., to exercise the talent of reason, in accordance with its universal principles (A838/B866).⁹

Prior to the emergence of Kant’s own critical system, the concept of philosophy was merely scholastic, devoted to systematic unity of science for its own sake. Now, however, there “is given” a “cosmic concept” that has

always implicitly grounded the term “philosophy” (or lover of wisdom), especially as personified in the figure of the philosopher as legislator of “the essential ends of human reason”—an ideal in which scientific knowledge is subordinated to the moral goal, or “entire vocation of man” (A841/B869).

Kant makes no claims to be a philosopher in this morally exalted sense. As a teacher of philosophy his task is, rather (and with a view to that entire vocation), to elicit in the student a living confrontation with the dialectic of reason:

However (a) knowledge may originally be given . . . it is still, for him who possesses it, merely historical, if he knows it only at the level and as much as has been given to him from outside . . . whether through immediate experience or narration, or even instruction ([of] general knowledge). He who has in the strict sense learned a system of philosophy . . . knows and judges only what has been given to him. . . . He has formed himself [*bildet sich*] according to a foreign reason, but the imitative [*nachbildende*] faculty is not the productive one; that is, knowledge has not arisen in him out of reason, and although objectively it may be rational knowledge, subjectively it is merely historical. He has grasped and retained, i.e., learned [in the sense of “learned by heart”], and is a plaster caste on which the shape of a living man has been imposed (A836/B864).

Knowledge that is rational objectively is so subjectively only when derived from “sources from which there can also arise the criticism and rejection” of what is learned, and hence only when at home in and able to withstand conflict. The student must reflectively reenact that philosophic state of nature out of which arises criticism, understood as “reason’s true tribunal.” As Kant put it in a famous passage:

In the absence of . . . critique reason is in a state of nature, so to speak, and can make its assertions and claims valid only through war. The critique, on the other hand, which derives all decisions from the ground rules of its own installation, whose authority [*Ansehen*] no one can doubt procures for us the peace [*Ruhe*] of a legal condition, in which our controversy [*Streitigkeit*] can be carried in no way other than by *juridical process*. What ends the quarrels of the former state is victory to which both parties lay claim, and which is generally followed by an uncertain peace, established by a magistrate who places himself in the middle; in the latter, however, the *sentence* must se-

cure eternal peace, because it touches the source of controversy itself. The endless controversies of a merely dogmatic reason thus finally necessitate a seeking of rest in some critique of this reason itself, and on the lawgiving on which it grounds itself. As Hobbes says: the state of nature is a state of injustice and violence, and one must necessarily forsake it in order to submit oneself to lawful compulsion that limits our freedom only to the extent that allows it to coexist with that of others and with the general good (A751–2/B779–80).

The freedom guaranteed by this state of peace entails, above all, the right “to submit openly for estimation the thoughts and doubts that one cannot oneself resolve, without being shouted down as a warlike and dangerous citizen,”¹⁰ a freedom implied by the “original right of human reason, which recognizes no other judge than that of universal human reason, in which each one has a voice” (A752/B780). Such peace requires a “lowering of tone” from apodictic certainty, in moral matters that concern the very foundations of the public welfare, to one of merely practical conviction (A749/B777), and hence a (perpetual) willingness to give skepticism a hearing. The primary purpose of free discussion for Kant is therefore not, as some contemporary defenders of “public reason” would have it, to further the advancement of knowledge on a democratic basis, but a more immediately pedagogical one: to re-create, in each (would-be) rational citizen, that struggle between dogmatic assertion and skeptical doubt which “necessitates” the critical solution. To enter the fray and attempt to shout down one’s skeptical opponents with countervailing proofs (as dogmatic moralists are wont to do) is to keep “youthful reason” in a state of morally dangerous tutelage, a danger that speaks especially for the importance of the academy:

Must not at least the young, who are entrusted to academic instruction . . . be preserved from a premature knowledge of such dangerous (skeptical) propositions, . . . until the teaching one would ground in them is so rooted as to . . . withstand all opposition? . . . Nothing would be emptier or more fruitless than thus to keep youthful reason for a time under tutelage.

This, would indeed, guard against corruption [*Verführung*] for a while. And yet:

Whoever, in order to withstand the attacks of his opponent, brings

only dogmatic weapons, and doesn't know how to develop the hidden dialectic that lies no less in his own breast than in that of his opponent, sees apparent reasons with the advantage of novelty advance against apparent reasons which no longer have that advantage, but rather arouse all the more suspicion of abusing the credulity of youth. He believes he can in no way better show that he has outgrown childish discipline than by throwing aside . . . well-meant warnings and, accustomed to dogmatism, drinks deeply from the draught of poison that dogmatically destroys his principles (A754–5/B782–3).

Dogma, like poison, has its use, but only as a counter-poison.¹¹ In describing the thoughts of an imagined student, Kant also portrays the contours of man's generic adolescence—the “adventurous and self-reliant reason” (A850/B878) that in mistaking discipline for servitude jeopardizes its own self-grounded development.

Kant's own approach to academic teaching is the opposite one of exposing the student to the most powerful skeptical arguments available:

For in order to bring its principles as early as possible into application, and to show their sufficiency even when dialectical illusion is at its height, it is absolutely necessary that the attacks which seem so fearful to the dogmatist be allowed to direct themselves against the student's reason, which is still weak but critically enlightened, and that the latter be allowed to make the experiment of testing the groundless assertions of his opponent, one by one, with reference to those (critically established) principles (A755/B783).

Suspicion—corrupting in the ill-taught student—is saving when youthful reason turns it inward. Properly exposed to skepticism, the student “soon begins to feel his own power [*Kraft*] fully to secure himself against harmful deceptions” so that at last “all illusion for him must vanish” (A756/B784).

Free discussion is thus crucial, less as a means of arriving at the truth (for the most important theoretical and practical truths are already known)¹² than as a means of fortifying reason at a moment of developmental crisis. The student may not neutrally stand aside while dogmatists and skeptics trade blow for blow, but must recognize that struggle in his own breast, i.e., as intrinsic to reason itself. The struggle is thus best described, not as a true polemic, but as a tension or conflict (as in conflict of interest), which (as in the case of the Hobbesian civil contract) is reconcilable with

peace to the extent that each party lays down its claim to absolute possession. In the case at hand, each party must recognize the right of the other party to a voice, and thus lay down its claim to unrestricted knowledge.¹³ This “lowering of tone” (which allows the other to be heard)¹⁴ coincides with the maturing student’s own ability to register the conflict that betrays reason’s competing interests (A466/B494 ff.). Only by exposing, and bringing to a climax, this “natural dialectic” does he come to *feel* within himself the force that actualizes philosophy’s tribunal. It is not enough, in other words, to grasp critical philosophy conceptually; (here, as elsewhere) the juridical condition obtains only when force puts the juridical concept into action by enforcing the law against the conflicting interests of opposing parties.

The struggle against philosophic death—and, especially, against the skeptical hopelessness that, according to Kant, threatens pure reason’s very life (A743/B771)—must be conducted allopathically. Put otherwise, skepticism must be unleashed in order to be vanquished:

Speculative reason in its transcendental use is dialectical *in itself*. The objections that are to be feared lie in ourselves. We must search them out like old but never superannuated claims in order to ground eternal peace on their annihilation. External quiet is merely apparent. The seed of these troubles, which lies in the nature of human reason, must be uprooted; *but how can we uproot it, if we do not give it freedom, and even nourishment, to send out shoots, so that we can thereby discover it and subsequently iradicate it with the root?* Devise, therefore, objections yourself, that have not yet occurred to any opponent, and even lend him the weapons or concede him the most favorable position that he could wish! There is nothing in this fear, but rather *much to hope*, namely, that you will come into a possession that in all the future will never more be disturbed (A777–8/B805–6; emphasis added).¹⁵

This digression can help us see part of what was at stake for Kant in the conflict between theological and philosophic faculties. The development of reason as a talent requires not only the positive instruction of “culture and doctrine,” but, equally importantly, the negative instruction of a protracted and ongoing “discipline,” by which alone reason’s “habitual tendency to disobey” can be extirpated (A709/B737). Reason’s “natural antithetic” wakes it up but also tempts it either to abandon itself to skeptical hopelessness or to fall back into dogmatic obstinacy. The primary justifi-

cation for free and open discussion is not the discovery of new knowledge but preservation of moral truth from skeptical despair. Without the “felt” experience of self-correction, right reasoning falls on deaf ears.

V. The Conflict of the Faculties

Armed with some understanding of what philosophic education means for Kant, and assured that dispute is not only consistent with but also necessary to philosophic peace, we turn to Kant's distinction between an illegal and a legal conflict of the faculties.

A conflict of opinion is formally illegal if either party tries to prevail, not by objective argument, but by subjective appeals to inclination, i.e., through threats or flattery. Faculties engage in conflict in order (*ex hypothesi*) to influence the people by appealing (as all civil engines must do) to the interest people necessarily take in their own welfare [*Heil*]. But the people left to themselves (owing to an aversion to “self-exertion” in which their tractability essentially consists) recognize their welfare not in freedom, as the faculty of philosophy would urge, but in natural ends (perpetual happiness, secure possession, and a long and healthy life). In keeping with their “preference for enjoyment and their unwillingness to work for it,” they invite the higher faculties to provide them with easy, if immoral, shortcuts (such as the view that one may be saved by faith alone) (7:30, 31).

Practitioners will always be such trimming and accommodating “miracle workers,” Kant maintains, unless the faculty of philosophy is allowed to publicly work against them [*entgegen zu arbeiten*], not by overthrowing them, but instead by “countering the magic force” that the people, owing to their laziness, superstitiously attributes to them, “as if by passively giving themselves over to such artful directors” [*Führer*] the people would be “led with great convenience” to achieve their desired ends, and “excused from any self activity” (7:31). The original sin, so to speak, of the people is their desire “to be led”—a deliberate inertness consistent with Kant's comparison of primitive men to sheep and which goes together with a natural disinclination to “exert oneself.” The infantile in man naturally craves a return to the Paradise where wishes are deeds and work (even that required to breathe) unnecessary.

If the higher faculties themselves adopt such principles, they contest the lower faculty *illegally*, by virtue of the fact that they themselves treat

lawbreaking not as hindrances (as they can and should), but as wished-for opportunities. Not only would a government that sanctioned such principles authorize anarchy (and be, in this respect self-canceling); it would also destroy the lower faculty, ending the conflict in a medically heroic sense, i.e., by killing the university (as with the “euthanasia of pure reason” described above) in order to save it (5:32, 34 n.).

The *legal* conflict of the faculties, on the other hand, arises from the necessary gap, given human imperfection, between what reason dictates and what governments choose, through the higher faculties, to sanction (7:32). Government’s own interest in the truth (if only to carry out its end of influencing the people) necessitates a faculty of philosophy as potential—and, given the higher faculties’ inherent desire to rule, uniquely trustworthy—opposition. As such, the philosophy faculty has the right and duty to criticize teachings that are officially sanctioned, be their source reason, history (e.g., Biblical testimony), or “aesthetics” (e.g., “pious feeling of supernatural influence”) (7:33). Although the legal conflict of the faculties is “unending,” it could happen that government will one day come to see philosophic freedom, rather than its own absolute authority, as the better way of securing its own ends, so that the “lower” faculty would become the “higher” (7:35). For the conflict of the faculties is not a war, that is, “a dispute arising from opposing final aims” (for all parties aim at influencing the people); it is, rather, a reflection of the university’s dynamic combination of freedom and obedience, reason and statutory authority. The university, one could say, institutionalizes the tension between knowledge and power, in a manner calculated to bring them into ever closer alignment. It thus replaces, at least to the extent that it succeeds, the merger, at once necessary and impossible, of philosophy and kingship.¹⁶

VI. The Exemplary Conflict between the Faculties of Theology and Philosophy

The relation between the theological and philosophic faculties sheds special and “exemplary” light upon the university as a kind of living faculty¹⁷—a clarification of the “conflict of the faculties” to which the difference between “Biblical” and “rational” theology is central. A *biblical theologian* “properly speaking” is learned in Scriptures with a view to church faith, which depends for its authority on statutes, or the *Willkür* of

another; a *rational theologian*, on the other hand, is learned in the faith of reason, which is “based on inner laws that allow themselves to be developed out of every man’s own reason” (7:36). The latter faith alone may properly be called one of *religion*, which is not (as generally thought) the sum of all teachings understood as divine revelations, but “the sum of all our duties . . . understood as divine commands.” Whereas there can be many kinds of faith, which are ways of representing divine will in sensible form so as to exercise “influence” upon the mind [*Gemüth*], there can be only one religion, which consists in use of the Idea of God, which is “engendered by reason itself,” to give morality “influence on man’s will toward the fulfillment of all his duties.” Religion, then, differs from morality “only formally,” each having the same goal (fulfillment of all man’s duties) but differing inasmuch as morality through religion exercises particular “influence” on the will.

Among church faiths, Kant crucially adds, Christianity is “so far as we know” the most suitable [*schicklichste*] in form for “influencing the mind.” Hedged by the uncertainty (“so far as we know”) of a merely historical assertion, the (alleged) superiority of Christianity as a means of exercising mental “influence” marks the ground on which ecclesiastical and moral faith both meet, and over which the two faculties (from Kant’s own point of view) stake their competing claims to best accomplish that aim.¹⁸

Christianity consists of two “heterogeneous pieces” [*Stücken*]: a “canon,” or pure religious faith, and its “vehicle” or “organon”—which “can be called” church faith. The organon of “church faith” rests entirely on statutes that “must be taken for divine revelation if they are to count as holy teachings and prescriptions for living” (7:37). But the moral duty to “use” church faith, provided that it can be adapted as a “guiding tool” [*Leitzeug*] to moral ends, suggests that the true “organon” [literally, tool] of Christianity may be of an entirely different sort. It is on precisely this question—in what way and in what sense Scripture is to be regarded as a tool to salvation [*Heil*] or “eternal life”—that the conflict between the theological and philosophical faculties will turn.

Although that which matters most—“concepts and principles of eternal life”—cannot, strictly speaking, be taught, i.e., learned from others, their development out of one’s own reason can be “prompted” [*veranlassen*], Kant says, by discourse [i.e., *Vortrag*, literally “something said/carried forward”]. Scripture contains this much (a pure doctrine of moral religion) but also something more: i.e., an historical faith

which can be useful, in certain times and places, to pure faith as a “merely sensible vehicle,” but is not essential. The theology faculty’s *mingling* of the two is just what the faculty of philosophy contests [*widerstreitet*].

To this inessential vehicle belongs a method of teaching that is not a matter of revelation, and that is valid according to the thinking of the times. Certain apostolic doctrines, for example, are to be understood as mere concessions, which do not essentially oppose true religion, to then-prevalent illusions (e.g., demonic “possession”) or to Jewish partiality to the old faith that must come to an end (e.g., “interpreting the history of the old covenant as a fore-image [*Vorbild*] of the new”). As we shall see, Kant is at special pains to destroy this Jewish “piece” (which concerns the entire story of the Jewish covenant, from Abraham’s idol-shattering, to the circumcision and sacrifice of Isaac, to the Mosaic decalogue (“though shall have before thee no graven image . . .”) as a “wearisome” and “agonizing” remnant [*reliquia*] (7:38).¹⁹

From the ensuing difficulties arises an interpretive conflict in which the faculty of theology favors dogma and letter, while the faculty of philosophy defends religion and the (moral) spirit. The exegetical principles by which the conflict can be settled, however, must themselves, as principles, be rationally (hence, morally) grounded.

According to those principles:

1) Morally indifferent scriptural doctrines that transcend (i.e., cannot be grasped by) reason, *may* be interpreted morally. Doctrines that conflict with moral concepts *must* be interpreted morally.

2) Faith in scripture that depends on revelation is not meritorious, nor is the lack of such faith blameworthy.

3) Action must be represented as “springing from man’s own use of his moral forces,” and not as the “effect” of “higher influence.”

4) If man cannot justify himself by deeds (which follows from nature of conscience itself), reason is permitted to assume (lest effort flag) a supernatural supplement [*Ergänzung*].

These principles are calculated to promote a single end—to bring moral exertion to the highest possible pitch. They do this both by insisting on the responsibility of the individual for his own action, and as importantly, by denying traditional church teachings that tend to diminish moral effort. To wit:

1. The claim that Christ is God incarnate (a claim that is both transcendent and morally indifferent) *may* be morally interpreted (so as to render it “practically effective”) as a representation of the rational idea of “humanity in its full moral perfection” (7:39), i.e., as the portrayal, in a sensibly represented form, of the supreme goal of rational human striving.²⁰ Paul’s teaching concerning predestination, on the other hand, which directly contradicts concepts of freedom and the imputability of action (and so of the “whole of morality”), *must* be interpreted away, as a vestige of the “Mosaic-messianic scriptures” (cf. 7:66).

2. The claim that faith in revealed doctrine is not only meritorious in itself, but also lifted up [*erhöben*] over morally good works must be taken to apply to *moral* faith, i.e., “that which improves through reason and elevates [*erhebende*] the soul.” (Such doctrines may, however, be considered as essential to a church faith in which the state, as defender of the public peace, takes legitimate interest. In this case, the teacher should not attack these doctrines directly, but instead warn against treating such doctrines as intrinsically holy,²¹ with a view to their gradual disappearance [7:42].)

3. The doctrine of grace, which, according to the letter of the Bible, claims that man is passively sanctified, must be interpreted to refer to nothing other than “the nature of man, insofar as he is determined to action through his own inner, but supersensible, principle (the representation of duty)” — an original predisposition [*Anlage*] to the good that we did not ourselves establish:

Grace, i.e., hope in the development of the good—hope enlivened through faith in the original disposition in us to the good and through the example of humanity well pleasing to God in his Son—can and should become yet more powerful in us (as free) if only we let it affect us, i.e., let the disposition to conduct resembling that holy example become active (7:43).

It must, in other words, be made clear “that *we ourselves must work*” to develop the original disposition to the good, a disposition that points to a divine source, and is not itself meritorious.

4. Human reason may additionally assume, to avoid a morally enervating despair, that God will supplement what man cannot do to justify himself by his own acts. However (seemingly) at odds with principle no. 3,

this assumption is warranted by the fact that, without it, we cannot grasp [*fassen*] the courage and firm disposition to live rightly—given the limitless nature of our task (i.e., our subjective awareness, intrinsic to human frailty, of our guilt as unremitting) (7:44, 47).

Even the principle of imputation—that one may be credited (or blamed) only for one's own actions, on which Kant's moral system seems otherwise to turn—bends (at least subjectively) to human reason's "need" to prompt itself to labor. This need arises from the necessarily ambiguous voice with which the moral law speaks to finite creatures like ourselves. Conscience is informed by a tension, intrinsic to human consciousness itself, between idea (the final end that conscience tells us we can arrive at, as in "ought implies can"), and its limitless approximation in space and time (the only way we can conceive of carrying out our duty). The idea by which reason sets itself a task is accompanied by the impossibility of conceiving its completion. The idea is thus shadowed by despair—a "wearying" feature that is itself linked to the unreasonable basis of reason's dogmatic employment.²²

VII. Sectarian Division

Christianity so understood is "the idea of religion" (7:44), which is one, necessary, and universal. Church faith, on the other hand, breaks down into sects, which mistake religion's vehicle for religion itself. Nevertheless, to the extent that sects *are in fact* religion's vehicle, they are divisible (in ways to be considered below) according to an a prior plan (7:48, 49). If, on the other hand, Christianity is understood as an historical faith in a Messiah, then enumeration of its sects would be a "Herculean and thankless labor," inasmuch as Christianity would then merely be one national messianic faith as opposed to others—e.g., Judaism in the final period of its "complete domination over the Jewish people." Christianity, so construed, would only be the (statutory) faith of a particular people, and could not claim to be the final, universally valid revelation (7:49). A rational enumeration of all possible sects that can serve as vehicles of the true religion (and concomitant avoidance of a work as futile as it is heroic) goes hand in hand with an understanding of Christianity as other than a messianic faith linked to the history of a particular people.

Sects in general, according to Kant, are "admixtures" of religion and

paganism, the latter of which he defines as a “passing off” of the “external” as religiously essential.²³ Empirically speaking, sects are as innumerable as settled nations (themselves a combination of rational and merely statutory authority), with which sects are historically connected (see 7:50 n.). The relation of religion and paganism is thus one of (more or less) affined degree rather than of simple opposition. Paganism in a sect can, however, come to loom so large that the “entire religion” is transformed (literally, “goes over”) into something wholly pagan, which “passes off usages for laws,” and in which nothing (morally) rational remains (7:50–1).

There arises in sects (not completely pagan) an effectual tension, then, between the “influence” upon the soul of “pure, religious faith”—influence which, since it is “bound up with consciousness of freedom,” cannot be lost—and the “violence” [*Gewalt*] exercised on conscience by what is merely church faith (7:51). Accordingly, a plurality of sects is morally benign only where there is universal agreement about the essential maxims of religion, and disagreement only as to the “greater or lesser appropriateness [*Schicklichkeit*] or inappropriateness of the vehicle” (7:52).²⁴ Catholics and Protestants could, in this way, benignly coexist as “brothers in faith” (leaving it to time, and government forbearance in allowing freedom of belief, to adjust the “formalities”). Jews, on the other hand, can join that brotherhood (or “invisible church”) only by publicly becoming Christians—Christians, that is, in the peculiarly Kantian sense. We may gather, then, that Judaism, as Kant sees it, is (now) “wholly pagan”:

Without dreaming of a general Jewish conversion (to Christianity as a messianic faith), this (invisible church) is possible even in regard to Jews, if among them, as is now happening, purified religious concepts awaken, and the clothing [*Kleid*] of the old cult that no longer serves for anything, and, what is more, represses [*verdrängen*] all true religions disposition, is thrown off. Since they have now for so long had a garment without a man [*Mann*] (church without religion), and since, moreover, a man without a garment (religion without a church) is not well preserved, they need certain formalities of a church that is most suitable to lead them in their present place/situation: so we can consider [*halten*] the proposal of Ben David, a very good mind of that nation, to publicly accept the religion of Jesus (presumably with its vehicle, the Gospel), not only very fortunate but also the only proposal whose carrying out would soon make this people no-

ticeable [*bemerklich machen*] as one that is learned, well-civilized and ready for all civil rights, whose faith could also be sanctioned by the government, without their having to amalgamate with others in things of faith; [in keeping with this proposal] they would surely have to be left free to interpret the Bible [Torah and Gospels] so as to distinguish the way that Jesus spoke as a Jew to Jews from the way he spoke as a moral teacher of men in general.—The euthanasia of Judaism is pure moral religion with the leaving behind of all ancient regulatory teachings, some of which had to remain held back [*zurück behalten bleiben müssen*] in Christianity (as messianic faith): this division of sects must, yet, also finally disappear, and so the conclusion of what is called the great religious drama of religious change on earth. [The restoration of all things] leads hereby, at least in spirit, to there being only one shepherd and one herd (7:52–3).²⁵

This extraordinary passage, so retrospectively ominous,²⁶ sheds remarkable light, not only on Kant's own complex attitude toward Judaism and Jews,²⁷ but also on the nature, as he sees it, of religion as such. For present purposes, two features of his discussion are especially important: the highlighted connection between sectarian division, governmental sanction, and religious progress (the "great drama"); and the peculiar status of Judaism and the Jews.

On the first point: the statutory sanction (not just toleration) of religion is here not just an evil to be put up with—i.e., a temporary concession to historical conditions that would make opposition to established religion dangerous or politically destructive—but a positive good, the means by which enlightenment as such (in the terms laid out earlier in the essay) can be stably and lawfully advanced. Through its sanction of religious sects (for its own nonreligious ends), government can regulate, through the faculties it authorizes, those religious practitioners who deal with the people in a direct way, and who can therefore do most to help or hinder instruction at the most popular level. That Kant's most immediate political foes, in his own struggles with government repression, were often from precisely that group (as earlier defined) underscores the importance of this consideration to Kant's overall idea of a workable solution to the problem of education. By this mechanism, Kant is able to reconcile, as he believes, religious freedom (from such "coercive" measures as would, by playing on "fear and hope," destroy the "true religious disposition" of the people) with the steady pressure necessary to regularize (not to say, routinize) moral and religious progress (*Anthropology*, 7:327–8).²⁸ The

government's nonmoral interest in effective rule (i.e., in public order) is here harnessed to a moral purpose, through the intermediation of the faculties, whose interests (truth and ruling) are at once united and divided. The university can thus be thought of as a kind of coupling by which force tending one way is redirected in another—a spiritual mechanism, powered by “conflicting” human interests, that “providentially” addresses the problem of educating the human species, a problem that is otherwise, as Kant elsewhere states, “insoluble.”

The second point—Kant's attitude toward the Jews—is best considered in light of his “a priori” treatment of how Christian sects set about making men morally better (so that Christianity is “really encountered in their hearts”) (7:53). Whereas all Christians agree as to the end, there can be disagreement over means, from which arises an a priori “inward” division of religion (despite its essential unity) into two *religious* (not just historical, or nationally distinctive) sects. (Here, at last, is that a priori guidance of which Kant earlier spoke.)

The task of Christian (viz., moral) education—to work towards a (moral) disposition—must be supersensible, rather than empirical. What is more, because conscience (inasmuch as it “imposes” itself on us) is witness to the fact that we are not (now) good, this “working” must (or must be thought of as) bringing about a complete revolution, or transformation, as expressed by the question: “How can we be not just better men, but other men?”

For those who take everything *supersensible* for *supernatural*, that question becomes: “How is rebirth (whereby one becomes another, new man) possible through the immediate divine influence of God?” (7:54). This version of the question—one historically associated with Spener, who founded both the Pietist movement and (indirectly) the Moravian Brotherhood²⁹—is “rationally grounded,” as Kant insists, owing to its fundamental opposition to orthodoxy, i.e., the view that one can please God by merely carrying out prescribed observances. For those who pose the question in this manner, and who thus consider it a “small thing to ascribe natural effects to supernatural causes,” the answer is necessarily mystical (cf. 7:89). That mysticism takes the form of calling man's *natural* predisposition to the good (a predisposition to which conscience also bears witness) “mere *flesh*” [*Fleisch*]. People make this error (i.e., identifying nature with flesh) because they lack a clear conception of the supersensible (which is distinct from flesh without necessarily ceasing to be natural), and therefore think that a true moral revolution is possible

only through a *supernatural* effect.³⁰

For such people, for whom man's *nature* is hopelessly corrupt, and therefore "man cannot hope to improve by his own powers," there are two ways of "feeling" such conversion understood as the result of supernatural influence: for the first group, man's heart is crushed with remorse; for the second, it dissolves in blessed community with God (7:55). According to the first hypothesis, good in human nature separates from evil through a supernatural "operation" felt as a penitential "breaking and crushing" of the heart. That is to say, conversion begins with a miracle and ends with rational (i.e., morally good) conduct. Yet "even in the highest flights of mystical imagination," moral reasoning asserts itself. Unwilling or unable to regard themselves as mere machines, believers take a kind of credit for the very penitence that they also attribute wholly to grace, and thus contradict in practice their professed despair concerning the power of unassisted human nature (7:56).

On the second view, man makes the first step to moral improvement on his own, through rational awareness of his guilt in relation to the moral law; carrying out his resolution to make the law his maxim, however, is considered to be a miracle, requiring "awareness of continual intercourse with a heavenly spirit," which can be maintained only by "continuous cultivation through prayer" (7:56).

From such opposed representations of the feeling by which divine influence communicates itself (one anguished, the other so "sanctified" as to brand one as a holy fool)³¹ may well arise, Kant notes, two actual sects, of which the German Pietism and the Moravian Brotherhood (movements in which his own parents participated)³² are respective examples (7:55, 57, 57 n.). In each case, conscience, which is essentially bound up with consciousness of freedom, is "repressed" by a kind of force or violence, itself associated with lack of hope concerning the power of human reason to bring about the goal it aims at (see 7:51).

That violent repression is facilitated, in turn, by a mistaken belief in the possibility of bearing witness through one's feelings to supernatural influence. Because "the idea of God lies merely in reason," the claim "to *feel* the immediate influence as God" as such an influence is "self-contradictory." (The chasm between reason and sense, in other words, cannot be breached by mere assertion.) Each side exemplifies, from opposite poles of sorrow and joy, the historic struggle of pure reason's "force" against the "drag" of that spiritual inertia that takes it to be "no small thing" to take what is supersensible for supernatural. Neither side can

answer Spener's question adequately because each (out of a kind of intellectual lassitude)³³ relaxes, rather than striving to resolve, the tension between idea (by which human reason sets itself a task) and feeling.

There is, however, a "biblical" principle that is a *solution* [*Auflösung*] to Spener's problem. "In deed," it not only falls upon the eye that such a principle is encountered in the Bible, it is also "convincingly certain," Kant says, that only by virtue of such a principle, "contained in the Christianity of the Bible," could the latter have acquired so extensive an influence and enduring an effect upon the world. This effect no other teaching of "revelation (as such)"—no faith in miracles nor "united penitential voices"—could ever have brought about "because they would not have been drawn from the soul of man as such" and thus "would always have remained foreign to him" (7:58).

Kant's "proof" of his solution is thus a moral proof, based on the nature of the human soul as revealed by conscience, by way of "sound common understanding":

There is, namely, something in us we can never cease to wonder at once we have seen it [*ins Auge gefaßt habe*], the same thing that lifts up *humanity* in the idea to a dignity that one would never have suspected in *man* as an object of experience. That we are subject to the moral laws and destined through reason . . . to sacrifice (to them) all conflicting pleasures of life one doesn't wonder at, for (this) . . . lies objectively in the natural order of things as an object of pure reason, . . . so that it would never occur [*einfällt*] to sound common understanding to ask (with a view to evading them) whence such laws come to us or even to doubt their truth.—But that we also have the *ability* to offer to morality, with our sensible nature, so great a sacrifice, that we *can* do what we easily and clearly grasp that we *ought* to do—this ascendancy [*Überlegenheit*] of the *supersensible* man in us over the *sensible*, an ascendancy against which (if it comes to a conflict) the latter is *nothing*, though in its own eyes [*seinen eigenen Augen*] it is *everything*—this morality, from a disposition inseparable from our humanity, is an object of the highest *wonder*, which only rises higher the longer one regards [*ansehen*] this true (not fabricated) ideal (7:58–9).

The object of the constant and ever mounting wonder, once it is "seized upon in our eyes," is the ability or faculty of one's soul to sacrifice as nothing, "with one's sensible nature," that which to it (to me?) is every-

thing; i.e., to entirely negate, with my sensible nature, the value of what my sensible nature (I?) most esteem. That this negation is possible the moral imperative assures us; how it is possible is a question that is both dizzyingly unanswerable,³⁴ and ever increasingly uplifting once the vision [*Ansehen*] of that capacity is seized on. Wonder is the morally sublime, or something like a rational affect, however difficult it may be to bring the two together.³⁵

The “inconceivability” [*Unbegreiflichkeit*] of the “supersensible in us” which is yet “practical,” must and can, as practical, be “seized on” [*gefasst*].³⁶ Nevertheless, this inconceivability “excuses” those seduced into taking [*halten*] the supersensible *in* us for the supernatural *beyond* us. They are, however, “very wrong” in this, because were the source of the effect external, “the effect of this faculty would not be our own deed, and could not be imputed to us” (7:59).

The problem of the “new man” thus reduces to the question: How can we be the source of the effect by which we are made better without already being what we are thereby to become? The “genuine solution” consists in:

The putting into use of the idea of this faculty that cohabits us in an inconceivable manner, and its inculcation [*Anseherzlegung*; literally, “laying into the heart”]³⁷ beginning in earliest youth and continuing on in public discourse. And even the Bible seems to have nothing else in view [*vor Augen*], that is to say, it seems to point, not to supernatural experiences and fanatical feelings which should effect [*bewirken*] this revolution in reason’s place; but rather, to the spirit of Christ, which he demonstrated in teachings and examples that we might make it our own, or more exactly, since it already lies within us with the original moral disposition, that we only procure it room. And so between soulless orthodoxy and reason-killing mysticism there is the faith teaching of the Bible, which can, by means of reason, be developed out of ourselves, which, working [*hinwirken*] with divine force upon all human hearts toward fundamental betterment, and uniting them in a universal (although invisible) church, is the true religious teaching, grounded in the critique of practical reason (7:59).³⁸

Given the excusable susceptibility, on the part of ordinary decent folk, to “seductive” shortcuts—a failure that left to itself would bring about the “death of reason”—the chief vehicle of popular moral education is, then, the Bible (albeit, properly, or “scientifically,” interpreted). And govern-

ment, for its part, owing to its own interest in “tractable and morally good (i.e., externally faithful) subjects,” will refuse to sanction “merely natural” religion (or church faith without the Bible); instead, it will use its influence on the theological faculty to bind “public teaches of the people” to an “orthodoxy” (i.e., faith in the Bible as interpreted by teachers university trained and statutorily sanctioned) that (for once) isn’t “soulless” (7:60). In thus turning against itself the main obstacle to popular enlightenment (an obstacle manifested in the people’s willingness to be “led,” and which essentially consists in a lack of spiritual stamina), Kant’s idea of the university negotiates (heroically?) the narrow path between a religious tolerance that places no breaks on popular fanaticism (and thus brings about the death of reason) and a soul-destroying religious establishment.

VIII. Settlement of the Conflict

The conflict of the faculties, accordingly, looks forward to a peace (one shepherd and one herd) without arriving at it “in the foreseeable future.” Kant’s solution to the crisis of reason avoids “heroism” (in the medical sense) by prolonging the remedy indefinitely (see 7:64).

There is, however, one additional hurdle: the Bible’s claim to authority on the basis of a *Mosaic-messianic* faith, which reads the Bible as God’s literal covenant with Abraham and Moses.³⁹ Mosaic-messianism (outside of its original Jewish context) leaves the difficult matter of interpretation not to sharp-witted experts, where it belongs, but rather to the “natural reason” of the laity, who are tempted by irrational credulity (7:62, 61).⁴⁰ That this tendency flourishes especially in the religious sects just qualifiedly praised, and under the banner of “freedom of belief,” makes the danger all the more pressingly acute.

Kant’s strategy against such faith is simple: to trace its presence within historical Christianity to a Judaism that has outlived its spiritual usefulness—the “wearying remnant” of which reason (and Kant himself) earlier complained.⁴¹ The interpretive subtlety of that nation (see 7:66 n.)—whose art might still be enlisted in the moral cause were they to publicly accept “the religion of Jesus”—instead involves them in prophetic complications⁴² at once extravagant and morally empty (i.e., clothing without the man) (7:63).⁴³

Both acclaimed for its sublimity⁴⁴ and condemned to “euthanasia,”

Judaism assumes a special—indeed, crucial—place of favor and disfavor in Kant’s great drama of religious change on earth. The very sublimity of the Jewish law against idol worship (or sensible representation of the idea) aimed, it seems, too high; unsupported by the protective mediation (or what he elsewhere calls “scaffolding”)⁴⁵ of representation, the idea of religion, for them, simply collapsed, leaving behind a “wearying remnant” about which reason may rightly complain. Judaism, one could go so far as to say, is for Kant the very embodiment of the human problem of despair, if not the human problem simply. Both figuratively and literally, Kantian religion is the circumcision (or “cutting off”) from humanity of Judaism as Judaism.

Unlike the Jew’s messianic covenant with God (“marked” by the sacrifice or cutting off of Isaac), Christianity, as Kant understands it, abjures direct communication between God and man. Here, then, is one reason for Kant’s pronounced abhorrence, in religion as elsewhere, of communicative merger. The messianic covenant—i.e., the promise of deliverance by a superior and external voice on which one (therefore) cannot count—is the example, *par excellence*, of that “contagion” of sound common understanding of which Kant elsewhere warns.⁴⁶

If the authenticity of a divine statutory code, perfectly harmonizing in its end with morally practical reason, could be *documented*, then it would be, Kant grants, “the most forceful tool” for human guidance. But the difficulty of knowing whether God is actually speaking to one, given the impossibility of our apprehending the infinite by our senses, means that the only way that such a voice can be directly authenticated is through a negative comparison with what is morally certain: so that any being, however “majestic and surpassing of nature,” if it commands what, contrary to the moral law, shall be held an illusion (7:63).⁴⁷

The positive certification [*Beglaubigung*] of the Bible is neither the immediate voice of God, nor the allegedly divine learning of its author (who, being human, could err), but the “most forceful influence” it exercises upon the people—an influence which (since their preachers are as scientifically “incompetent” as they are), can only stem from the pure source of universal rational religion (7:63). But that this source dwells in “every common man” is not something that every man acknowledges. Hence the need, for the foreseeable future, of the Bible as a statutory form, which, so far as its “spirit” or moral content—as distinguished from its “letter”—is concerned, is “self-accrediting” (7:64).

Only by thus understanding the Bible’s “work” as “providential” in

the specific Kantian sense, i.e., as the “effect of nature” in conjunction with the “result” of a “progressing moral culture,” can one avoid the twin threats of religious skepticism (which reduces the Bible wholly to nature) and superstition (which raises the Bible wholly above it). The problem of popular education is soluble, in other words, in so far as history is grasped, not according to the (superstitiously worshipped) letter of the Bible, but as what Kant calls the “great religious drama of religious change on earth” (7:64; cf. 7:52–3). Of such providential (not messianic) history, he presents the following example:

The disciples of the Mosaic-messianic faith saw their hope in God's covenant with Abraham completely sink with the death of Jesus (we had hoped that he would deliver [*erlösen*] Israel); for salvation was in their Bible promised only to the children of Abraham. Now it happened that when the disciples were gathered at Pentecost, one of them fell upon the happy notion [*Einfall*] (cf. 7:17), in keeping with the subtle Jewish art of exegesis [*Auslegungskunst*], that pagans (Greeks and Romans) too could be considered as accepted in this covenant, if they believed in the sacrifice of his only son that Abraham was willing to offer God (as the sensible image of the sacrifice of the world-savior [*des Weltheilendes*]); for then they would be children of Abraham in faith (first with, but afterwards without circumcision [*Beschneidung*]). It is no miracle [*Wunder*] that this discovery, which, in a great gathering of people, opened up so immeasurable a prospect, was accepted with the greatest jubilation and as if it were the immediate effect of the Holy Spirit, and was held to be a miracle [*Wunder*] and came into biblical (Apostolic) history as such; by this, however, it belongs in no way to religion to believe [in these events] as a fact [*Factum*] and to press it upon [*aufdringen*]⁴⁸ natural human reason. Obedience necessitated through fear in esteem [*Ansehung*] for such church faith as required for salvation is thus superstition (7:65–6 n.).

In Kant's “great drama of religious change on earth” (which substitutes, explicitly, for the “book containing the entire history of Jewish covenant”) the letter of Christianity (i.e., the New Testament), rather than Judaism, is the remnant [*Vermächtung*] worth holding onto [*aufbehalten*]: like “old skin,” that remnant may be illegible in places, and yet serves humanity as a guiding band [*Band*],⁴⁹ “just as if it were a divine revelation” (7:65). Against the Mosaic-messianic faith in God's promise as literally written—a belief that stands behind the traditionally Christian view that “his-

torical faith is necessary to salvation”—Kant proposes an “evangelical-messianic faith” that is consistent with and preparatory to the moral transfiguration of the Gospel (7:66).⁵⁰

Historical Christianity, with its “dream” of Jewish conversion, is, rightly understood, the *Vorbild* of the true religion. Like the cutting off of Isaac (and/or his flesh) by/from his father, which marked the Jewish covenant with God,⁵¹ religion must eventually cut itself off from its protective clothing; the “scaffolding” must one day “fall away.”⁵² But the cut is not an easy or a clean one, inasmuch as religion needs such vehicles “for the foreseeable (but not perpetual?) future.” Given that need, moral progress will continue to depend upon arbitrary statutes and, hence, continue to court misadventure.⁵³

IX. Conclusion

In the *Conflict of the Faculties*, Kant works out in its clearest and most explicit form his answer to the question: How is (moral) education possible “wholesale,” i.e., with a view to the moral improvement of the people generally? His solution looks both backward, to traditional notions of Christian rebirth, and forward, to the modern organization of the factory. It also takes, not only as given but also as providential, a certain historical development of Christianity that finds its clearest expression in German religious sectarianism of the seventeenth and eighteenth centuries. His solution is thus, in an important sense, a national solution, as befits a conception of the university as, essentially, a state institution.⁵⁴ To be sure, sectarian warfare, so tragically intertwined with the German past toward which *Perpetual Peace* ironically gestures, is part of what Germany and Europe must put behind them, if theirs is ever to be an enlightened age. And yet, as Kant here labors to make clear, German sectarian division, rightly deployed and organized through the conflict of the faculties, provides a singular occasion (or *Veranlassung*) to effect the heart *en masse*.

The key to that deployment is the peculiar indirection with which the university manages popular education, without quite leading it. Readers familiar with Kant’s famous paeans to equal dignity may be surprised at his emphatic insistence on intellectual *inequality*, not only, and notoriously, between men and women, but also, and more crucially, between scholarly or scientific heads and ordinary human understanding. The pe-

culiar weaknesses of common understanding require (despite the latter's practical strengths) considerable and ongoing concessions for the indefinite future to historical modes of clothing the morally ideal in "sensible" garb. The university, as indirect alliance between philosophy (or reason) and state power, is more than a marriage of necessity or convenience, but the only way pragmatically conceivable of lifting on a sufficiently massive scale the drag upon enlightenment that is exerted by popular credulity and servility. Indirectly at first, then more directly, the university becomes the true visible church—dedicated both immediately, and through the disciplining of the professions, to the promotion of a national ethical culture. (Popular opinion is not so much led as liberated [or delivered] from excusable but very wrong seductions.) So understood, the university is the providential engine which, in the words of the first *Critique*, "guides" and "influences" humanity collectively, and without prejudice to human freedom.

Still, human reason, flummoxed by the impossibility of representing the fulfillment of the task it sets itself, is prone, always, to a deflating premonition of its own futility. Hence Kant's peculiar antagonism toward (contemporary) Judaism, which not only passively awaits an external deliverer, but also absolutely rejects the representation of holiness, in the person of Jesus, by which the true law makes itself more lovable and hence less burdensome. If there is a place for Judaism in an enlightened culture, then it is only as Jewish Christianity, i.e., a people to whom Jesus spoke individually as well as universally. If there is a place for Judaism in Germany, in other words, it is as a Christian sect rather than a German one. And yet (ancient) Judaism is in its way the very model of a people stirred (like the spectators mentioned in Part Two, 7:88) to ideal enthusiasm.

In thus "solving" the riddle with which, on Kant's account, Rousseau leaves us, Kant helped initiate a new kind of transformative pedagogy, which, for better and for worse, has deeply marked the nineteenth and twentieth centuries. He also gave liberalism new-found soul and liberal education a new lease on life; for it is also accurate to say that Kant's "idea" indirectly inspired, through Humboldt and his imitators, the modern university—devoted at once to research and to 'citizenship'—that America, especially, has embraced as model. But Kantian education also harbors darker currents, absent from Rousseau, who never sought such universal transformation. The contemporary "crisis" in university education—one that we have thus far been unable to resolve—may well reflect tensions inherent in its founding.

Notes

1. *Essays on Kant and Hume* (New Haven: Yale University Press, 1978), 204.
2. *Ibid.*, 210.
3. Kant's most famous statement to this effect is in the *Critique of Practical Reason* (5:55–6).
4. For a discussion of this problem, see *Conflict of the Faculties*, Part Two, 7:81 ff.
5. See, for example, Rousseau, *Emile*, tr. A. Bloom (New York: Basic Books, 1979), 49–50.
6. Compare A821–2/B849–50; Kant there defines belief or faith [*Glaube*] as a subjectively sufficient “holding to be true” that is at the same time taken as objectively insufficient, i.e., not necessarily held for true by every rational being, and hence not “commanded” by reason as such. All transcendental use of reason, to the extent that it holds something to be true, does so “as a matter of faith,” and from the standpoint of practice, be it that of skill or of morality. The term “faith” specifically refers here “to the guidance that the idea gives me and the subjective influence on the advancement of the actions of my reason, which holds me fast to them even though I am not in a position to give an account or justification of it from a speculative standpoint” (A827/B855). There is, Kant admits, a kind of faith—e.g., in the existence of God, insofar as we are led to it by our investigation of nature—which is not, strictly speaking, practical. There is in such belief, however, “something unstable,” given the susceptibility of such belief to “speculative difficulties” which “put one off,” at least for a time. With moral faith (e.g., in God's existence), on the other hand, the end to which faith is known to be a necessary means is itself fixed and absolutely necessary; here one is sure “that nothing could make this belief unstable, because thereby my ethical grounding principles would also be overturned, principles that I cannot renounce without being worthy of abhorrence in my own eyes” (A828/B856). In the *Critique of Practical Reason*, the “end” in question is not only complete conformity of my will to law but also realization of the highest good. Faith arises from reason's inability otherwise to satisfy its need to make the object of the law “intelligible” to itself (5:126).
7. The world used for “well-being” is in each case *Heil*, the same term Kant will later use to mean “salvation.”
8. The explicit task of Kant's *Religion within the Limits of Reason Alone* was to “mark” the relation between religion and the good and evil dispositions of human nature, and in so doing, “define” the concept of religion. Kant was there so bold as to suggest his essay as matter for a “final course” for students of the theological faculty, which possessed, as Kant acknowledged, the right to censor philosophic examinations of religion. Kant's attempt at a direct intrusion into the faculty of theology came to immediate grief, to be replaced, in the *Conflict of the*

Faculties, by an external appeal hardly more likely to appease (cf. his stated intention "to control [the faculty of theology] in order to be useful to it"). The larger question concerning the role of representation (or "marking") as an analog of that marking (by circumcision) that accompanied God's original promise to Abraham, will be taken up below. (See also Genesis, 17:11; Romans, 2: 25–9.)

9. Education [*Bildung*], on this account, is not imitative [*nachbildende*] "self-formation," but the rationally generated (rather than abortive) image of philosophy itself.

10. Cf. Kant's earlier definition of philosophy as the *Urbild* for the estimation of all attempts to philosophize.

11. See A743/B771: "Even poisons have their use. They serve to counteract other poisons generated in our bodily humors, and must have a place in every complete *pharmacopoeia*."

12. I.e., the moral law and, thanks to Kant's own teaching, the limits of theoretical reason.

13. See also A407/B434: "[Reason's natural antithetic] certainly guards reason . . . from the slumber of an imaginary conviction, but at the same time tempts it either to abandon itself to skeptical hopelessness, or to assume a dogmatic obstinacy, setting its head stiffly on certain assertions without allowing the grounds of the opposing party a hearing and to be done justice. Both are the death of a healthy philosophy, though the former might also be called the *euthanasia* of pure reason." Kant's notion of "critique" links judgment, in a juridical sense, with the related medical notion of a "crisis" as the moment of "judgment" in which the body either prevails over or succumbs to its own unbalanced humors. Hence the aptness of "counter-poison" as an ingredient in Kant's own critical arsenal. On the "self-preservation" of reason, see also *Critique of Judgment*, 5:261. In contrast with the more optimistic account of human development proposed by Herder, mental and physical puberty, for Kant, are fraught with danger in which the self-preservation of pure reason is itself at stake. For a different but related treatment of "engrafting" as a potentially life-threatening procedure, see Kant's discussion of vaccination (6:424).

14. Cf. Kant's later reference to "united penitential voices," in which, presumably, the individual (with his own particular moral history) is no longer heard as such. On the relation between philosophy and lowering one's tone, also *On a Newly Arisen Superior Tone in Philosophy* (1796). "Tutelage" is literally understood by Kant as an inability to speak for oneself.

15. Kant's example of such an objection, against which "weapons of war" are permitted, is the apparent proof by experience that the elevation or derangement of our spiritual forces are merely different modifications of our organs. The "force" of that proof can be weakened by the hypothesis that the body is nothing more than a "fundamental appearance." Though it exceeds the conditions of real knowledge, this speculative hypothesis is "polemically" permitted as a weapon against

the “hopelessness” that such an understanding of the relation between body and soul would otherwise entail. Without such an excuse, that hypothesis, on Kant’s own account, would be “visionary” if not deranged (see A769–82/B797–810). On Kant’s general concern with this issue see Shell, *The Embodiment of Reason* (Chicago: University of Chicago Press, 1996).

16. See *Idea for a Universal Natural History*, 8:23.

17. See the title to the Appendix, 7:36.

18. The specific competence of the theological faculty is the study and transmission of those aspects of historical Christianity that best promote the influence of true moral religion. The theological faculty’s very human tendency to inflate its own importance (in this case, by taking its own secondary status for primary) means, however, that it will always be in need of philosophic criticism.

19. Kant here quotes Cicero (*Ad familiam* 12:4), the original of which reads: *Nunc me reliquiae vestrae exercent* (“Now your vestral remains weary me”). Kant changes the singular to plural and omits *vestrae* (clothing). It seems plausible that he does so because, on his account, Judaism is nothing but the garment.

20. Viewing Christ as the human incarnation of God, who by his existence redeems us (without any effort on our part), warrants the ravings of one “Postellus,” who thought that he had found the special redeemer of women in one young and pious maiden of Venice. If the idea of humanity is taken to be a “real human being,” we are confronted with the difficulty that human beings come in two sexes, whose “weaknesses and transgressions” are “specifically different” from each other (7:39 n.).

21. Such doctrines, in other words, must not be finessed (literally, “gone over with dry feet” [*mit Trockenem Fuß übergangen werden*]), and yet should be surpassed (as one may urge) without delay (7:42). “Dry feet” may be an allusion to the Biblical crossing of the Jordan “dry-footed.”

22. See A467/B495: in addition to the “practical interest” that every well-disposed human being takes in the dogmatic determination of the cosmological ideas, an interest that is in some sense justified, Kant lists a “speculative interest” on reason’s part in gaining “unconditioned footing and support” in an unconditioned being; and, not least, a “popular advantage,” deriving from the “comfort” it affords the common understanding, which is disinclined to undertake that “restless ascent from the conditioned to the condition” in which “there can be no satisfaction.” Given that reason’s speculative interest in empiricism far outweighs its speculative interest in dogmatism, the endemic appeal of the latter finds its principal nonmoral source in intellectual inertia (or lack of “perplexity”) on the part of ordinary understanding. Kant carefully avoids identifying this “popular advantage” with a true “interest of reason.”

23. I.e., *das Äußerliche* (or, alternatively, “that outside the essence” [*das Außerwesentliche*]).

24. Otherwise, the existence in a state of many sects is good only as a “sign

that government allows the people freedom of belief," i.e., that the people are not, in matters of faith, externally coerced.

25. Cf. *Herdenmensch* as Nietzsche's contemptuous word for man as a herd animal.

26. See Mark Lilla, "Kant's Theological-Political Revolution," *Review of Metaphysics* 52:1998, 397–434.

27. Kant counted a number of Jews among his closest students and admirers, including Lazarus Ben David, David Friedlander, and Marcus Herz, who served as Kant's second during his *Inaugural Dissertation* defense. Kant once told Herz that, of all Kant's students, he was the one "who penetrated [his thought . . . most] deeply." See Kant's letter to Herz of May 1, 1781 (10:266–7).

28. On the political and theological effect in Germany of this transformation of liberal thought, see Lilla, *op. cit.*

29. See Introduction to Philip Jacob Spener, *Pia Desideria*, ed. and tr. T.G. Tappert (Philadelphia: Fortress Press, 1964).

30. Establishment of the supersensible (not supernatural) as a plausible concept is a guiding theme of Kant's *Critique of Judgment*.

31. See Kant's suggestion that a whole people educated in one of these sects would reflect it in their "national physiognomy." "That such would emerge," he says, "is not to be doubted, for repeated impressions, especially contrary to nature, upon the mind, externalize themselves in gesture and tone of speech, and expressions finally become permanent facial features. *Sanctified*, or, as Herr Nicolai calls them, *divinely blessed*, faces would distinguish such a people from other civilized and developed peoples (not exactly to its advantage); for this is a sketch of piety in caricature." Evidently, Kant regarded collective religious experience (taking "religion" here in its more common sense), and the unusual and peculiar feelings it may evoke, as a potential contributor to differences in national character.

32. Although Kant's parents were technically "Pietists," his praise of their serenity also links them with the Moravian Brotherhood as here described. On the generally cheerful and moderate character of Pietism in Königsberg, see Michel Despland, *Kant on History and Religion* (Montreal: McGill Queens University Press, 1973), 101–2.

33. See A473/B501, which also ascribes this tendency of ordinary human understanding to a certain vanity, inasmuch as both the more and less intelligent are equally at a loss, once the limits of human reason are exceeded. For an earlier version of Kant's antagonism toward intellectual lassitude, see his early *Universal Natural History and Theory of the Heavens* (1755), Part Three, 1:349 ff.

34. As indicated by the many "eyes" or viewpoints contained in the above passage. Cf. Martin Luther, *Freedom*, Works, V.31, 357.

35. Cf. *Critique of Judgment*, 5:272; and *Critique of Practical Reason*, 5:161. See also *Anthropology*, 6:261, on the related notion of "astonishment" (or perplexity before the unexpected), an "affect" according to Kant that is excited wholly

by reason. It is precisely this perplexity that ordinary understanding (which “does not know what it means to conceive”) is inclined to miss.

36. Here, beyond “conceivability,” popular and scholarly intelligence are, at least potentially, on equal footing.

37. Cf. *Auslegung* meaning “interpretation.”

38. A fuller account of how this good disposition cohabits with an evil one is taken up more fully in Kant’s *Religion within the Limits of Reason Alone*, 6:11, whose explicit aim is to make “notable” the relation of religion to a human nature loaded with both good and evil *Anlagen*, through the representation of good and evil principles as “two self-subsistent, inflowingly-effectual causes influencing men.” On the “idea” in us that gives us “force,” but of which we are not ourselves the author, see 6:61 f.

39. That this threat comes largely from the Protestant side gives Kant another reason to endorse the ongoing coexistence of Protestantism and Catholicism in the sense earlier described.

40. The Jews, during their “ethical” period, would seem to have avoided this fate owing, in part, to the natural interpretive gifts of their nation, in part to the “theocratic” organization of their constitution.

41. The modern practice of attacking Christianity by attributing what one finds disagreeable in it to Judaism (a practice that Voltaire perfected, and of which Marx’s “On the Jewish Question” is a late and notorious example) seems to originate with Spinoza. See Leo Strauss, “Preface,” *Spinoza’s Critique of Religion* (New York: Schocken, 1965), 16–17.

42. As Kant makes clear (7:62–3 n.), this tendency is responsible for the “Pietist” interest in historical “signs”—a problem to which Part Two of the *Conflict of the Faculties* specifically addresses itself.

43. See also Kant’s *Reflexionen zur Religionsphilosophie*, 19:652. Among the few handwritten notations in Kant’s personal Bible, the editors of the *Akademie* edition note the following: “old clothing” [*alt Kleid*] (at Matthew 9:16) is underlined in black, over which is written *Judenthum* (19:652). (Luther’s text reads: “Niemand flickt ein altes Kleid mit einem Lappen von neuem Tuch; denn der Lappen reißt doch wieder vom Kleid, und der Riß wird ärger.” [No one puts a patch of new cloth on an old garment for the patch tears away from the garment and leaves a more vexing tear.])

44. See *Critique of Judgment*, 5:274: “Perhaps there is no more sublime place in the law book of the Jews than the commandment: “Thou shalt make no image [*Bildniß*], or any likeness either of what is in the heavens or on the earth nor under the earth,” etc. This commandment alone can explain the enthusiasm that the Jewish people during their ethical epoch felt for their religion, when it compared itself to other people.

45. See Kant’s letters to Johann Caspar Lavater of April 28, 1775, and the following one (10:175–9).

46. By the same token, lying (the “one rotten stain of our species”) is morally intolerable because, over time, it puts “moral judgment out of tune” by damaging one’s ability to credit one’s own inner voice, rendering “imputability entirely uncertain.” (See *Religion*, 6:38–9.)

47. Kant’s example of such contrariety is the “myth” of the sacrifice of Isaac (the only use of the word “myth” in the essay). Kant goes out of his way here to emphasize (in a manner that calls to mind both the traditional blood libel and traditional portrayals of the Christian Passion) both the brutality of Abraham’s act, which Kant refers to as “the butchering and burning of his only son,” and the pathetic innocence of Isaac (who unknowingly carried the burden of his own wood); Abraham, Kant adds, “should have answered this supposedly divine voice, even if it should ring down from (visible heaven)”: “that I should not kill my good son is altogether certain; but that thou who appearest to me to be God, of that I am not certain and also can never become certain.” The story of the sacrifice of Isaac is, of course, the lightning rod of Kierkegaard’s later criticism, in the name of “authentic” Christianity, of a moral “legalism” he associates (following Kant and Hegel) with philosophy as such. See especially his *Fear and Trembling*.

48. Cf. the *verdringen* or pressing down by the Judaic remnant of all true religion (7:52–3).

49. I.e., *Leitband*, a verbal compromise, it seems, between two other terms of which Kant makes frequent use: *Leitfaden* for the guiding thread that reason spontaneously projects, and *Gängelband* for the leading strings that hobble the unenlightened. For Kant’s most famous treatment of the latter, see *What is Enlightenment?* (8:33–42). The related *Band* is a word for “volume” or “book.”

50. Biblical faith [*biblische Glaube*] (as distinguished from Bible faith [*Bibelglaube*] of the kind Kant favors) is a “messianic historical faith, that has at its basis a book of God’s covenant with Abraham, and consists in a Mosaic-messianic and an Evangelical-messianic church faith, and gives such a complete account of the origin and fate of God’s people . . . (from the beginning to the Apocalypse) . . . that one might expect a divine author”—but for the presence of questionable Cabalistic numbers . . . which must somewhat weaken” one’s faith in its authenticity (7:62). On the moral meaning for Kant of “Gospel” (literally, “good news” [*frühe Botschaft*]), the following late note (1799) may be relevant: “it is impossible that a man without religion be joyful in life”; *Reflexionen zur Religionsphilosophie*, 19:649.

51. See Genesis 17:11.

52. See Kant’s letter to Lavater of April 28, 1775 (10:176; cf. 179). The distinction between religion’s dispensable “scaffolding” and the indispensable “fortification” is arguably one that Kant never made entirely clear. See also in this regard the morally ambiguous “parerga” in *Religion*, which touch upon without being part of religion within the limits of mere reason. “Parerga” (or *Nebenwerken*: literally, “works beside the main work”) were classically associated with the sec-

ondary labors of Hercules. Kant uses the term elsewhere to designate such ornamental and protective works bordering the main work as frames, decorative columns, and drapery. See *Anthropology*, 7:243; and *Critique of Judgment*, 5:226.

53. This theme is taken up at length in Part Two, which explicitly confronts the despair-inducing view of history that Kant here, as elsewhere, associates with Moses Mendelssohn. Kant earlier questioned the “good will” (but not the “cleverness”) of his (former) friend, when Mendelssohn responded to Ben David’s proposal with “So long as God does not from Mount Sinai revoke our laws as festively as he (with thunder and lightning) gave them to us—that is, until the end of time—we are bound to them.” By this, Kant says, he apparently meant: “First get rid of the Judaism *in your own faith*, and then we will relinquish that in ours.” Kant here charges Mendelssohn with “cutting off” hope of the least lightening of the “pressing burden” of his fellow believers (7:52 n.). Intentionally or not, Kant seems himself to have misrepresented the position of Ben David, who consistently opposed Jewish conversion. On the latter point, see the *Encyclopedia Judaica* (Jerusalem: Keter) 4:474–5. Lazarus Ben David lectured and published on Kantian philosophy and was a forceful proponent of the latter’s views.

54. Although Kant was himself a member of a Prussian university and a subject of the Prussian state, his reliance on other German universities to allow him to publish elsewhere what he was not permitted to publish in Prussia suggests the emerging outlines of an institutionalized German cultural life. On Kant’s early interest in German national culture, see his essay *Observations on the Feeling of the Beautiful and the Sublime* (1764), 2:205–56.

The *Quid Facti* and *Quid Juris* in Kant's Critique of Taste

Henry E. Allison

In a highly influential essay, Dieter Henrich has argued that Kant's deductions were modeled on the so-called *Deduktionschriften*, which were still widely used in the eighteenth century to justify various legal claims within the Holy Roman Empire.¹ According to Henrich, these writings addressed the two questions to which Kant himself refers in the beginning of the Transcendental Deduction in the first *Critique*: the *quid facti* and the *quid juris*, and the resolution of each turned on a question of origin, albeit in different ways.² For example, if one were to claim to be the rightful owner of a piece of property, one would have both to document the acquisition (*quid facti*) and show that it stemmed from a legally valid contract (*quid juris*). Moreover, though he places special emphasis on the deduction of the second *Critique*, with its notorious appeal to the "fact of reason," Henrich also suggests that this legal model provides the essential basis for understanding all of Kant's deductions.³

My present concern is to explore this hypothesis in connection with a deduction that Henrich himself largely ignores, namely that of pure judgments of taste in the third *Critique*.⁴ As Kant emphasizes, a deduction is here required because a claim of taste involves a demand for universal agreement that is analogous to the demand built into a cognitive judgment. Thus, in asserting that a particular object (whether of nature or art) is beautiful, one is claiming to speak with a "universal voice," even though one's judgment is based on a feeling. Accordingly, the *quid juris* in the domain of taste concerns this presumed right. And since it is only the pure judgment of taste that lays claim to such a right, the *quid facti* must be whether a given judgment of taste is pure, which is straightforwardly a matter of origin.

These questions are addressed in the Deduction and the Analytic of the Beautiful respectively. At least in broad outline, then, the *Deduktionsschrift* model corresponds to the division of labor in Kant's treatment of the normativity of taste. Beyond that, however, I shall also try to show that

this model underscores a feature of Kant's account that is often either glossed over or misconstrued; viz., the undeterminability of the *quid facti*. As we shall see, the problem is that, though the Analytic clearly specifies the conditions under which a judgment of taste may lay claim to purity, it turns out that we can never be in a position to determine with certainty that these conditions have been met in a given instance. But rather than viewing this as a weakness in Kant's argument (much less a fatal flaw), I shall claim that it is an inevitable consequence of the premise (shared by Kant and virtually all of his contemporaries) that judgments of beauty are matters of taste.

I

For Kant, the above means that judgments of beauty are aesthetic, that is, based on feeling rather than concepts. Starting with this assumption, he explores in the Analytic of the Beautiful the conditions that must be met by such judgments, if they are to lay claim to purity. Since an adequate account of these conditions would require a monograph, I shall here confine myself to a rapid overview of the basic points insofar as they bear on the *quid facti*.

The first of these conditions is disinterestedness. That interest independence is a *sine qua non* for a pure judgment of taste follows from the twin assumptions that such a judgment is aesthetic and that it claims normativity. Since all interest for Kant concerns either the agreeable or the morally good, if a judgment of taste were based on an interest, it would be reducible either to a judgment of agreeableness or of moral goodness. In the former case it could not claim any normativity (since what is agreeable is purely a private matter), while in the latter case it would no longer count as an aesthetic judgment. Consequently, disinterestedness matters because it is a condition of the autonomy of taste, which, in turn, is a condition of its claim to normativity.

But rather than arguing directly for this, Kant contends at the beginning of the Second Moment for the converse thesis that the universality of a judgment of taste follows from its disinterestedness. Although this argument has been rejected as a *non-sequitur*, I believe that its main thrust is defensible, at least within a Kantian framework.⁵ Once again, the essential point is that all interested liking is either for the agreeable or the morally good, with the former based on inclination and the latter on pure practical reason. And since a disinterested liking would, *ex hypothesi*, not be based on inclination, there remains nothing merely private on which it

could be based. To suggest that it might be based on something purely idiosyncratic, say a fondness for yellow, is beside the point, since within the Kantian framework, this would have to be regarded as an idiosyncratic inclination. Given this, the remainder of the Second Moment is devoted to the exploration of the nature of this universality (showing that it is merely subjective in the sense of applying to the sphere of judging subjects), and to the location of the ground of a putative, universally valid liking in the free harmony of the faculties. The latter is crucial to both the *quid facti* and *quid juris*, and is a topic to which we shall return.

The Third Moment makes another essential contribution to the determination of the conditions of the pure judgment of taste, this time from the side of the object. It does so by arguing, albeit circuitously, that only the form of an object or its representation can occasion the free harmony of the faculties. And from this it is inferred that only a judgment based on this form alone counts as pure. This, too, is a much criticized move, since it seems to commit Kant to a highly restrictive aesthetic formalism, according to which only spatio-temporal configurations are relevant to judgments of taste.⁶ Once again, however, I think that the move is at least partly defensible. Here the key point is that the harmony in question is a harmony *in reflection*, which means that only an ordering of the sensory materials (sensations) given to the imagination to apprehend, and not isolated sensations themselves, can occasion such a harmony. Although this entails that the pure judgment of taste must attend only to formal features of an object or its representation, it does not require us (as Kant apparently thought) to restrict these to spatio-temporal configurations. Moreover, so construed, the Second and Third Moments complement one another, since the former addresses the issue of the origin of the disinterested liking affirmed in the pure judgment of taste from the side of the subject, and the latter from the side of the object.

The Fourth Moment is concerned with the necessity claim built into judgments of taste. As in the previous analysis of universality, Kant notes that this necessity must be subjective, since it concerns a feeling. Kant himself characterizes it as "*exemplary*," which is glossed as "a necessity of the assent of everyone to a judgment that is regarded as an example of a universal rule that we are unable to state" (*Judgment*, 5:237; 85).⁷ In other words, when I judge an object to be beautiful, I am demanding the agreement of everyone, because I take my judgment to be normative. And I take it to be normative because I assume that it instantiates a universal rule, which, since the judgment is aesthetic rather than cognitive (based

on a feeling rather than a concept), cannot be stated. Moreover, Kant argues, the very intelligibility of such a demand, as well as of the assumption that one's judgment instantiates an unstatable (because non-conceptual) rule, presupposes the idea of a common sense, which is initially defined as "the effect arising from the free play of our cognitive faculties" (*Judgment*, 5:238; 87).

In addition to the inherent difficulties stemming from the ambiguity of this idea of a common sense,⁸ Kant's appeal to it at this point creates special problems for an interpretation such as the one advanced here. For one thing, it does not seem to provide an additional constraint on the pure judgment of taste beyond those already provided in the three preceding Moments. Thus, whereas we might examine a putatively pure judgment of taste to see if it meets the conditions of disinterestedness, of being based on the harmony of the faculties, and of being directed merely to the purposiveness of form of the object, it is not clear what else one might be looking for if one asks whether it presupposes the idea of a common sense. In short, this idea points to a general condition of the possibility of a pure judgment of taste, but apparently not one that might be used to distinguish the pure from the nonpure variety.

Second, and potentially even more damaging, in the Fourth Moment, Kant goes beyond affirming that the idea of a common sense must be presupposed as a condition of the possibility of pure judgments of taste; he also provides in section 21 an argument purporting to show that we have grounds (based on an analysis of the nature and conditions of cognition) for assuming that there actually is such a sense. Moreover, this argument has frequently been viewed in the literature as itself providing a deduction or the "first deduction" of taste.⁹

In response to this obvious line of objection, I shall make two interpretive suggestions, which I cannot hope to justify here. The first is that the function of the appeal to the idea of a common sense in the Fourth Moment is not to provide an additional constraint on the purity of a judgment of taste, but rather to supply a kind of focal or unifying point, which brings together all of the factors that have previously been analyzed in isolation. In other words, the thought of a judgment possessing the characteristics and conditions delineated in the first three Moments already brings with it the idea of a common sense. Moreover, this is just what Kant himself seems to suggest, when, at the very end of the Fourth Moment, after raising the perplexing questions about whether common sense is constitutive or regulative, and taste a natural or acquired faculty, he remarks that these is-

sues are not to be considered at this point because "for the present our task is only to analyze the faculty of taste into its elements and unite them finally in the idea of a common sense" (*Judgment*, 5:240; 90).

The second suggestion is far more controversial, namely that the argument in section 21 purporting to show that we have grounds, derived from an analysis of cognition, for assuming a common sense, should not be read as an attempt to establish a common sense in the sense stipulated at the end of section 20; viz., "the effect arising from the free play of our cognitive faculties." Part of the justification for this suggestion is that when read in this way (which is admittedly also the most natural way) the argument is almost universally judged a failure. Moreover, the failure is generally thought to lie in the move from the cognitive to the aesthetic senses of a common sense. For even assuming that Kant there succeeds in showing that a common sense must be presupposed as a condition of the universal communicability of cognition, and that the latter is required if skepticism is to be avoided, it is by no means clear that this requires anything like what Kant later refers to as a "*sensus communis aestheticus*" (*Judgment*, 5:295 n.; 162).¹⁰ Accordingly, the argument seems much more plausible if it is taken as merely indicating that we have grounds for assuming something like a cognitive common sense.

Why, then, would Kant bother to include such an argument in the Analytic of the Beautiful, where the explicit concern is with the conditions of taste rather than cognition? The answer, in my view, lies in the apparently paradoxical nature of the idea of a common sense [*Gemeinsinn*].¹¹ Much like the perplexing notion of an a priori intuition, the idea of a common sense combines apparently incompatible properties, namely of being a sense or feeling (which is essentially private), and of being directed to what is common or universally communicable. And given this, it does seem reasonable for Kant to have attempted to show that the idea of a such a sense is not an impossible fiction by indicating that something like it plays an essential role in cognition. This would not amount to a deduction of a principle of taste, but it would nonetheless be a useful prolegomenon to such a deduction.

II

Kant unambiguously links the Deduction with the *quid juris* by identifying its task as the establishment of "the *universal validity*, for the faculty

of judgment as such, of a *singular* judgment that expresses the subjective purposiveness of an empirical representation of the form of an object" (*Judgment*, 5:281; 143–4). The proposed method is to focus on what Kant terms the two "logical peculiarities" of a judgment of taste, which turn out to be its putative subjective universality and necessity. According to Kant:

If we resolve [*die Auflösung*] these logical peculiarities . . . we shall have done all that is necessary in order to deduce this strange faculty, provided that at the outset we abstract from all content of the judgment . . . and merely compare the aesthetic form with the form of objective judgments as prescribed by logic (*Judgment*, 5:281; 144).

Since Kant had already dealt with these "peculiarities" in the *Analytic*, the above might seem somewhat surprising, particularly given the present insistence on a clear division of labor between the *Analytic* and the *Deduction*. Indeed, sections 32 and 33, which are devoted to the consideration of these peculiarities, appear to provide nothing more than a restatement of the results of the *Analytic*, albeit in somewhat different terms. Moreover, precisely these considerations underlie the critique of Jens Kulenkampff, who argues that Kant's very method prevents him from providing anything like a genuine deduction in the realm of taste. This method, as he describes it, is that of "uncovering analysis" [*entdeckender Analyse*], by which he means essentially the analytic method. Thus, in sharp contrast to those who maintain that the *Analytic* already contains a "first deduction," Kulenkampff contends that Kant's purported deduction is best described as a "second *Analytic*," since it never gets beyond a regressive account of the conditions of the intelligibility of judgments of taste, such as is provided in the *Analytic*.¹²

Clearly, everything turns on how one construes the *Auflösung* of these peculiarities. If, as Kulenkampff assumes, it consists in nothing more than the determination of the conditions of the possibility of the judgments possessing them, then there would no more be a deduction here than in the *Groundwork*'s contention that the autonomy of the will is the supreme condition of the possibility of a categorical imperative, or the argument of the *Prolegomena* that a pure intuition of space is a necessary condition of geometry, conceived as a body of synthetic a priori knowledge. In that case one would not have gotten beyond the *quid facti*. If, on the other hand, the resolution consists in articulating and grounding a

principle capable of *justifying* (rather than merely explaining the possibility of) judgments with these peculiarities, then the *quid juris* will have been addressed. In other words, what is required is the authorization of a right to speak with a universal voice in matters of taste (on the condition that one's judgment is pure), and this can be provided only by some higher-order normative principle under which the first-order judgment of taste is subsumed.

Moreover, this is precisely how Kant characterizes his project in the Deduction. Thus, after examining these peculiarities in sections 32 and 33, Kant turns in section 34 to the question of a principle [*Princip*] of taste. As one would expect, he begins by denying the possibility of an objective principle on the grounds of its incompatibility with the aesthetic nature of the judgment of taste. Kant also suggests, however, that this does not preclude the possibility of a subjective principle. In fact, the claim is that insofar as the critique of taste has a transcendental dimension, its task is to “set forth [*entwickeln*] and justify [*rechtfertigen*] the subjective principle of taste as an a priori principle of the faculty of judgment” (*Judgment*, 5:286; 150). Accordingly, it seems reasonable to regard the deduction as consisting of two steps: the first sets forth or presents the subjective principle that licenses the demand implicit in pure judgments of taste, and the second attempts to ground this second-order principle. These are the tasks of sections 35 and 38 respectively, while sections 36 and 37, which I shall not discuss here, provide a transition from the first to the second step, and section 39, which I must likewise neglect, is a kind of appendix.¹³

III

Kant identifies the required subjective principle in the heading of section 35 as the “subjective principle of the faculty of judgment as such,” and at the end of the section he refers back to it as the “legitimizing basis” [*Rechtsgrund*] (*Judgment*, 5:287; 151). The intervening argument is compressed into a single dense paragraph and formulated entirely in the language of subsumption. Given Kant's characterization of judgment as the faculty of subsuming particulars under general concepts, this language might seem appropriate. The problem, however, is that this characterization was formulated with cognitive judgments in mind, and it is not immediately apparent, to say the least, how it is supposed to apply to judg-

ments of taste, which are not based on concepts. In short, we need an account of what is supposed to be subsumed under what in such judgments, such that the end product can make a rightful claim to normativity.

Kant's answer, as usual, is by elimination. Since concepts are precluded *ex hypothesi*, the only remaining alternative is that the subsumption be under the "subjective formal condition of judgment as such," which is identified with the faculty of judgment itself (*Judgment*, 5:287; 151). In other words, in a judgment of taste it is the *faculty of judgment* that plays the role assigned to a concept in a cognitive judgment; so the representation of the object is subsumed under this faculty, which therefore serves as the subjective principle of such judgments. But what does it mean to subsume a representation under a cognitive faculty in general and under judgment in particular? Unless a clear sense can be given to such "subsumption," Kant's solution must remain a purely verbal one, without any power to explicate the normativity of taste.

The essential point here is that the normativity of taste, like all normativity for Kant, must be rooted in the condition of the successful activity of the faculty in question. Thus, by subsumption under a faculty is to be understood subsumption under such a condition (or conditions), which, in the present case, means the condition of subsumption. Kant describes this sought-for condition as the "subjective formal condition of judgment as such." It is subjective in that it is a constraint on judgment's capacity to perform its characteristic subsumptive activity rather than on what may be encountered in experience.¹⁴ It is formal in that it is a constraint on the form of what is apprehended insofar as it allows for the possibility of subsumption (its subjective purposiveness for judgment). Moreover, since subsumption requires that what is given in intuition and apprehended through the imagination be brought under concepts of the understanding, this condition turns out to be nothing other than the harmonious interplay of the imagination and the understanding in their respective activities.

We initially encountered this harmony in connection with the discussion of the *quid facti* in the *Analytic of the Beautiful*. It was there noted that a free harmony is the source of the pleasure that is claimed to be universally communicable in the pure judgment of taste. As such, it is also the subjective condition of the purity of such a judgment; the "objective" condition being its correlate, the purposiveness of the form of the object or its representation. We learn in section 35 that such a harmony is normative for taste precisely because harmony in general is normative

for judgment as such. Moreover, this normativity is connected explicitly with the subsumptive activity of judgment. For Kant now claims that the judgment of taste consists in "the subsumption of the imagination itself . . . under the condition that the understanding requires in order to proceed in general from intuition to concepts" (*Judgment*, 5:287; 151). And since this condition is just what Kant means by the harmony of the faculties, it is this that functions as the norm for judgment, as well as that under which the intuitive representation of an object is subsumed in the judgment of taste.

Kant's own accounts of harmony in general and free harmony in particular tend to be notoriously cryptic and metaphorical. And this has proven to be the major stumbling-block to the understanding, not to mention the evaluation, of his theory of taste. At least with respect to the general notion of the harmony of the faculties, however, there is a particularly helpful, albeit metaphorical, discussion contained in a student's transcript of Kant's lectures on logic. According to this transcript, Kant states:

Imagination and understanding are two friends who cannot do without one another but cannot stand one another either, for one always harms the other. The more universal the understanding is in its rules, the more perfect it is, but if it wants to consider things *in concreto* then [it] absolutely cannot do without the imagination.¹⁵

As the passage clearly indicates, the "friendship" between imagination and understanding is not without a certain tension, which results from the fact that they pull in opposite directions: the understanding towards universality and the imagination towards specificity. Accordingly, though the understanding requires the imagination to exhibit intuitively what is thought in its concept (concepts without intuitions are empty), and the imagination presumably needs the understanding to give it direction (intuitions without concepts are blind), they nevertheless often work at cross purposes. This occurs either when the understanding in its endemic quest for universality produces a concept that is too general and indeterminate to be represented adequately by any particular instance, e.g., the concept of a thing, or when the particular imaginatively apprehended is too idiosyncratic or atypical to represent fully what is thought in the concept, e.g., the image of a three-legged dog.

This further suggests that the general notion of a harmony of the faculties for Kant may be taken in either a minimal or a maximal and ideal

sense. Harmony, minimally construed, occurs whenever there is any cognitive fit between concept and intuitive representation, that is, whenever the latter is subsumable under the former. And in that sense harmony is a necessary condition of cognition. By contrast, the same harmony maximally or ideally construed occurs when the fit between universal and particular is extremely close, that is, when the understanding's concept is not too indeterminate for the imagination and, conversely, the latter's image exhibits all of the essential features thought in the concept. In that case, the two faculties work together smoothly, like two well-meshed gears with little friction, and the subsumption accordingly proceeds without difficulty. Moreover, it is clearly harmony in this ideal sense that functions as a norm for judgment and therefore as the subjective principle of taste.

But harmony, so conceived, is an explicitly cognitive norm and it remains to be seen how it could serve as a principle for a noncognitive judgment of taste. In order to understand this, we must consider the concept of a free harmony, which cannot simply be equated with harmony as explicated above.¹⁶ Such a free harmony can occur only in "mere reflection," that is, an aesthetic contemplation in which the normal cognitive concerns with classifying objects under determinate concepts or explaining them by bringing them under covering laws are suspended.¹⁷ As before, this must concern the manner in which an object is presented in intuition by the imagination. But rather than presenting something that exemplifies a kind or indicates what it is to be of a certain kind, the imagination now produces an image that pleases in virtue of its being, as it were, itself a kind. In Kantian terms, this means that an intuitive representation simulates the exhibition of a concept, albeit no concept in particular (for the latter would necessarily involve comparison with other objects and an appeal to determinate concepts).

Kant describes the activity of the imagination in its free play in deeply suggestive fashion as "schematizing without a concept" (*Judgment*, 5:287; 151), by which we must understand the production of a form or organization in the sensory representation, such that it seems suitable to serve as the intuitive exhibition of a concept, though no particular concept is exhibited. In other words, the imagination produces what might be termed a "schema of a schema," that is, something like the form of the exhibitability of a concept in general (though, again, no concept in particular). And this is pleasing to judgment, since it requires that all concepts (empirical as well as pure) be exhibitable in intuition.

Although much more work is obviously required to fill in the details of this picture, particularly with respect to the free harmony underlying taste and the imagination's "schematizing without a concept," some points should already be clear, and these must suffice for present purposes. First, the familiar objection that if Kant's argument proves anything at all, it proves too much, since it entails that everything must be beautiful, is a red herring.¹⁸ For that objection assumes that by the harmony of the faculties is to be understood the minimal necessary condition of cognition, which accordingly must be met by any cognizable object. If the above account is correct, however, the harmony in question is an ideal rather than a minimal condition; and as such it is seldom, if ever, fully met in cognition. Moreover, even if it were to occur frequently in cognition, this would have no bearing on its occurrence in a free play of the faculties that does not aim at cognition.

This account of the harmony of the faculties also enables us to understand Kant's puzzling talk in section 21 about degrees of the "attunement" of the cognitive faculties, differences of "proportion" derived from differences in the objects cognized, and of an optimal attunement with a view to the "cognition of objects in general" (*Judgment*, 5:238–9; 88). These refer to differences in the degree to which the "two friends," imagination and understanding, manage to cooperate in pursuing their distinct cognitive goals. And the condition of optimal attunement that is maximally beneficial to both is simply the ideal state of harmony, where there is a virtually perfect fit between concept and intuition, universal and particular.

Finally, and most importantly with respect to the argument of section 35, we can see why the faculty of judgment itself provides the normative principle underlying claims of taste, and why this norm must take the form of a principle of subsumption. Judgment provides the norm for taste precisely because the pure judgment of taste is concerned in its own way (aesthetically) with the fit between universal and particular. And since this fit is entirely a matter of the subsumability of the latter under the former, this norm functions here as a principle of subsumption as well. As Kant points out, however, in the case of a pure judgment of taste it cannot be a principle of the subsumption of intuitions under concepts, "but, rather, the faculty of intuitions or exhibitions (i.e., the imagination) under the faculty of concepts (i.e., the understanding) insofar as the imagination *in its freedom* harmonizes with the understanding *in its lawfulness*" (*Judgment*, 5:287; 151). And, as already noted, Kant characterizes

this principle, which was arrived at through reflection on the logical peculiarities of judgments of taste, as the "legitimizing ground" of such judgments.

IV

Having thus formulated the subjective principle or legitimating ground of a pure judgment of taste, it still remains necessary to justify this principle, that is, to ground the normativity of the norm. This is the task of section 38, which is appropriately entitled: "Deduction of Judgments of Taste." For those who have struggled with the complexities of the two versions of the Transcendental Deduction in the first *Critique*, this one is remarkably simple and straightforward, a point on which Kant himself remarks in the comment attached to the argument (*Judgment*, 5:290; 156). It consists of three sentences, each of which constitutes a step in the argument. I shall first cite the steps and then comment on them in turn:

1) If it is granted that in a pure judgment of taste our liking for the object is connected with the mere judging of its form, then this liking is nothing but its subjective purposiveness for judgment, which we sense as connected in the mind with the representation of the object.

2) Now, since with regard to the formal rules of judging apart from all matter (whether sensation or concept), judgment can be directed only to the subjective conditions of the employment of judgment in general (which is restricted neither to a particular mode of sense nor a particular concept of the understanding), and hence to that subjective factor [*dasjenige Subjective*] that can be presupposed in all men (as is required for possible cognition in general), so it must be allowable to assume that the agreement of a representation with these conditions of judgment is valid for everyone a priori.

3) That is to say, the pleasure or subjective purposiveness of a representation for the relation of the cognitive faculties engaged in the judgment of a sensible object in general can with right be required of everyone.

Since the first of these sentences does nothing more than restate the results of the Analytic as they relate to the Deduction, while the third merely attempts to restate the main implication of the second, it is clear that virtually the entire weight of the argument falls on the second step. But

what does it claim? First, it argues by elimination that formal rules of judgment, lacking any "matter" (whether sensory or conceptual), can be directed merely to the subjective conditions of the employment of the faculty of judgment as such. In spite of the difference in language, this is basically a restatement of the thesis of section 35 that the principle of taste can be only "the subjective formal condition of judgment as such." Then, it simply asserts that we may presuppose these same conditions in everyone, on the grounds that they are conditions of cognition. And from this it is inferred that we are entitled to assume that a representation's harmony with this condition (or conditions) will be valid for everyone. Finally, in what has been characterized as step three, this entitlement claim is transferred from the harmony to the pleasure through which it is felt.

The crux of the argument appears to be a principle of the form: if x is subjectively purposive for me, then it must be subjectively purposive for everyone. This seems a reasonable enough claim, given the connection between subjective purposiveness and the conditions of judgment built into the very definition of such purposiveness, and the fact that the cost of denying the universality of these conditions would be a radical skepticism that would undermine any form of rational communication. Moreover, Kant himself indicates that this is basically what he had in mind both in a footnote attached to section 38, where he states that the initial thing that must be granted is that "in all people the subjective conditions of this faculty [judgment] are the same" (*Judgment*, 5:290 n.; 155), and in the following comment, where, by way of explaining why the deduction is "so easy," he tells us that it asserts only that "we are justified in presupposing universally . . . the same subjective conditions of the faculty of judgment that we find in ourselves" (*Judgment*, 5:290; 156).

V

This result establishes the principle that pure judgments of taste possess a rightful claim to normativity, but it does not suffice to justify any particular claim of taste. For, as Kant himself points out in the footnote, the latter also requires that "the judgment has taken into consideration merely this relation [of the cognitive faculties] (and hence the *formal condition* of the faculty of judgment) and is pure, i.e., mingled neither with concepts of the object nor with sensations as the judgment's determining grounds" (*Judgment*, 5:290 n.; 155). And, in connection with the second passage,

he adds that a claim of taste also asserts that "we have subsumed the given object correctly under these conditions" (*Judgment*, 5:290; 156).

In both of these passages, then, Kant clearly distinguishes between the *quid juris* and the *quid facti*, and suggests that they are jointly required to legitimize any given judgment of taste. In other words, one must be assured both that pure judgments of taste are normative and that one's judgment is, in fact, pure. But, as he also notes in the second passage, precisely because the judgment of taste is based on feeling rather than concepts, determining the correctness of the subsumption "involves unavoidable difficulties" and "may easily be illusory" (*Judgment*, 5:290–1; 156). In fact, though he does not say it in so many words, I think the clear implication of Kant's analysis is that we can never be certain in any instance that we have made the correct subsumption, that is, that one's judgment is based solely on the relation of the faculties in free play.¹⁹ We can, of course, take pains to abstract from charm, emotion, and the like, and presumably the need to do this is at least part of what Kant is hinting at in section 22, with his oblique reference to taste (identified with the *sensus communis*) as a regulative rather than a constitutive principle. But though we may be required to strive towards this ideal of aesthetic purity, we can never be certain that we have attained it, just as in the moral realm we can never be sure that we have acted from duty alone.²⁰

For present purposes, however, what is of particular interest is that Kant does not seem to have been at all disturbed by this result. On the contrary, he remarks in the footnote that, "even if a mistake be made on the latter point, this amounts to nothing but an incorrect application, in a particular case, of an authority given to us by a law, and in no way annuls the authority" (*Judgment*, 5:290 n.; 155). Moreover, he makes essentially the same point in the remark to section 38, insisting that problems concerning the correctness of subsumption, which also occur, though in a lesser degree, in cognitive judgments, do not affect the legitimacy of the "principle of judging validly for everyone from subjective sources" (*Judgment*, 5:291; 156). Clearly, then, Kant's position is that the Deduction establishes merely the *quid juris*, and that this stands regardless of the difficulties involved in the determination of the *quid facti*. Indeed, it stands even if (as I believe to be the case for Kant) we can never be certain in a particular instance that the conditions of a pure judgment of taste have been met. For we at least know that (and on what basis) such a thing as a pure judgment of taste is possible and (through the Deduction) that it has normative force.

Admittedly, the result of the Deduction on this reading is rather modest, since it leaves questions of the normativity of particular claims of taste undecidable. Moreover, in this respect it differs sharply from that of many interpreters, most notably Paul Guyer, who takes the task of a Kantian deduction of taste to be to provide something like a warrant to expect or demand a particular response from others, and then proceeds to criticize the argument for failing to provide it.²¹ I cannot discuss the merits of Guyer's reading here; but I would like to close with two brief comments regarding the significance of the Deduction on the alternative reading sketched in the present essay.

The first is that the conclusion here attributed to the Deduction, though relatively modest when compared to that of the Transcendental Deduction in the first *Critique*, is far from trivial. For if sound, it establishes something that is frequently denied, namely that there is normativity in the domain of taste, and that this normativity is based on a principle unique to judgment. Consequently, judgments of taste are reducible neither to judgments of agreeableness, which lack normativity altogether (the empiricism of taste), nor to judgments of perfection, whose normativity is based on a principle extrinsic to taste (the rationalism of taste).²² Moreover, since this is tantamount to establishing the autonomy of taste, such a result would seem to provide ample justification for a separate *Critique* of the faculty of taste.

The second and final point is that this outcome is also the best that could be expected, given the nature of taste. For if, as Kant assumes without arguing, judgments of taste are aesthetic, then any normativity pertaining to them must govern feeling, and this already precludes the possibility of a rule or decision procedure for determining whether any particular judgment conforms to this norm. Indeed, as I hope should be clear by now, the interest of Kant's theory of taste consists largely in the fact that it attempts to preserve a space for their normativity without denying the aesthetic nature of judgments of taste.

Notes

1. Dieter Henrich, "Kant's Notion of a Deduction and the Methodological Background of the First *Critique*," in *Kant's Transcendental Deductions: The Three "Critiques" and the "Opus Postumum"*, ed. E. Förster (Stanford: Stanford University Press, 1989), 32.

2. Ibid., 35.

3. Ibid., 30.

4. In this paper I use the term 'deduction' to refer both to the argument intended to resolve the *quid juris* and the sections of the *Critique* in which it occurs (sections 30–9). To disambiguate, I shall use the lower case when referring to the former and the upper case when referring to the latter.

5. For a critique of this argument, see Paul Guyer, *Kant and the Claims of Taste* (Cambridge: Harvard University Press, 1979), 168–70. The above may be seen as an attempt to meet his objections.

6. See *ibid.*, esp. 224–37.

7. The last number in parentheses gives the page reference to the English translation of the *Critique of Judgment* by Werner S. Pluhar (Indianapolis: Hackett, 1987). Although I make use of Pluhar's translation, I frequently modify it, sometimes significantly.

8. For differing accounts of this ambiguity, see Donald Crawford, *Kant's Aesthetic Theory* (Madison: The University of Wisconsin Press, 1974), 128–9; and Guyer, *op. cit.*, 280–2.

9. See Guyer, *op. cit.*, 279–307; and Karl Ameriks, "How to Save Kant's Deduction of Taste," *Journal of Value Inquiry* 16:1982, 295–302.

10. See, for example, Anthony Savile, *Aesthetic Reconstructions: The Seminal Writings of Lessing, Kant and Schiller* (Oxford: Basil Blackwell, 1987), 145–6, 185; and Christel Fricke, *Kants Theorie des reinen Geschmacksurteils* (Berlin: Walter de Gruyter, 1990), 118–19, 173.

11. This appearance of paradox does not, of course, attach to the expression '*gemeinen Menschenverstand*' [common or ordinary human understanding], which is also frequently rendered in English as 'common sense'.

12. Jens Kulenkampff, *Kants Logik der ästhetischen Urteils* (Frankfurt/M.: Vittorio Klostermann, 1978), 106–10.

13. For this view of the structure of Kant's argument, although not the details, I am indebted to the work of Christel Fricke, *op. cit.*, 152 ff.

14. This is based on the account of judgment given in the Introductions. As Kant there argues, it is only reflective judgment that can have a principle of its own; but this principle must be subjective, since it is merely a rule for reflection, which, as such, does not bring with it any ontological implications. See, for example, the discussions of the heautonomy of judgment (5:185; 25, and 20:225; 414), which are intended to make the point that judgment (in contrast to understanding and reason) legislates merely to itself rather than either to nature or freedom.

15. Kant, *Logik Dohna-Wundlacken*, 24:710; English translation by Michael Young, *Lectures on Logic* (The Cambridge Edition of the Works of Immanuel Kant, New York: Cambridge University Press, 1992), 447.

16. Although I cannot develop the point here, I take the distinction between harmony of the faculties in general and a free harmony, which occurs only in free

play, to be essential to the proper understanding of Kant's position. Thus, there can be a harmony without free play (in the case of a cognitive judgment) and a free play without harmony (in the case of an aesthetic judgment of ugliness). Unfortunately, this point tends to be ignored in the literature, with the expressions "free play" and "harmony" being taken as virtually equivalent. One author who does emphasize the distinction and its significance is Fricke, *op. cit.*, 50 ff.

17. Of course, it is not *only* cognitive concerns that are suspended, since, as we learn from the First Moment, such contemplation must be disinterested.

18. See, for example, Guyer, *op. cit.*, 86 ff.; Kulenkampff, *op. cit.*, 95–6; Ralf Meerbote, "Reflections on Beauty," in *Essays in Kant's Aesthetics*, eds. T. Cohen and P. Guyer (Chicago: University of Chicago Press, 1982), 82; and Ameriks, *op. cit.*, 299–300.

19. A possible objection to this interpretation is provided by the discussion of the idea of a universal voice in section 8 of the Analytic. Kant there states that, "Whether someone who believes he is making a judgment of taste is in fact judging in conformity with that idea may be uncertain; but by using the term beauty he indicates that he is at least referring his judging to that idea, and hence that he intends it to be a judgment of taste. For himself, however, he can easily attain certainty on this point [*davon*], by merely being conscious that he is separating whatever belongs to the agreeable and the good from the liking that remains to him after that" (*Judgment*, 5:216; 60). Admittedly, this passage is ambiguous, but I think it reasonable to take Kant to be claiming that what we can be certain of by attending to our mental actions is whether we have even intended to make a judgment of taste (since this requires abstraction from the agreeable etc.), not that we have in fact judged according to the universal voice. Otherwise expressed, what we can be certain about is that we have made a judgment of taste, not that we have made a *pure* judgment.

20. I discuss the issue of a moral requirement to develop taste in "Beauty and Duty in Kant's *Critique of Judgment*," *Kantian Review* 1:1997, 53–81. For my views on the inscrutability of moral motivation, see *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), 118, 159, and 176–7.

21. See Guyer, *Kant and the Claims of Taste*, 260, where he states that the task of the Deduction is to justify "the epistemological presumption of claiming to know about the subjective states of others."

22. In the recent literature, the readings of the Deduction closest to the one advocated here are by Claude MacMillan, "Kant's Deduction of Pure Aesthetic Judgments," *Kant-Studien* 76:1985, 43–54; and Fricke, *op. cit.* On page 152, note 1, Fricke remarks that the deduction of the principle of taste is closer to that of the second than of the first *Critique*, since it consists in its legitimization against skeptical objections from the side of empiricism rather than the proof of a validity claim. I am in general agreement with this, but would add that Kant's target here is not only empiricism, since he is equally concerned to defend the legitimacy of the principle of taste against those rationalists who wish to ground judgments of taste in a principle of perfection.

Kant in the 1760s: Contextualizing the “Popular” Turn

John H. Zammito

Suppose Kant had not been Professor of Logic and Metaphysics in the University of Königsberg. Suppose rather that, like his friend Berens, he had become a businessman; or like his pupil Herz a physician; or like his neighbor Hippel a civic administrator; or like his brother Johann Heinrich a pastor. Suppose, further—and this is easier to suppose—he had like the first three followed a literary avocation. We may presume that he would have had a reputation in his time as good as that of many men we have mentioned in this book, and would now be as forgotten as they.

—Lewis White Beck¹

I. Rethinking the Popular Philosophy of *Hochaufklärung*

It is only right to laugh along with Heinrich Heine’s witticism that Kant had no life and therefore required no biography or history. At the same time, it is important to make a place in contemporary Kant scholarship for historical contextualization alongside contemplation of the “timeless” arguments of the *Königsberg Aufklärer*.² Kant had a history and a biography; he lived in his time, embodying and contesting it in richly nuanced ways. My essay is a contribution to the historical reconstruction of Kant’s thought.³ It is also homage to one of America’s most discerning contextualizers of Kant, Lewis White Beck.⁴

Concretely, I am interested in approximately one decade in Kant’s history and biography (1762–73), during which his compass wobbled wildly from the *telos* of the critical philosophy we have ever since enshrined. Everyone knows Kant had a moment, in the 1760s, when his skepticism

concerning metaphysics seemed to be turning him into a "Prussian Hume."⁵ Relieved by Kant's "critical turn," historians of philosophy have passed over this moment as a mere aberration after which Kant found his true path and went on, consistently for the next thirty years, to create the critical system. Even so historically sensitive a Kant scholar as Giorgio Tonelli may be taken as exemplary of all this:

The essayistic style became fashionable only at the time of the so-called 'popular philosophers' such as Mendelssohn, Meiners, Tetens, etc. Kant, in his pre-critical period, shared this view; but he was soon converted to a very new kind of systematic spirit, which would henceforth inform his critical philosophy.⁶

Who could dispute that the Kant of critical philosophy is a figure of world-historical significance? What, on the other hand, of that Kant of the 1760s, of whom Lewis Beck once offered us such a charming, yet at the same time dismissive conjecture? A Kant who never made the "critical turn," Beck assures us, would have been a thoroughly representative figure of German Enlightenment thought, but scarcely worth a mention in the history of philosophy.⁷ Would he really have been as insignificant as Beck makes out? I propose to reconsider this intractably precritical Kant, to ask whether he does not deserve a larger stature in the history of the German *Aufklärung*.

I suspect that we have misplaced something hitherto in our conception of the critical turn. We have not been careful enough in determining the question: What was Kant turning away from?⁸ The traditional claim has been that it was Wolffian school metaphysics. But that turn had happened already by 1762–3. Dieter Henrich has given us a very persuasive argument that in 1762–3 Kant achieved the first systematic form of his philosophy. But in a passage that is crucial for this essay, he adds: "that Kant, the metaphysician, went silent before the end of 1763 and hid himself for the next seven years in elegant and popular occasional pieces, should be just as strong an indication that Kant's system-conception of 1762 went into crisis."⁹ Ernst Cassirer portrays it virtually as a calamity: "The phase that Kant's evolution as thinker and author entered after the writings of 1763 dashed the hopes of the world and of his friends most strikingly."¹⁰ I am interested in a less dire view of that crisis and the seven years of a "different" Kant that ensued before Kant made the "critical turn."

I propose to consider whether what Kant turned away from in 1769–70 was not in fact the path of a popular philosopher. Had Kant pursued this other path, I am convinced, it would not be nearly so easy to dismiss the whole idea of popular philosophy in Germany as trivial.¹¹ *Aufklärung* took a decisive turn in Germany around the middle of the eighteenth century. *Schulphilosophie*, especially in its orthodox Wolffian form, ceded leadership to what came to be called *Popularphilosophie*.¹² A way to put the central dispute over *Schulphilosophie* is to say that by mid-century its critics doubted that it could serve the grand purpose of *Aufklärung* any longer. Philosophy seemed too important to be left to such philosophers. We must be careful, in assessing this moment of change in the middle of the eighteenth century associated with the rise of *populäre Philosophie*, that the latter not be underestimated, construed merely as vulgar philosophy, as philosophy conducted with insufficient intellectual rigor.

The image of popular philosophy has been notoriously colored by Kant's vehement rebuttal, in the *Prolegomena* (4:372–83), of the Garve-Feder review of his first *Critique*. Beck's handling of the popular philosophers is a typical modern instance; he can see them as little more than "lightweight opponents of Kant." Their concerns, Beck condescends, "found suitable expression in pretty essays, diaries, private correspondence destined to be published and autobiographies." Their "superficial conception of the role and scope of philosophy was destroyed by Kant," he asserts.¹³ There are grounds for a more nuanced historical and philosophical treatment. Let us rivet our attention on the idea of philosophy for the people, i.e., intellectual advocacy of the self-constitution of society, that sense of Enlightenment that Kant was eventually to advocate publicly as "emergence from self-inflicted immaturity."¹⁴ *Popularphilosophie* signified "Philosophy for the World," a redefinition of philosophy's mission, away from the traditional preoccupation with epistemology and metaphysics, with theoretical knowledge and its certainty, to a new ethical and socio-political agency for change and progress. Those very terms, change and progress, rendered historical the standards of philosophy and set as the highest criterion precisely the "vocation of man."

Periodization can provide a very useful structure here. We can distinguish three phases of "popular philosophy" over the German eighteenth century. The protracted sniping at Kantianism in the *Spätaufklärung* for which popular philosophy is best known appears only as its third (and least interesting) phase. The first phase was associated with Christian

Thomasius and the *Frühaufklärung*. Thomasius called for “an easy manner, comprehensible to all rational persons of whatever station or gender.”¹⁵ That entailed a radically pragmatic revision of reason: “Philosophy became with Thomasius the critical application of sound human understanding to historical circumstances.”¹⁶ Thomasius wanted to create a general German culture, and to do so he felt compelled to break free from the stranglehold of theologically orthodox university culture. This was the same motive that led him, roughly simultaneously, to offer the first lecture in the German language at a German university, and to establish the first popular journal in German. His sustained efforts to revolutionize the university earned him a clutch of enemies within the *Gelehrtenstand*; but his reputation for pedagogical innovation and political relevance made him very attractive to the Great Elector of Prussia when the latter was thinking about establishing a new university. Thomasius became instrumental in every aspect of the creation of the University of Halle.¹⁷ His eclecticism, orientation to practical reason, and emphasis on jurisprudence fit perfectly with the aggrandized ambitions of the state.

Thomasius launched a movement, and Beck has traced its lineage: “Four men stand out in a direct line of succession from Thomasius: [Johann Franz] Budde [1667–1729], [Andreas] Rüdiger [1673–1731], [Adolf Friedrich] Hoffmann [1707–1741], and [Christian August] Crusius [1715?–1775]. Budde was a colleague of Thomasius, Rüdiger a student, Hoffmann was Rüdiger’s student, and Crusius was Hoffmann’s.”¹⁸ For a span, Thomasius’ eclectic program of enlightenment worked in tandem at Halle with Pietism, which resisted dogma in the name of concrete moral-religious life. But Pietism itself soon hardened into orthodoxy. Concurrently, Thomasians on the Halle faculty began developing metaphysical predilections of their own, putting a strain on pragmatic pedagogy. Already Budde reverted to the type of the Polyhistor, generating yet another “‘school philosophy’ with many scholastic trimmings,” in the words of Beck.¹⁹ The result was the famous quarrel between Wolff and the Pietists that resulted in his dismissal from Halle in 1723.²⁰

Ironically, this dismissal made a martyr and a hero of Wolff, and his ideas became even more influential. For a span of some thirty years, from 1720 through 1750, *Gelehrsamkeit* in the guise of Wolff’s new rigor or theoreticity prevailed. Such rigorous science could only conceive itself as systematic, requiring demonstrative proofs of all propositions. Thomasius’ idea of a practical, eclectic, engaged style of reason went into eclipse. To enter the higher realm of truth, Wolff and his school main-

tained, one had to get beyond mere *cognitio historica* to the autonomous realm of the a priori, i.e., to *cognitio philosophica* and *cognitio mathematica*.²¹ Wolff's background was in mathematics and the natural sciences, and he brought with him into his philosophizing the ideal of *l'esprit géométrique* so central to Descartes and the seventeenth century metaphysicians.²² Philosophy and natural science needed to approximate mathematics in its structure of argumentation and proof. Grounding—*begründen*—was what distinguished philosophy, drawing it close to the lucidity of mathematics. Though not identical, Wolff saw more in common between metaphysics and mathematics than would his critics. By contrast, mere *cognitio historica* suffered two fatal deficiencies. Either it was literally derivative—second-hand or "book-knowledge"—or it was all-too-unreflected, "*nuda facti notitia*," immediate sense experience.²³ In both instances, *cognitio historica* lacked rigorous formulation of arguments or explanations according to principles. Kant was entirely Wolffian in this regard; he celebrated Wolff just for this in the B-Preface to the *Critique of Pure Reason* (Bxxxvi), as the figure who had secured the "*Geist der Gründlichkeit*" in Germany.²⁴

Wolff challenged the traditional ascendancy of theology, jurisprudence, and medicine in the university by claiming they offered merely a historical knowledge of their fields, whereas "only philosophy could raise this knowledge to the theoretical level of a knowledge of principles."²⁵ But his proposals for revision could not dislodge the traditional university organization of the eighteenth century. Nor was traditionalism the only reason he failed; Wolff's vision had an energetic, creative rival. German universities in the mid-eighteenth century turned towards empiricism and especially toward "an increasingly historical orientation in all of the sciences," according to Hammerstein, and this blocked Wolff's ambitions for philosophical hegemony.²⁶

By 1750, Wolff's grip had slackened. Three impulses conjoined to spawn opposition even before his death. First came the immanent philosophical objections to his system, especially to the idea of the mathematical method in philosophy.²⁷ Second came the massive incursion of foreign thought, both French and British.²⁸ Finally, there was the retrieval of Thomasius' idea of "eclecticism" precisely as a resource to bring the philosopher, the university scholar in general, "down to earth" or back into the "world."²⁹

From around mid-century, a second phase of popular philosophy rose up against Wolffian *Schulphilosophie*, and it dominated the

Hochauflklärung (1750–80).³⁰ An event of 1754 has drawn historical attention both for its coincidence with the death of Wolff and for its explicit advocacy of “popular philosophy.” In his Rectoral Address at the Thomasschule in Leipzig, entitled *De Philosophia Popolari* (*Of Popular Philosophy*), Johann August Ernesti (1707–1781) took up the call of the key French philosophe, Denis Diderot, “*Hâtons-nous de rendre la philosophie populaire* (Let us hasten to make philosophy popular).”³¹ Mortier notes that for a major classical philologist at Leipzig specializing in Cicero, Ernesti was radical in attaching no privilege to Latin over against vernacular languages. He urged that esoteric metaphysics be replaced by a more accessible and a more useful style, a humanism.³² Ernesti insisted that “true philosophy” had to be “a [mode of] thought in close contact with society . . . necessarily open to non-specialists.”³³ Thus he took up the criticism of the “scholastic” university and fashioned a conception of philosophy that accorded far better with a *philosophe* of the style of Diderot than with the established German *Gelehrte* like Wolff (though Ernesti did not mention Wolff by name).

Even more central to the genesis of *Popularphilosophie* was the new University of Göttingen. From the outset Göttingen (est. 1737) was hostile to Wolffian *Schulphilosophie*. Its founder, Freiherr von Münchhausen, a product of Thomasius’ University of Halle, deliberately set about creating a university which would pursue “a sort of new eclecticism.”³⁴ Beck notes that “the atmosphere of the university . . . concentrated on the *realia*, especially history, philology, pedagogy, science, and law,” and hence in his view made no contribution to disciplinary philosophy.³⁵ Bodeker stresses the segmentation of disciplines within the philosophical faculty at Göttingen.³⁶ This reconstitution of disciplinary inquiry was most prominent in philology, in history, and in the natural sciences.³⁷ The University of Göttingen was also decidedly shaped by the Hannoverian dynastic connection with England, so that intellectual styles of research and of discourse took on a decidedly British tone of empiricism and anti-speculative “common sense.”³⁸

Insistence upon “down-to-earth” empirical-historical research characterized the scientific canon of Göttingen in its ascendancy in the mid-eighteenth century.³⁹ It is not surprising that it would become a center of opposition to *Schulphilosophie*. Especially in the hands of Feder and Meiners after 1768, the “Göttingen program,” as it has been called, represented the most aggressive agenda for *Popularphilosophie* in Germany in the *Hochauflklärung*.⁴⁰ Indeed, this Göttingen school had been prominent

for a generation before its infamous and catastrophic tangle with Immanuel Kant's critical philosophy, 1781–1783.⁴¹ Meiners is especially pertinent for grasping the Göttingen agenda before Kant's challenge. What Zimmerli calls an "explosive element" in the Göttingen program found articulation above all in Meiners's (anonymous) *Revision der Philosophie* of 1772:

Once one can show that no philosophical opinion in the tradition has any advantage over any other, that all of philosophy can be transformed into the relativism of history, one forces oneself and others to think independently in coming to a position: "The great advantage of this method, transforming all of philosophy into mere philosophical stories, would be without question the healthiest imposition which one can make upon one's audience to think for themselves."⁴²

The challenge of "historicity" to philosophy could not be stated more explicitly.

The opposition to Wolffian *Schulphilosophie* was simultaneously and more essentially an opposition to university *Gelehrsamkeit* and to its representatives, the *Gelehrtenstand*: "a clearly articulated, public hostility toward the erudite scholarship of the universities."⁴³ What was at stake is the grander question of *Bildung* versus *Gelehrsamkeit*, but the specific philosophical position of Wolff and his disciples got caught in this crossfire. That was captured in a negative twist to the key phrase, *Schulphilosophie*, which came to mean, first and foremost, enclosed thinking: "closed minds" associated with closeting in the schoolroom, and with guild restrictions in society.⁴⁴ Turner elaborates: "In the name of *Aufklärung* critics denounced the universities for their outmoded, medieval constitutions and their pedantic curriculum still mired in Wolffian philosophy, theological dogmatism, and the Latin *imitatis*."⁴⁵ In this context, Wolffian *Schulphilosophie* became identified with a far older image of intellectual life, that of the "Schulfuchs," and by the mid-eighteenth century no longer esteemed fit for the wider context into which intellectual discourse needed to insert itself. It suffered continuous persiflage: "[b]etween 1750 and 1780 the very word *Gelehrsamkeit* was degraded to a term of satirical abuse."⁴⁶ Lessing contributed to this vein with his early play, *Der junge Gelehrte* (1748). C.G. Salzmann's novel, *Carl von Carlsberg oder über das menschliche Elend* (1784) was a prominent late instance of such satire. Herder, not surprisingly, proved eminently proficient in this vein, but what may surprise us is how accomplished the early Kant proved in it.

The point of this "general satire of the scholars" was "to take a stand against the [ideal of] scholarship created by study in the arts faculties [and identified concretely] with professors and masters." Thus the whole "chastisement and contempt of the scholar" presupposed the vantage of the wider bourgeois order, and its demands for a participatory role in learning.⁴⁷ The core idea here was that culture must be "fundamentally available to all."⁴⁸ Both the style of articulation and disputation, and the institutional framework of the *Gelehrtenstand*, came under such fierce attack over the second half of the eighteenth century in Germany that neither escaped drastic reformation.⁴⁹ In the words of Werner Schneiders, "School philosophy became conscious of its open flank to the world."⁵⁰ It needed to transform itself into the vanguard articulation of a new public's aspirations to practical success and happiness in the world.⁵¹ Philosophers were not to restrict themselves to discourse with other philosophers only, but rather address themselves to the wider public and thereby acknowledge a different identity, namely that of a "*reflektierende Schriftsteller*."⁵² The sphere of intellectual expression and disputation expanded forever beyond the walls of the university, and the style and organization of that expression altered dramatically to accommodate the wider public sphere into which it fit. It was to be a "*Philosophie für die Welt*" (where, to be sure, world meant really the bourgeois reading public).⁵³

Beyond the university, things were stirring. The impulses that Thomasius set in motion with his popular journal gathered into a huge tide of such journals, the "moral weeklies," between 1740 and 1770, creating the reading public that Thomasius could only envision. One of the most important achievements of the "moral weeklies" in the early eighteenth century was to wean this new bourgeois reading public away from the religious confines of traditional culture on the one hand, and from acquiescence to representations of courtly and aristocratic preeminence, on the other.⁵⁴ There were dramatic shifts in the structure of reading. Religious texts came to be displaced by cultural ones, Latin texts by German ones. Rolf Engelsing has suggested that a shift took place from "intensive" to "extensive" reading, i.e., from the close and repeated reading of a few texts to the continuous assimilation of ephemeral reading material, above all in the form of periodical literature.⁵⁵

Not only was this the period in which Germany experienced a dramatic expansion of literacy; it was the time in which a literary market of modern institutional form took shape.⁵⁶ The moral weeklies having largely

completed their work by mid-century, a new, more critical style of periodical literature came to succeed them.⁵⁷ Berlin became the prime center for the emergence of such journals of the "beautiful arts" and the "beautiful sciences." There Nicolai, Lessing, and Mendelssohn created an intellectual context for the German public's reception of ideas from all Europe. This represented a decisive turn in the cultural history of Germany. In the 1750s, these three figures, and above all Mendelssohn, became the voice of popular philosophy in Germany.⁵⁸ They—with their allies, Sulzer, Resewitz, Abbt and others—established the program of *Hochaufklärung*; that is, they had a definite agenda for progress, freedom of expression, and an informed public opinion.⁵⁹

One unity that tied the new reading public together was that education was indispensable for its practices, and not only in the economic or occupational sense. Rather, education was the personal distinction, the figuring of self, out of which individuality and career achievement could be wrought. Reading was essential to identity-formation: that one read, how one read, what one read—all these marked one as *gebildet*.⁶⁰ Reading was for self-discovery or self-formation (*Bildung* had that sense at its core); but it was also for orientation to the world (*Weltkenntnis*, a term of decisive importance for Kant), and through these first two processes it led to the achievement of urbanity (*Weltbürgerlichkeit* in the widest sense of "cosmopolitanism"), a self-confident sense of attunement to, and participation in current affairs, in the "public sphere."⁶¹

The new *Bürger* needed to achieve independence in judgment: *Selbstdenken*. To be capable of thinking for oneself was to have achieved "maturity," *Mündigkeit*.⁶² To achieve *Selbstdenken*, to become *mündig*, was a matter of personal knowledge or the cultivation of individual judgment, which could only be successful by drawing upon all of one's own life experience, the physical and the emotional alongside the logical-cognitive. Here such terms as *Sinnlichkeit*, *Beobachtung*, and above all *Erfahrung*, whose technical philosophical senses are central for Kant, took their decisive social place. In all these, the new mentality stressed the necessity for sharp observation, for a keen notation of particular instances, and hence for a process of learning that built upon empirical and historical accrual.⁶³ Here *cognitio historica* was not inferior, it was indispensable. And the most important point of all was that this process of learning constituted character-formation in that distinctively pragmatic-moral sense that proved central to the specific undertakings of Kant. The pragmatic end was to achieve fulfillment or happiness within a commu-

nity that obviously posed competitive challenges and constraints, even as it offered recognition and reward. The means were the cleverness and the effectiveness to cope and indeed prosper in this world, along with the (moral) wisdom to do so with integrity. Here was the crux of the critique of *Gelehrsamkeit*. Philosophy needed to be reoriented, made "pragmatic." Of course, "pragmatic" is not identical with "practical" in Kant's sense. The tension between virtue and success, the relation of character to prosperity, proved the central theme of empirical psychology, of anthropology and of ethics for *Aufklärung*, and for Kant in particular, over the second half of the eighteenth century.⁶⁴

But the problem was not merely theoretical; it was existentially urgent and it was socially mediated. One could neither pursue nor achieve *Bildung* in isolation. Identity is always social. *Mündigkeit* could only be demonstrated in practice. The ideal of sociability was of immeasurable importance for the emergent *gebildeten Stände* in the *Hochaufklärung*.⁶⁵ The impetus of reading was outward from the subject into the social world. It is crucial, then, that the first and foremost instrumentality of this wider social practice was the formation of *Lesegesellschaften*, "reading societies."⁶⁶ The reading of periodical literature was a collective process: "what had been read was also discussed in a critical manner."⁶⁷ Training in listening and in both agreeing and disagreeing helped constitute participatory-democratic political orientations.⁶⁸ These impulses were actualized further in other forms of voluntary association that moved increasingly towards social practice and concrete politics.⁶⁹

This is especially true of the so-called "patriotic societies." The "patriotic societies" sprang up especially in the 1760s with demands for civic improvement, social welfare, and political reform. They are a decisive indication of the stirrings towards politicization in the German middle class.⁷⁰ Moses Mendelssohn, invited to join the Patriotic Society of Bern, responded with some striking observations in Letter 138 of *Briefe, die neueste Literatur betreffend* (1761):

Philosophy is at home in Germany. . . . Nor do we lack systems of political science. . . . It is just that our political philosophy still retains an element of timidity, which it will naturally only lay aside in a free state. Its true fatherland is there where it does not have to shrink from openly opposing arbitrary violence and hypocrisy. . . . We must admit, in all fairness, that several Swiss writers were among the first Germans to begin looking at people in the context of a wider political society from a truly philosophical viewpoint.⁷¹

The announcement by this Bern patriotic society of a prize competition on the role of philosophy in promoting the good of the people proved quite important to Mendelssohn and Herder.⁷² Certainly, such ideas and ideals were in considerable tension with the principles of territorial absolutism which otherwise prevailed in the Germanies. Beck notes this only under the negative rubric of the "unpolitical German":

Unlike the French philosophes to whom they are sometimes compared, the German popular philosophers' interests were focused on the problems of private sentiment and virtue, religion, education, and art. . . . [T]he concern was with the individual who needed, perhaps, a new kind of education but not a new kind of government.⁷³

Concentration on the private world, the world of human social engagement apart from political authority, has been the source of a longstanding interpretive controversy over the "unpolitical" German.⁷⁴ As a persuasive rejoinder, we can invoke Rudolf Vierhaus:

The development of political consciousness . . . involve[s] a complicated process of reconstruction of experiences, of learning and carrying forward, of sensitization to things that earlier appeared self-evident or not even noticeable, of recognition of contradictions, of making opinions possible and utterable and of the formulation of goals. It takes place in a context of general intellectual changes and shifts in mentality, which are themselves inseparably embroiled in concrete socio-cultural and socio-economic change.⁷⁵

The cultural history of the *gebildeten Stände*, especially from the mid-point of the eighteenth century, demands this more sophisticated approach to "*Aufklärung* as politicization."

What Ernesti called for in 1754, in Leipzig, had already been underway for a number of years in Berlin—a city, it should be noted, with no university but with the most important Academy in the Germanies. One of the unique features of the Berlin Academy was that it had a section for speculative philosophy.⁷⁶ Maupertuis, its first President, had found that somewhat suspect, but it fit with his patron Frederick II's agenda, to use the Academy to undermine the established *Schulphilosophie* by drawing in the wider—i.e., the French—Enlightenment. This was one of Frederick's motives for requiring the language of the Academy to be French. (Another was his contempt for German as a literary language.) In any event,

Maupertuis took up the agenda, and the Academy's sessions, and above all its prize competitions, worked consistently to bring the established orthodoxies of German *Schulphilosophie* under scrutiny.⁷⁷ Particularly central in this endeavor of the Academy was the man in charge of the section for speculative philosophy for the 1760s, Johann Georg Sulzer (1720–1779), a Swiss by origin and one of the most important figures in the rise of popular philosophy.⁷⁸ Sulzer's *Kurzer Begriff aller Wissenschaften* (1st edition, 1745; 2nd rev. ed. 1759) offered exemplary formulations of the new approach. Sulzer made very clear what his agenda signified: "'Philosophy of the world' (gen. subject) is worldly cleverness or 'wisdom' which can only be learned by experience or 'socializing in the world'."⁷⁹ Packed into that one-sentence definition is a horde of crucial connections, the whole configuration of ideas that distinguishes popular philosophy. What all these popular philosophers did was draw up a coherent agenda, articulating the desire for a new public sphere committed to social and political improvement, and placing responsibility for its cultivation and advancement with philosophers: "philosophy for the world." For the newly emergent reading public, they proposed a new and apposite style. Over against *subtilitas* ("all too much precision") they offered *popularitas* ("general comprehensibility") and over against *siccitas* (dryness) they offered *suavitas* ("a certain charm").⁸⁰

While there were "literati" outside the university, we must never forget J.D. Michaelis' point that most of the writers and thinkers of Germany throughout the eighteenth century remained associated with the university: "The majority of our great scholars actually reside in the universities, and of the rest many either once were actually professors who taught in the university or in any event in their youth enjoyed the privilege of devoting themselves to the academic life."⁸¹

Gerd Ueding confirms this observation: "Most of the philosophical essayists were in fact at the same time school philosophers. They taught at universities."⁸² But they were calling upon themselves, even as they were being called upon by others, to think in a wider frame, with a new goal, in a new style. Turner puts it very well:

The crisis of the *Gelehrtenstand* followed not from overt rivalry with groups that stood outside it, but from the gradual adoption by its own members of new intellectual and cultural values that increasingly proved incompatible with its own traditional ones. The crisis of the

Gelehrtenstand was an internal crisis, a crisis of self-image and self-justification.⁸³

This is the decisive point, directly relevant to Kant.

II. An “Altogether Different” Kant: Contextualizing the *Beobachtungen*⁸⁴

The biographical literature clearly notes that something began shifting dramatically for Kant around 1762, making his 1760s turbulent years. Vorländer writes: “Around the year 1762 a change in Kant’s inner life set in.”⁸⁵ And Arsenij Gulyga makes the same point: “1762 was a crucial year in Kant’s life. . . . [It brought on] a revolution of all his life’s objectives.”⁸⁶ Summarizing the biographical literature dealing with the 1760s, Rolf George notes, “this is a period of liberation for Kant, which left its imprint on his intellectual make-up.”⁸⁷ Richard Velkley and Frederick Beiser argue that the impact of Rousseau in just these years profoundly and permanently refigured Kant’s conception of his philosophical mission.⁸⁸

Three aspects of Kant’s biographical context in these years deserve more consideration. First, we must ask after Kant’s sense of his intellectual career. How was he doing? What were his expectations and his frustrations in this vein, not only in regards to income but in regards to status in the profession? What about Kant’s identity as a writer, i.e., what name, what fame was he making for himself in circles beyond Königsberg, indeed, beyond the professional scholarly community? How are we to interpret the nomination of Kant for a professorship of rhetoric in the mid-1760s? We know he declined the chair; what we need to ask is how he could have been considered for it in the first place, and what that tells us about Kant’s image. On this same line, what was it that so excited the faculty and students at the University of Erlangen in 1769 at the prospect Kant would accept its chair in philosophy? It was not *Schulphilosophie*. Only one book by Kant went through eight printings in his own lifetime, and it was not *Schulphilosophie* either. *Beobachtungen über das Gefühl des Schönen und Erhabenen* [*Observations on the Feeling of the Beautiful and Sublime*] became Kant’s “best read” and most successful piece of writing. It thrust Kant to the forefront of “popular philosophy” at the time, made him an exemplary “*reflektierende Schriftsteller*.”⁸⁹ Kant, in short, was being fitted for a new hat, and the question is whether he intended this to happen. What can we determine about all this?

Kant turned forty in 1764. We must consider this in light of Kant's prolonged wait for a professorial chair. He may well have been the "gallant *Magister*," but he was not yet a fully established member of the guild. In our own day, forty years of age often represents an important psychological turning point, a point of self-reckoning and even of anxiety. Is there any evidence of something like our "mid-life crisis" in Kant in these years? Did his fortieth birthday pass tranquilly, or might we detect signs that all was not entirely well? As it happens, Kant devoted a great deal of consideration to the psychology of this turning point, much of it, one can plausibly infer, derived from introspection.⁹⁰

Second, what about Kant as reader? That is, how did reading figure in Kant's self-construction as a *gesitteter Mensch*? There is a considerable volume of work on the role of reading in identity formation among the *gebildeten Stände* of the eighteenth century in Germany that deserves to be brought to bear upon its cultural producers, not just its cultural consumers.⁹¹ In particular, the role of novels in self-examination and self-conception has proven very illuminating, and this appears the case for Kant as well.⁹² But far and away the most important reading Kant ever did, if we consider its impact upon his identity and project, came at this moment in his life: his dramatic encounter with the writings of Jean Jacques Rousseau. To this we must add his important but more elusive encounter with David Hume.

Finally, and not at all unrelated to the first two considerations, Kant's socializing [*Umgang*] dramatically intensified in these years. "He was much in demand, and he never refused an invitation. As a clever and lively conversationalist Kant was the life of the party. He was natural, easygoing, quick-witted. . . ."⁹³ Kant's interest in gaming, in table talk, in hobnobbing with the rich and powerful, proved so substantial that Hamann worried whether he would produce anything. Was this merely a matter of "distraction" or were there more intimate and passionate dimensions to the life of the "gallant *Magister*"? Exploring these three questions sheds interesting light on Kant's turn to compose the *Beobachtungen über das Gefühl des Schönen und Erhabenen* in summer 1763.

Kant's success in the Berlin Academy Prize competition of 1763 did not go unnoticed by the Prussian Ministry of Culture in Berlin, or perhaps even by King Frederick II himself, though of this one has more reason for skepticism. Kant's text appeared, under the Academy's auspices, in April 1764. Several months earlier, *Beobachtungen*, Kant's most successful book, came to market. It received no fewer than fifteen re-

views in the German periodical press and would go into a second printing in 1766.⁹⁴ Thus, when in July 1764 the Professor of Rhetoric and Rector of Königsberg University, Johann Georg Bock, died, the Prussian Ministry in Berlin sent a directive to the local administration in Königsberg that made it clear that Kant was someone in whom it was interested: "We have become acquainted with a certain *Magister* Immanuel Kant through a few of his writings, from which shines forth a very profound learning. . ." ⁹⁵ The Ministry was seeking someone who would not only be able to dash off a few ceremonial poems in German, but who could strengthen the entire philosophical faculty and especially help ground students in classical—Latin and Greek—literature. It instructed its local officials to establish whether Kant was qualified for such a position and whether he would be interested. The official on the scene, Braxein, contacted Kant personally and reported back to the Ministry on October 19, 1764, "*Magister* Kant teaches with great effectiveness to general acclaim." ⁹⁶ Unfortunately, he went on, it did not appear that Kant was suited for—and he definitely indicated he was not interested in—the position. ⁹⁷ Nonetheless, Braxein proposed that Kant should be considered for the next position as Professor of Logic and Metaphysics, which he believed would soon come available. ⁹⁸ The Ministry in the King's name appointed Kant's friend, Lindner, to the position, and instructed the local administration "for the good and enhancement of your university in whatever other circumstance that might arise to find a place for *Magister* Immanuel Kant." ⁹⁹ A few days later, a second directive followed, taking note of Braxein's report: "the very competent—and according to your previous report—also highly acclaimed teacher at your university, *Magister* Kant, should be advanced at the very first opportunity." ¹⁰⁰ Kant was provided written confirmation of this preferment, November 14, 1764. ¹⁰¹ His reputation began to spread even beyond Prussia. In 1765 the University of Jena included Kant in its list of candidates for a chair in philosophy (see 13:41). Yet one cannot help reflecting that while all this was happening, Kant was writing *Träume eines Geistersehers* [*Dreams of a Spirit-Seer*], his most negative statement about professional philosophy and the academic community. Do we have here a case of too little too late? After all, Kant actually got no job offer, only the promise of "preferment" in the future.

Kant's economic circumstances figure significantly in the context of his estrangement from the academic community. They played this role especially since Kant believed economic independence was indispens-

able to personal dignity.¹⁰² In *Beobachtungen*, composed summer 1763, Kant wrote (2:213): "What increases the evil of poverty is contempt, which cannot be completely overcome even by merit, at least not before common eyes." There is a famous incident from Kant's later years in which he became offended when a French *homme des lettres* included in his "literary history of Prussia under Frederick II" an account of Kant's ostensible penury in his years as *magister legens*.¹⁰³ Particularly galling was the suggestion that Kant was invited to dine with the merchant Green as a kind of charity case. Kant responded angrily that these were the most pleasant years of his life and that he had no real economic difficulties.¹⁰⁴ Despite—or perhaps in light of so much—protest on his part, there is reason to suspect Kant's pride in later years may well have clouded his recollection of the reality of his circumstances in these earlier years.

When he took up his position as *magister legens* in 1755, Heller argues, "he could permit himself no illusions regarding his material situation even for the future."¹⁰⁵ Just attaining his initial university position proved expensive. "Considerable expense was connected with his promotion, such as the payment of fees, and the printing of the dissertations; in meeting this he was aided by his uncle Richter."¹⁰⁶ As *magister legens* Kant received no salary from the University; faculty of his rank depended entirely upon student subscriptions to their courses for income. These courses were not even listed in the university catalogue at the time Kant was serving in this rank. Instead, instructors had to print up their own course advertisements, obviously at their own expense. Jachmann notes: "In the first years of his private-docenture at the university the earnings from his lectures were very small and he had to take such abstemious care of himself that not infrequently he became embarrassed over his living conditions."¹⁰⁷

To make enough money, Kant undertook an unbearable teaching load. Gulyga writes: "While a *magister* he had to teach four to six subjects a semester. Each week he taught at least sixteen hours. . . . In the second half of the 1750s he writes hardly anything; lectures exhaust his time and his strength."¹⁰⁸ Jachmann reports that Kant had managed to set aside 20 gold pieces as his emergency fund, which he would not touch. But everything else had to go, including a nice library he had built up in his years as private tutor.¹⁰⁹ Ritzel questions the plausibility of this report, doubting that Kant could have earned enough as *Hauslehrer* to have accumulated a library, and doubting further whether selling it off could have been enough to cover Kant's expenses.¹¹⁰ But that only deepens the mystery of Kant's early livelihood. Arnoldt offers the suggestion that, in these years, Kant

had a pensioner, a student who lived with him while attending university to help defray his costs.¹¹¹ There is the prospect, too, of private tutoring of students at the university. But the fact is that we do not know more than that Kant's means were very strained.

Still, Gulyga argues that, by the end of the 1750s, Kant "no longer has financial worries. The *Privatdozent* can afford a servant: he takes a retired soldier, Martin Lampe."¹¹² Vorländer reports that Kant took on the ex-soldier, Lampe, by the early 1760s, and a letter from Kant to Borowski from 1761 confirms that he did have a servant by that point.¹¹³ What changed Kant's financial circumstances? A strong likelihood is that it was the patronage of the Russian military. Moreover, once Kant had established this connection, it proved easy to maintain even after the Russians left East Prussia, by switching to the Prussian military and to the higher aristocracy. Nevertheless, these private sessions were unpredictable and they also represented an additional burden on top of all the other work Kant had to undertake. There was simply no secure source of income. Even if student enrollments in his privately offered courses grew over the 1760s, Emil Arnoldt, who studied Kant's teaching at the University of Königsberg more closely than anyone, observed that Kant's enrollments were never substantial, even after he was already *Ordinarius*. Thus we have reason to question whether Kant was being honest in claiming that he was richly compensated by his student fees in the 1760s.

The best reason we can have is Kant's own (self-interested) claim from these very years that his situation was very strained. We have that claim in the letters Kant addressed to King Frederick II and to Freiherr von Furst und Kupferberg in 1765, in connection with his application for an appointment as assistant librarian at the state library in Königsberg. Kant wrote to the king requesting "a gracious contribution to the lightening of my very wanting subsistence at the university here."¹¹⁴ To Freiherr von Furst und Kupferberg Kant wrote of "my very uncertain academic subsistence."¹¹⁵ Granting that Kant needed to plead his case, one should consider that Kant would go to such lengths for a mere 62 thaler per year.¹¹⁶ When he declined the Prussian Ministry's offer of a chair in rhetoric in 1764, Kant had to contemplate a future of protracted economic limitations. For no other reason did he take on the unpleasant job of assistant librarian, for that Kant disliked the job all his biographers attest. He dropped it as soon as his income permitted.

Even after his appointment to the professorship Kant was hardly prosperous. Kant's salary in 1770 was 400 thaler per year.¹¹⁷ He supplemented

that with the student subscriptions to his private lectures, and for the first two years he retained his library position as well. From what we know of academic salaries in Germany at that period, this cannot appear a particularly lavish income. Indeed, the offer from Erlangen (1769) would have brought Kant substantially more. Kant's salary should be compared with the data we have of the salaries for professors in the higher faculties in Germany, and with the data for those in the philosophy faculty in other universities. The philosophical faculty earned salaries drastically inferior to those of the higher faculties.¹¹⁸ Moreover, the latter faculties could supplement these salaries substantially by practicing their professions outside the university.¹¹⁹ All the philosophical faculty could do of a parallel nature was teach in the lower schools, but these were even more miserably paid than the university. The result was an economic and even status divide which was severe and widening through the mid-eighteenth century. In Hammerstein's words, "At the bottom of the ladder then as ever stood the arts faculty, and indeed so much so that some of its members barely earned enough to feed themselves."¹²⁰ They needed to string together a series of appointments to create a standard of living that was bearable. Thus professors accumulated chairs and other sinecures to make up for the inadequate economics of the universities generally. At Erfurt in 1778, for example, half the philosophy faculty also taught in secondary school and roughly the same percentage held more than one sinecure.¹²¹ Only at Göttingen was this practice somewhat overcome.¹²² In that light, it hardly appears that Kant was as "richly compensated" as he would have liked to pretend even after attaining his professorship.

In sum, we have every reason to believe Kant to be growing discontent with his career, perhaps even with his calling. Reading Rousseau proved catalytic for precipitating this discontent. That more than a trace of anti-intellectualism characterized the writings of Jean-Jacques Rousseau should be no surprise. Kant has been said never to have had an anti-intellectual moment.¹²³ I think that is quite wrong. He did have at least one. Rousseau did not elicit it, but he did offer decisive confirmation for it. One of the central features not only of Kant's *Bemerkungen* but of his university lectures at least through the 1760s was his vehement questioning of the place of "science" [*Wissenschaft*—with the wider sense in German of "scholarship"] among human concerns, and more specifically of the costs of a personal "scientific" commitment. What calls for examination is the individual form in which Kant articulated his disaffection from the scientific vocation, his personal experience of intellectual *acedia*. Certainly,

the impact of Rousseau in this aspect of Kant's identity is not to be denied. In a famous passage from his *Bemerkungen*, Kant observed that he was "by disposition a researcher," but that this had inspired in him an unjustified elitism, a contempt for the "masses without learning." Explicitly he acknowledged Rousseau's influence in "setting him right" (*Bemerkungen*, 38). I wish to suggest the possibility that Rousseau—the gadfly, the critic of Enlightenment—seemed to offer Kant the model of an alternative intellectual identity. As Goetschel has noted, "In his encounter with Rousseau, Kant formulated his concept of himself as a writer."¹²⁴

That Kant had Rousseau in mind in his remarks on science even in his Herder Lectures is clear, since he mentioned him by name: "Were [man] to be concerned with life he would never need science. Woe to him! (Rousseau) He finds [only] phantasms" (28:892). Yet there were already a variety of impulses in Kant's own life experience carrying him towards that conversion moment for which, to be sure, Rousseau was the decisive catalyst. To grasp this dialectical interweaving of theoretical and personal dimensions, we must link Kant's own critique of the "scientific vocation" with his assimilation of Rousseau.

Rousseau's *First Discourse* (1748; published in 1751) had a tremendous impact in Germany. It made its impact first through a review by Lessing in 1751, whereupon a translation speedily followed (1752), widening access to the "paradoxical" text.¹²⁵ Kant seems hardly likely to have missed the hoopla.¹²⁶ The literature on Kant's reception of Rousseau has concentrated, appropriately enough, on the mid 1760s, but that does not signify that Kant would have been unaware of him earlier.¹²⁷ Indeed, there is evidence—e.g., the casual mention of Rousseau's work by Hamann in a letter to Kant from 1759—that points the other way.¹²⁸ The remarks on the hypertrophy of knowledge in modernity which played such a striking role in Kant's "*Spitzfindigkeit*" essay of 1762 seem unquestionably linked to the Rousseau of the *First Discourse*.¹²⁹ What we must ask is whether Rousseau's works might have helped fuel the mounting wave of discontent that would impel Kant from a primarily natural-scientific identity in the 1750s to preponderantly practical pursuits in the 1760s.

Certainly we must believe that Kant was already enthusiastically embroiled with Rousseau when Herder first joined his classes in August 1762, for the latter completely reconstructed his regimen around Rousseau explicitly under Kantian inspiration.¹³⁰ By then, Kant had certainly digested the two *Discourses*, together, in all likelihood, with the *Letter to*

D'Alembert. He had also read Rousseau's novel, *Julie, ou La Nouvelle Héloïse* (1761).¹³¹ In fact, just before Herder began attending Kant's classes, in the late summer of 1762, Kant obtained his copy of Rousseau's *Emile*, and plunged into his legendary absorption in and of that work.¹³² Significantly, Kant's reception of Rousseau was far closer to the general public's enthusiasm than to the scholarly community's hostility.¹³³ Indeed, we have every reason to believe that Kant was never so enthusiastic in his reading of any other thinker. This does not mean that Kant would not become a critical and discerning reader of Rousseau, or that he assimilated the latter's ideas indiscriminately. It does mean, however, that we do not find in Kant, as we do in so many of his contemporaries—especially in Germany—a scandalized concern over Rousseau's impieties, sacred and profane. As Raymond Trousson observes, "From 1762 to 1766 the German press was indeed very occupied with Rousseau but, it is essential to recall, in a manner entirely negative."¹³⁴ Kant in fact represented one of the most important positive avenues of Rousseau reception in Germany. No other German could have written, in that epoch, what Kant wrote in his *Bemerkungen* (37–8) around 1764:

The first impression of the writings of Mr. J.-J. Rousseau received by a knowledgeable reader, who is reading for something more than vanity or to kill time, is that he is encountering a lucidity of mind, a noble impulse of genius and a sensitive soul of such a high level that perhaps never an author of whatever epoch or of whatever people has been able to possess in combination.¹³⁵

Yet that very same passage continued with criticism of Rousseau that took account of what troubled so many of his contemporaries (*Bemerkungen*, 38):

The impression that immediately follows is bewilderment over the strange and contradictory opinions, which [so] oppose those which are in general circulation that one can easily come to the suspicion that the author, by virtue of his extraordinary talent, wishes to show off only the force of his bewitching wit and through the magic of rhetoric make himself something apart who through captivating novelties stands out among all rivals at wit.¹³⁶

Kant's reception of Rousseau had to pass from an initial stylistic bedazzlement gradually to a cool, critical appraisal (*Bemerkungen*, 28). The cli-

max of Kant's crucial *Auseinandersetzung* with Rousseau came in 1764–5.

The impact of Rousseau on Kant was so enormous—Schmucker calls it a "revolution" and describes it as "overwhelming" Kant, and Henrich agrees—that it cannot be encompassed here.¹³⁷ What is of importance for this essay is the way in which Rousseau reinforced Kant's disaffection from the learned estate by offering a theory of history and culture in which scientific pursuits were denuded of a great deal of their conventional glory. Kant was of course quite proud of his scholarly achievements. He saw them as a mark of his "natural superiority" and at the same time as the fruit of his disciplined application (thus, implicitly the legitimation for his upward social mobility). "Use of the understanding is a skill [*Fertigkeit*] not just a capacity [*Fähigkeit*]. [N]aturally inferior: the one who does not have this skill (to be distinguished from inferior socially: many who are socially inferior can be naturally superior and conversely)" (*Metaphysik Herder*, 28.2:871). Kant had parlayed learning into social importance: *Bildung* made him a *Bürger*. Nevertheless, the scholarly estate had its demerits. *Gelehrsamkeit*, the entire *Hochaufklärung* contended, was not the same as *Bildung*, and sometimes it got in the way, both for the individual and for his community.¹³⁸

The Herder lectures formulate the issue in terms of the difference between "distraction" and "abstraction." The former was a widely shared human trait, and one which was, Kant averred in a variety of contexts, healthy. Abstraction, by contrast, was a trait that could be attained only by assiduous effort. "If I really think deeply, then I think little; I abstract with great effort, in order to make something clear" (28.2:871). At the very outset of his metaphysics lectures, as Kant introduced students to the field, he noted: "Nothing makes one more weary than the effort to understand correctly. The more removed [thought] gets from the sensual and the more preoccupied it is with [the abstract], the more difficult become its considerations" (28.1:6). Kant insisted that "a great vehemence (in studies, in intellectual activities) cannot be sustained: [only] a certain measure keeps up ones energies" (28.2:881). Interludes of relief proved indispensable. "One relaxes oneself not by not thinking but by thinking a lot: . . . Distraction is pleasant and necessary: thus high society distracts (when I am among thousands then I am really alone. Addison) then one is freer, more solitary, than in isolation" (28.2:871). Kant seemed to be affirming distraction, and *Umgang* as a vehicle for it, as the necessary antidote to scholarship. "The drive for action: are all the pleasures that it brings just to still the one drive toward science?" (28.2:893).

Science, Kant asserted unequivocally, was an unnatural pursuit. In the *Praktische Philosophie Herder*, he observed: "In general that is unnatural which is against the drives of nature. *Re*: the drive to science: it is not just something opposed to the self-preservation drive, but in particular [it is] against the sexual drive" (27.1:63).¹³⁹ He elaborated in *Bemerkungen* (34):

When something is not in keeping with the length of a [human] lifetime, nor with its epochs, nor with the larger part of mankind, but is ultimately a matter very much of chance and only possible with the greatest difficulty, it does not belong to the happiness and perfection of the human race. How many centuries went by before the first real sciences appeared and how many nations are there in the world which will never possess them.

From this global perspective, Kant turned to the level of the individual to continue his critique: "The poor accommodation of science with mankind is primarily to be seen in this: that the vast majority of those who wish to ornament themselves [with learning] achieve not a whit of improvement of the understanding but rather a perversion of the same, not to mention that for the majority science serves only as a mechanism for vanity" (*Bemerkungen*, 35).

In *Bemerkungen* Kant operated throughout with the Rousseauian distinction of natural man from the man of civilization. In that light, he argued that natural man had no need for science. "In natural circumstances one can be good without virtue and rational without science" (*Bemerkungen*, 14). But man was no longer merely natural. Though "[t]he drive to know does not lie in Nature; it has become indispensable for us now, . . . through long exercise" (27.1:63). Science—and, more generally, civilization—had become the human norm, however artificially established.

Animals [have] only capacities for life: instinctual organs, capacities in which there is nothing superfluous, but rather everything already in use in the present. . . . But the human spirit [has] a strong curiosity and passion which when it is developed make all the things of life trivial. Many more noble, great (useless) questions captivate him which have no use in this world. In these capacities the Creator seems to have handled even the wisest man with the greatest contempt (28.2:893).

The question is whether the pursuit of abstraction was not just strenuous but pathological. It is surprising how vehement Kant's tone got on this matter:

Scholars . . . have developed too many drives which could never be satisfied and become disturbed, become suicidal. . . . The scholar, gaunt, hypochondriac about useless things. . . . [M]ore and more useless drives are being awakened—what unhappiness faced with such a life! What a disproportion between the length of life and sciences (28.2:894).

Desire for things that we can never reach makes us crazy in this life. . . . No science in itself can make up for its own discontents; they become more and more harmful to themselves—useless to the public. . . . [S]cience terrifies, tears us away from pleasures, and the only satisfaction it offers is a means of prediction. Life is too short for this: there is just no proportion between them. Beavers do not build for the centuries. Man is insatiable in the pursuit of science and [then] he dies. His successor takes up the matter eagerly and he dies; everything gets interrupted. Newton died in the middle [breaks off] . . . (28.2:892).

Newton instead of being a human became an ape or an angel—lost enjoyment/peace—never satisfied himself—became childish—got laughed at, then died (28.2:894).

When we consider the reverence that Kant felt for Newton, these are stunning passages.

Kant argued that whatever did not answer to the actual purposes of nature destroyed its harmony, and he continued: “In the manner that one holds the arts and the sciences in such high regard, one makes contemptible those who are not in possession of them and [this] carries us into an injustice” (*Bemerkungen* (33–4). Thus “scholars think everything is made for their sake” (*Bemerkungen*, 34). This injustice that science incites was very much on Kant’s conscience. “It seems to us nowadays that the human race has virtually no worth apart from the great artists and scientists it boasts, and thus country folk and peasants appear to be worth nothing and of value only as a means to support the former. The injustice of this judgment is enough to prove its falsity” (*Bemerkungen*, 23–4). This was the insight for which Kant expressed his gratitude to Rousseau, for he had himself been guilty of this injustice. With Rousseau having “set him right,” Kant “learned to honor mankind and I would find myself more useless than the common laborer if I did not believe that my reflections could provide the rest [of mankind] with value, by formulating the rights of mankind” (*Bemerkungen*, 38).

The danger was that scholarship could interfere with moral capacities. "Speculations weaken the moral feeling" (28.2:895). Pedagogically, "one of the greatest harms caused by science is that it takes so much time that young people's morals get neglected" (*Bemerkungen*, 37). That made the concern for balance and moderation even more important. "Sciences . . . in all too vast a measure and development are harmful: hence what we need is an antidote: the human soul will live—according to the order of nature" (28.2:895). Yet Kant argued that one should not simply wish to revert to the natural condition. "Therefore, [pursue] science, etc., not as a blind thirst (consequently not to escape boredom, not to become unsociable, not to be contemptuous of the unlearned, but judging them generously), but with distance, as a purpose" (28.2:895).

A great deal of the language of this denunciation of science is clearly derived from Rousseau, but it would be unwise to presume that this was Kant's only source. Indeed, the whole array of moralists to whom Kant had turned his attention over the last years of the 1750s and the early 1760s, from Montaigne and Bacon through Shaftesbury and Hutcheson to Rousseau and Hume, to say nothing of the German tradition from Thomasius to Crusius, all provided him with insights on this score.¹⁴⁰ I would particularly wish to highlight the importance of Hume in this, for the efforts of Kant scholars to locate that famous but elusive moment when Hume awakened Kant from his "dogmatic slumbers" on the question of necessary causality incline them to minimize unduly the prominence of Hume as a cultural critic of scholarship in general and of academic philosophy in particular.¹⁴¹ Hume's impact on Kant, the "popular philosopher," was as seminal as Hume's later impact on Kant, the "transcendental philosopher."¹⁴² This is unequivocal in the case of Kant's concern with "national character."¹⁴³ But Kant hearkened even more to the fierce command Hume put into the mouth of Nature:

Indulge your passion for science, says she, but let your science be human, and such as may have a direct reference to action and society. Abstruse thought and profound researches I prohibit, and will severely punish, by the pensive melancholy which they introduce, by the endless uncertainty in which they involve you, and by the cold reception which your pretended discoveries shall meet with, when communicated. Be a philosopher, but amidst all your philosophy, be still a man.¹⁴⁴

Indeed, one may suspect that Kant was in closer sympathy on this matter with Hume than with Rousseau, for the latter's posture seemed to hold

out no place for science, whereas Hume at least considered a proper sort of science beneficial.¹⁴⁵ This, it turns out, was Kant's view as well. He stressed repeatedly the idea that while sciences did contribute to the corruption of civilized man, the advanced cultivation of these same sciences conduced to remedying their own evils (*Bemerkungen*, 35). He suggested that science could, along with the sense of looming death and of consolation in the next life, serve as an antidote to decadent inclinations (*Bemerkungen*, 18). If "effeminacy in mores, idleness and vanity produce the sciences[,] the latter give to the whole a new decor, keep away much that is bad and when they have achieved a certain level they ameliorate the evil they have themselves occasioned" (*Bemerkungen*, 37; see 80). What Kant had in mind in this last context comes clearer in another note: "Contestations in philosophy have the utility of fostering the freedom of the understanding and of arousing distrust of the very dogmatic edifice one proposed to build upon the ruins of another" (*Bemerkungen*, 112). Kant goes on—altogether in the spirit of Hume—to adopt a (moderate) skepticism in method: "the doubt of delaying. (Zetetics. . .) The method of doubt is useful in so far as it preserves the mind in acting not according to speculation but according to sound understanding and sentiment" (*Bemerkungen*, 130). Thus Kant redefines the project of philosophy: "One could say that metaphysics is the science of the limits of human reason" (*Bemerkungen*, 134). Hume saw his task in precisely that manner.¹⁴⁶

Kant, in what Richard Velkley has claimed is the decisive reorientation to the "primacy of practical reason," develops an argument in the *Bemerkungen* that science must always be appraised in terms of its human moral value.¹⁴⁷ "Every purpose of science is either erudition (Memory) or speculation (Reason). Both must lead to the result that man becomes more understanding (cleverer, wiser) in the world appropriate to human nature in general and thus more satisfied" (*Bemerkungen*, 10). That is, "if there is any science which man needs it is the one which teaches him how to take up appropriately the role in creation to which he has been assigned and through which he can learn what he must be in order to be a man" (*Bemerkungen*, 39). Elsewhere Kant notes: "the taste which is moral requires that a science which does not improve [mankind] be held in low esteem" (*Bemerkungen*, 11). Thus we see in Kant all the impulses of the critique of *Gelehrsamkeit* and the turn to the practical matter of "educating mankind" that are the hallmarks of *Popularphilosophie* and the *Hochaufklärung*. This is only reinforced when we turn to his dramatic change in life style.

The Russian occupation of Königsberg, which commenced in January 1758 and thus spanned most of the Seven Years' War, permanently altered some essential features of the city's social whirl.¹⁴⁸ Above all it transformed the relations among the elites of Königsberg society. Kurt Stavenhagen has written most vividly of this:

Above all, these [Russian] officers found totally incomprehensible the restrictions which kept the aristocracy and higher bureaucracy separated from the university faculty, and these in their turn from the commercial circles, and they did not feel the need to adapt to the prevailing etiquette of estates and cliques. . . . Thus arose for the first time that society which brought together all intellectual energies and soared beyond the lingering circles of estate life . . . This aristocratic society which could keep intellectual pace with its epoch is the Königsberg to which Kant, the elegant *Magister* with the sure presence of a man of the world, applied himself.¹⁴⁹

Königsberg did become a significant "center of the *Aufklärung*" over the balance of the eighteenth century, not least but also not only because of Kant.¹⁵⁰ The transformation that Stavenhagen portrays represented the social breakthrough of the Enlightenment, the creation of that eclectic social mix, the *gebildeten Stände*, out of which it emerged. The Russian occupation played a major role in cracking the frozen social relations of early modern life and opening up crucial interactions. Another vital contribution of the Russians was a revitalization of the "Dreikronenloge," the most important masonic lodge in Königsberg, making the city the center of a rich web of masonic interaction across the entire Baltic region and deep into Imperial Russia.¹⁵¹ While Kant held himself aloof from free-masonry, virtually everyone else in the Königsberg *Aufklärung* took part in it, including many of his intimate friends.¹⁵² It was also during the Russian occupation that the erratic but intrepid publisher and bookdealer Johann Kanter began to establish himself in Königsberg. He set about courting favor with the Prussian state in order, after the war, to secure a royal patent to found his own newspaper in Königsberg.¹⁵³ That newspaper, too, belongs in the institutional matrix of *Hochaufklärung*, and it proves a very important context for Kant, Herder, and the developments this study proposes to trace.¹⁵⁴

These were the years that Kant dramatically increased his socializing. Indeed, immediately after the Prussians resumed control over Königsberg, as Vorländer reports, Kant for the first time "entered the most distinguished

homes of the city, including especially the upper military."¹⁵⁵ Hamann wrote a very interesting letter in February 1764 about the Kant of that moment:

[Kant] is currently conducting a lecture course for General Meyer and his officers, which is bringing him great honor and profit, since almost every day he eats [with the General] and [is brought] there in a coach to give his lectures on mathematics and physical geography. Carried away in a whirl of social distractions, he has a number of projects in his head, *Morals*, *Attempt at a New Metaphysics*, an excerpt from his *Physical Geography*, and a number of small ideas from which I hope myself to profit. . . . If any of this comes to pass, which is still in doubt.¹⁵⁶

This Kant is quite a far remove from the one we have typically envisioned, far enough that even Hamann worried for his future productivity. All morning Kant devoted to his teaching, but the rest of the day was his. "After his lectures *Magister* Kant spent his time over a cup of coffee or a glass of wine; he played billiards and in the evening, cards. Often he made his way home after midnight, once—on his own admission—so drunk that he could not find his way to the street where he lived."¹⁵⁷

It is important that we take this intense social life seriously as a symptom of personal psychological restiveness. Kant needed to hob nob with the elite of the city. He needed to party with the military officers. He needed to seek out the society of ladies of high fashion. What was this need about?

"Socializing is the true root of life," Kant wrote, "and it makes a worthy man useful" (27:422). But *Umgang* was a trait Kant found particularly lacking among scholars. Wolfgang Ritzel paraphrases:

Many scholars are incapable of socializing and 'not conversational'; and yet just they ought to be 'able to converse with all stations [of society]' because they are 'beyond all their [particular] spheres . . . not so lowly before the grand, not too lofty for the commoner'. The true scholar is 'an academically trained person and thus best suited for conversation'. The unsociable nature of many scholars follows from a lack of 'worldly knowledge' and from a false estimation of science. 'Scholars think everything is arranged for their own pleasure'.¹⁵⁸

Vorländer observes much the same thing:

Thus it was not only the witty Hamann who mocked with regularity the petty rivalries and chicaneries as well as the arrogance of the Königsberg professorial circles, but there are also numerous statements in Kant's own writings regarding the presumptions of the scholars and their pedantry which can only be explained in terms of personal experience.¹⁵⁹

Ritzel, like Vorländer, registers Kant's sharp contempt for academics: that Kant "had something better in mind than the career of a school philosopher and scholar by occupation is certainly connected with his attitude towards the whole scholarly guild. . . . One does not have to search long to find further evidence of Kant's penchant for contempt, even malice."¹⁶⁰ Stavenhagen makes the same point: "in his circle of friends and acquaintances of the Sixties there was not one single scholar; instead every one was a representative of the *bios praktikos*."¹⁶¹ Weis corroborates the way Kant belittled "this estate of professional know-it-alls" he called the "scholar-mob": "it knows nothing, it understands nothing, but it talks about everything, and whatever it talks about it brags about."¹⁶² Vorländer writes:

[Kant] hated the pedantry of scholars just as much as he despised the cliquishness that so easily develops in closed occupational circles. He never joined over the course of his entire life any of the official scholarly associations, of which there were several in Königsberg, or the secret societies like the freemasons, although a number of his closest acquaintances belonged to them. He associated more with educated men and youths from different social backgrounds.¹⁶³

One of the most important such "youths from different social backgrounds" was a new student matriculated in the Albertina University in August 1762—Johann Herder—whom Kant "noticed immediately."¹⁶⁴ Kant was explicit in conveying to his favorite student that there was a necessary balance between study and society.

Just before Herder left Königsberg, Kant spoke with the nineteen-year-old and advised him that he should not pore so much over books, but rather far better follow his own example. He was very sociable and it was only in the world that one could cultivate himself. (In fact at that time *Magister* Kant was the most gallant man of the world, wore fancy clothes, a *postillon d'amour*, and visited all the coteries).¹⁶⁵

This is the passage which seems to have been the origin of the characterization of Kant as the "gallant *Magister*." Schöndörffer, having scoured the documentation and found none from the 1760s, chose instead the phrase "elegant *Magister*."¹⁶⁶ Hamann, and it would appear also students at the Albertina, called Kant the "little *Magister*," in token of his diminutive size, and perhaps also of his endearing youthfulness of personality.¹⁶⁷ On the other hand, this last formulation had an unfortunate linguistic connotation, for the phrase came very close to *petit maitre*, and this was a term Kant particularly despised and would have resented profoundly if ascribed to himself.¹⁶⁸ Even "gallant *Magister*" would have been problematic for him, for though he acknowledged and cultivated this style, he recognized it as well as a form of French rococo manners which he saw quite clearly as a feature of modern decadence (*Bemerkungen*, 100; see also 82).

Kant did become a favorite with the Prussian military officers of Königsberg over the balance of the 1760s. He was acquainted with several generals beside the General Meyer to whom Hamann referred above. Over the Autumn vacation of 1765 he accepted the invitation of General von Lossow to spend his time on the latter's estate on the East Prussian border with Russia. At the same time, Kant was a favorite of the Countess von Kayserling. It is striking that Kant got on so well with the higher military and aristocracy. Stavenhagen observes on that score: "Kant commanded the chivalrous forms of the cavalier of his time with mastery, because they fit so closely the chivalry of his own nature."¹⁶⁹ Hagiography, true; but there is something to this. Kant set himself very self-consciously to cultivate aristocratic manners; it was, in fact, part of the ambition of his class and even for the avant-garde of his profession. There was a deliberate effort to create a style at Göttingen that would make it a "university appropriate for cavaliers" and, indeed, Göttingen drew among the highest proportions of aristocrats into its student body achieved by the German universities of the eighteenth century.¹⁷⁰ Kant's dreary years of service as Hofmeister had not gone totally to waste. It was "during his time as a private tutor [that he] gradually picked up the fine social graces."¹⁷¹ In the 1760s Kant clearly possessed "worldly elegance and refined urbanity."¹⁷² Without them, the world of the officers and the aristocrats, the world of General von Lossow and Countess von Keyserling, would have been foreclosed to him.

But Kant was just as interested in the commercial elite of Königsberg. Most important in this connection was Kant's meeting, at least by 1766,

with the man who became his best friend, Joseph Green, a Scottish merchant settled in Königsberg. Green was Kant's paradigm of British national character, his informant on things British—and perhaps even on the tempestuous relations in those very years between Rousseau and Hume, two figures of ceaseless interest for Kant. Green was a man more methodical than Kant, and punctual enough to fulfill all the myths that have been saddled upon Kant himself. Indeed, Theodor von Hippel recognized him for an excellent subject and modeled his comic play, "A Man by the Clock," after him. Kant spent a great deal of time with Green, dining with him regularly. They talked about everything—even, Kant would have it, the details of composition of the *Critique of Pure Reason*.¹⁷³ There is little reason to suspect Green of metaphysical acumen, but every reason to see in him that *Weltkenntnis* which Kant prized perhaps more highly. Green's death in the mid-1780s left Kant genuinely bereft. Green was a bachelor, like Kant, but he left him some consolation. He had brought to Königsberg a young assistant and eventual partner, Robert Mothersby, who had the good sense to marry into one of the prominent commercial families of Königsberg. His family became a central part of Kant's later life, with the children of that household among his most cherished companions. Green and Mothersby were British. Johann Hamann, Theodor von Hippel and Johann Herder were avid students of British thought and letters. In a word, Kant in Königsberg was steeped in an Anglophile cultural milieu, and the affirmative image of British national character that fills his anthropology is not just something he derived from books.¹⁷⁴

Probably Kant's most important point of intellectual *Umgang* was in Kanter's bookstore, especially after 1766 when he took an apartment just upstairs. Here all the intellectually lively people in Königsberg gathered to check out the latest publications, read what they would not buy (Kant was prominent among those who did this), and talk about everything. It was a *Lesegesellschaft* without official organization. Hippel called it the "stock exchange of the scholarly world of the city."¹⁷⁵ Vorländer describes the reading room:

On a large table were laid out for all those interested to examine the newest releases from the book market. These, together with new political newspapers, stood ready for all visitors to read without charge, even for students twice a week. University professors and other scholars . . . met at this literary gathering place, in part to get informed, in part just to take a moment to enjoy the exciting entertainment.¹⁷⁶

Kanter had the painter Becker do portraits of the local literary lions as well as those from all over Prussia, e.g., Mendelssohn and Ramler, and these were hung on the various walls of the shop. Becker's portrait of Kant (1768) for this collection has been preserved. Since Kant lived just upstairs for eight of the years the bookshop played this role, 1766–74, he was one of its key attractions.

What our contextual analysis has brought into consideration, I propose, makes Kant's turn to a new style of expression and his pursuit of a new audience far more comprehensible. Kant set himself to be the *reflektierende Schriftsteller*, the *Beobachter*, called for by *Popularphilosophie*. The result was *Beobachtungen über das Gefühl des Schönen und Erhabenen*. It was also at this moment that Herder was studying with Kant, closer to him than any other observer, indeed modeling and projecting identities for his mentor as for himself in the broader campaign for enlightenment. For Herder, Kant promised to be the German Shaftesbury.¹⁷⁷ In that moment, deeply troubled about the direction of traditional German school metaphysics, as well as about the role of the *Gelehrtenstand* in the project of Enlightenment, Kant actively considered an alternative identity, that of the popular philosopher, and it was this project that Herder took up in his stead and in his image.

Notes

1. Lewis White Beck, *Early German Philosophy: Kant and His Predecessors* (1969; reprint: Bristol: Thoemmes, 1996), 426; henceforth *EGP*.

2. Heine, "History of Religion and Philosophy in Germany," in *The Romantic School and Other Essays* (New York: Continuum, 1985), 203. The issue about timelessness and contextualization is at the heart of the dispute within the discipline of philosophy over the place of its own history, or more precisely about how one should construe history from within the discipline of philosophy. See R. Rorty, J.B. Schneewind, and Q. Skinner, ed., *Philosophy in History* (Cambridge: Cambridge University Press, 1984).

3. There are numerous indications of interest in this domain. The title of a recent book by Hermann Schmitz posed the question, *Was wollte Kant?* (Bonn: Bouvier, 1989). In *Constituting Critique: Kant's Writing as Critical Praxis* (Durham: Duke University Press, 1994), Willi Goetschel has examined Kant's self-constitution as author, and in *The Embodiment of Reason* (Chicago: University of Chicago Press, 1996), Susan Shell has examined Kant's concern with his own physicality. A major contribution is James Schmidt, ed., *What is Enlighten-*

ment? *Eighteenth-Century Answers and Twentieth-Century Questions* (Berkeley: University of California Press, 1996).

4. I will express my great debt to Beck's erudition for orientation to problems in Kant scholarship in the time-honored form of challenging some of his interpretations.

5. Lewis White Beck, "A Prussian Hume and a Scottish Kant," in *Essays on Kant and Hume* (New Haven: Yale University Press, 1978), 111–29.

6. Giorgio Tonelli, "The 'Weakness' of Reason in the Age of Enlightenment," *Diderot Studies* 14:1971, 217–44, citing 241.

7. Beck, *EGP*, 426–9.

8. Rolf George recognizes the centrality of this question in "The Lives of Kant," *Philosophy and Phenomenological Research* 47:1987, 485–500, at 495.

9. Dieter Henrich, "Kants Denken 1762/3: Über den Ursprung der Unterscheidung analytischer und synthetischer Urteile," in *Studien zu Kants philosophischer Entwicklung* (Hildesheim: Olms, 1967), 9–38, citing 11.

10. Ernst Cassirer, *Kant's Life and Thought* (New Haven: Yale University Press, 1981), 77.

11. There is a mounting revisionism against this view, and my study proposes to add to this revisionism. See the penetrating analysis of Werner Schneiders, "Der Philosophiebegriff des philosophischen Zeitalters: Wandlungen im Selbstverständnis der Philosophie von Leibniz bis Kant," in *Wissenschaft im Zeitalter der Aufklärung*, ed. R. Vierhaus (Göttingen: Vandenhoeck and Ruprecht, 1985), 58–92; Helmut Holzhey, "Der Philosoph für die Welt—eine Chimäre der deutschen Aufklärung?" in *Esoterik und Exoterik der Philosophie: Beiträge zu Geschichte und Sinn philosophischer Selbstbestimmung*, ed. H. Holzhey and W. Zimmerli (Basel: Schwabe, 1977), 117–38; W. Zimmerli, "Arbeitsteilige Philosophie? Gedanken zur Teil-Rehabilitierung der *Popularphilosophie*," in *Wozu Philosophie? Stellungnahmen eines Arbeitskreises*, ed. H. Lübke (Berlin: Walter de Gruyter, 1978), 181–212; Zimmerli, "'Schulfuchsische' und 'handgreifliche' Rationalität—oder: Stehen dunkler Tiefsinn und Common sense im Widerspruch?" in *Wandel des Vernunftbegriffs*, ed. H. Poser (Freiburg: Albers, 1981), 137–76; Doris Bachmann-Medick, *Die ästhetische Ordnung des Handelns: Moralphilosophie und Ästhetik in der Popularphilosophie des 18. Jahrhunderts* (Stuttgart: Metzler, 1989).

12. Johan van der Zande, "In the Image of Cicero: German Philosophy between Wolff and Kant," *Journal of the History of Ideas* 56:1995, 419–42; J. van der Zande and R. Popkin, ed. *The Skeptical Tradition around 1800: Skepticism in Philosophy, Science and Society* (Dordrecht: Kluwer, 1998), 69–115.

13. Beck, *EGP*, 307, 320, and 323.

14. "Was ist Aufklärung?" 8:35.

15. Thomasius, *Einleitung zu der Vernunft-Lehre* (1691), cited in Holzhey, *op. cit.*, 122.

16. Bödeker, "Von der 'Magd der Theologie' zur 'Leitwissenschaft'," in *Popularphilosophie im 18. Jahrhundert: Das achtzehnte Jahrhundert* 14:1990, 19–57, citing 24.

17. Hammerstein, "Zur Geschichte der deutschen Universitäten im Zeitalter der Aufklärung," in *Universität und Gelehrtenstand 1400–1800*, ed. H. Rössler and G. Franz, (Limburg/Lahn: Starke, 1970), 145–82, citing 155.

18. Beck, *EGP*, 297.

19. Ibid., 298. See Budde, *Philosophischen Diskurs von dem Unterschied der Welt- und Schul-Gelehrtheit* (1709). On Budde's Thomasian eclecticism, see Schneiders, "Vernünftiger Zweifel und wahre Eklektik: Zur Entstehung des modernen Kritikbegriffs," *Studia Leibnitiana* 17:1985, 143–61, citing 154–5.

20. Eduard Zeller, "Wolffs Vertreibung aus Halle: der Kampf des Pietismus mit der Philosophie," *Preussische Jahrbücher* (1862) no. 2, 47–71; N. Hinske, ed., *Zentren der Aufklärung: I. Halle: Aufklärung und Pietismus* (Heidelberg: Lambert Schneider, 1989).

21. Norbert Hinske, "Wolffs Stellung in der deutschen Aufklärung," in *Christian Wolff 1679–1754*, ed. W. Schneiders (Hamburg: Meiner, 1983), 306–20.

22. H. de Vleeschauwer, "La genèse de la méthode mathématique de Wolff: Contribution à l'histoire des idées au XVIIIe siècle," *Revue Belge de Philologie et d'Histoire* 11:1932, 651–77.

23. Schneiders, "Philosophiebegriff," 69. Holzhey notes: "the concept of history encompassed . . . in addition to literary experience second hand also actual empiricism, sensual experience (*die eigentliche Empirie, die sinnliche Erfahrung*);" "Philosophie als Eklektik," *Studia Leibnitiana* 15:1983, 19–29, cited from 28.

24. See Michael Albrecht, "Kants Kritik der historischen Erkenntnis—ein Bekenntnis zu Wolff?" *Studia Leibnitiana* 14:1982, 1–24.

25. Bödeker, *op. cit.*, 32. Kant would raise all these issues himself at the end of the century in his *Streit der Fakultäten*. See esp. Günther Bien, "Kants Theorie der Universität und ihr geschichtlichen Ort," *Historische Zeitschrift* 219:1974, 551–77, esp. 556: Kant believed "the model of modern analytical natural sciences . . . directed all the modern disciplines, even those of practical philosophy." For evidence see the *Critique of Practical Reason*, 5:161–3.

26. He observes: "To be sure, Wolff demanded that the disciplines taught there [i.e., in the philosophical faculty] should be counted among the most important and leading ones in the canon of the sciences, that the philosophical faculty should be number one. But that's where matters stayed. Nothing changed"; Hammerstein, "Christian Wolff und die Universitäten," in *Christian Wolff 1679–1754*, 268.

27. Tonelli, "Der Streit über die mathematische Methode in der Philosophie in der ersten Hälfte des 18. Jahrhunderts und die Entstehung von Kants Schrift über die 'Deutlichkeit'," *Archiv für Philosophie* 9:1959, 37–66.

28. "The popular philosophy of Germany in the second half of the eighteenth

century emerged directly out of the decline of Wolffianism and the influx of English and French philosophy made easier [by that decline]"; Schneiders, "Zwischen Welt und Weisheit: Zur Verweltlichung der Philosophie in der frühen Moderne," *Studia Leibnitiana* 15:1983, 2–18, esp. 15.

29. Schneiders, "Vernünftiger Zweifel," 156; Holzhey, "Philosophie als Eklektik," 23. Even Kant saw himself as an "eclectic" in that moment, Holzhey suggests, referring to Kant's claim in the *Logic* (9:31) that an eclectic was an "independent thinker who acknowledged no allegiance to any school." One is reminded that Tonelli, in seeking to characterize the precritical Kant, could find no better description than: "anti-Woffian eclectic"; Tonelli, "Conditions in Königsberg and the Making of Kant's Philosophy," in *Bewußt sein: Gerhard Funke zu eigen*, ed. A. Bücher, H. Drüe, and T. Seebohm (Bonn: Bouvier, 1975), 126–44, citing 139. That translates, contextually, into "popular philosopher." But in *Reflexion* 1636 (16:61) Kant's reservations come to the fore: "The philosophers [*Weltweisen*] who philosophized free from the yoke of the schools created the eclectic approach. Until in the end beyond the schools and scholarly societies the *publicum* sets the tone, when a general shallowness gains the upper hand."

30. It was overthrown, in its turn, by Kant's new form of rigorous philosophy in 1783. That led to the third phase of popular philosophy, resistance to Kantianism throughout the *Spätaufklärung*. The total defeat of this popular philosophy has traditionally been linked to the triumph of German Idealism in the 1790s: yet another "school philosophy." But recently revisionism has been suggesting that the position taken up by Kant (and *a fortiori* the one taken up by the Idealists) was impossible to sustain and that a longer-term perspective suggests that at least elements of the position of popular philosophy proved stronger. This becomes more plausible if we see not Feder or even Garve as the principal exponent of this stance but rather Johann Gottfried Herder. Recent scholarship suggests that Herder may have been a far more weighty force in the late eighteenth century than a traditional Kantian conception of that period allows. Particularly crucial for current scholarship is Herder's role in late-eighteenth century "anthropology." See H.-J. Schings, ed., *Der ganze Mensch: Anthropologie und Literatur im 18. Jahrhundert* (Stuttgart: Metzler, 1994). Indeed, I contend that Herder gathered his vision of "anthropology" in the phase of popular philosophy associated with the *Hochaufklärung* in the 1760s and developed it over the balance of his career. His polemical confrontation with Kant, starting in the mid-1780s, led to his isolation from the German Idealists (not without having deeply influenced them despite themselves) and his reputation suffered deeply from this. See my essay, "Herder, Kant, Spinoza und die Ursprünge des deutschen Idealismus," in *Herder und die Philosophie des deutschen Idealismus*, ed. M. Heinz (Amsterdam: Rodopi, 1997 = *Fichte Studien, Supplementa*), 107–44.

31. In Diderot, *Pensées sur l'interprétation de la Nature* (1753). Ernesti's text, in Latin with a French translation, is reproduced in Roland Mortier, "Diderot,

Ernesti, et la 'Philosophie Populaire',," in *Essays on Diderot and the Enlightenment in Honor of Otis Fellows*, ed. J. Pappas (Geneva: Droz, 1974), 207–30. Mortier cites Diderot on 209.

32. Holzhey, "Der Philosoph für die Welt," 127.

33. Mortier, *op. cit.*, 228–9.

34. Schneiders, "Philosophiebegriff," 73.

35. Beck, *EGP*, 307.

36. "Characteristic of the development in Göttingen was the early and epochal differentiation of the traditional faculty of arts into individual sciences"; Bödeker, *op. cit.*, 42.

37. Iggers, "The University of Göttingen"; Peter Reill, "Die Geschichtswissenschaft um die Mitte des 18. Jahrhunderts," in *Wissenschaft im Zeitalter der Aufklärung*, 164–93; and more extensively, Reill, *The German Enlightenment and the Rise of Historicism* (Berkeley: University of California Press, 1975); Ulrich Muhlack, "Klassische Philologie zwischen Humanismus und Neuhumanismus," in *Wissenschaft im Zeitalter der Aufklärung*, 93–119; Timothy Lenoir, "The Göttingen School and the Development of Transcendental Naturphilosophie in the Romantic Era" (Baltimore: Johns Hopkins University Press, 1981 = *Studies on History of Biology* 5), 111–205; Frank Dougherty, "Nervenzmorphologie und-physiologie in den 80er Jahren des 18. Jahrhunderts: Göttinger Beiträge zur Forschung und Theorie der Neurologie in der vorgalvanischen Ära," in *Gehirn—Nerven—Seele*, ed. F. Dumont and G. Mann (Stuttgart: G. Fischer, 1988 = *Soemmerring-Forschungen*, Bd. 3), 55–91.

38. Michael Maurer, *Aufklärung und Anglophilie in Deutschland* (Göttingen: Vandenhoeck und Ruprecht, 1987); Manfred Kuehn, *Scottish Common Sense in Germany, 1768–1800: A Contribution to the History of Critical Philosophy* (Kingston: McGill-Queen's University Press, 1987); Fania Oz-Salzberger, *Translating the Enlightenment: Scottish Civic Discourse in Eighteenth-Century Germany* (Oxford: Clarendon, 1995).

39. The phrase "down to earth"—or more commonly, bringing matters back down to earth—is one of the most frequent and indicative tropes of cultural reorientation in this period. Schneiders notes it in "Zwischen Welt und Weisheit," 6. It plays a crucial role in Kant's *Dreams of a Spirit-Seer*, 2:368.

40. Walther Zimmerli, "'Schwere Rüstung' des Dogmatismus und 'anwendbare Eklektik': J.G.H. Feder und die Göttinger Philosophie im ausgehenden 18. Jahrhundert," *Studia Leibnitiana* 15:1983, 58–71.

41. See Günter Eberhard Schulz, "Christian Garve und Immanuel Kant: Gelehrten-Tugenden im 18. Jahrhundert," *Jahrbuch der Schlesischen Friedrich-Wilhelms-Universität zu Breslau* 5:1960, 123–88.

42. Meiners, *Revision der Philosophie* (anonymous; Göttingen und Gotha, 1772), 82, cited in Zimmerli, *op. cit.*, 69.

43. R. Steven Turner, "Historicism, *Kritik*, and the Prussian Professoriate, 1790–

1840," in *Philologie und Hermeneutik im 19. Jahrhundert II / Philologie et herméneutique au 19ème siècle II*, ed. M. Bollack, H. Wismann, and T. Lindken (Göttingen: Vandenhoeck and Ruprecht, 1983), 455. And: "one major theme of German social history in the eighteenth century was to be the precipitous dissolution of the *Allgemeine Gelehrtentum*"; *ibid.*, 453.

44. Schneiders, "Zwischen Welt und Weisheit," 5.

45. R.S. Turner, "University Reformers and Professorial Scholarship in Germany 1760–1806," in *The University in Society*, vol. II, ed. L. Stone (Princeton: Princeton University Press, 1974), 495–531, citing 501.

46. R.S. Turner, "The *Bildungsbürgertum* and the Learned Professions in Prussia, 1770–1830: The Origins of a Class," *Histoire Sociale/Social History* 13:1980, 108–35, cited from 111; Martens, "Von Thomasius bis Lichtenberg: Zur Gelehrten satire der Aufklärung," *Lessing Yearbook* 10:1978, 7–34.

47. Martens, "Gelehrten satire," 11, 28–9. See also Gunter Grimm, *Literatur und Gelehrtentum in Deutschland* (Tübingen: Niemeyer, 1983).

48. Thomas Nipperdey, "Verein als Sozialstruktur in Deutschland im späten 18. und frühen 19. Jahrhundert. Eine Fallstudie zur Modernisierung," in Nipperdey, *Gesellschaft, Kultur, Theorie* (Göttingen: Vandenhoeck and Ruprecht, 1976), 174–205, 439–47, here 180–1.

49. "In the second half of the eighteenth century the intellectual culture which sustained the old learned estate underwent a catalysmic dissolution before the forces of the Enlightenment"; Turner, "The *Bildungsbürgertum* and the Learned Professions," 110. "By 1790 the *Allgemeine Gelehrtentum* had become a *Stand* which had lost, not its function, but its exclusiveness and with it all social definition"; Turner, "Prussian Professoriate," 459.

50. Schneiders, "Zwischen Welt und Weisheit," 13.

51. Rolf Grimminger, "Aufklärung, Absolutismus und bürgerliche Individuen: Über den notwendigen Zusammenhang von Literatur, Gesellschaft und Staat in der Geschichte des 18. Jahrhundert," in Grimminger, ed., *Hanser Sozialgeschichte der deutschen Literatur vom 16. Jahrhundert bis zur Gegenwart. Bd. III: Deutsche Aufklärung bis zur Französischen Revolution, 1680–1789* (München: Hanser, 1980), 15–99.

52. Bödeker, *op. cit.*, 27.

53. Holzhey observes that "world" really meant the bourgeois reading public created and cultivated by the moral weeklies; "Der Philosoph für die Welt," 123.

54. W. Martens, *Botschaft der Tugend. Die Aufklärung im Spiegel der deutschen Moralischen Wochenschriften* (Stuttgart: Metzler, 1968); Pamela Currie, "Moral Weeklies and the Reading Public in Germany, 1711–1750," *Oxford German Studies* 3:1968, 69–86.

55. Rolf Engelsing, *Der Bürger als Leser: Lesergeschichte in Deutschland 1500–1800* (Stuttgart: Metzler, 1974); but see René König, "Geschichte und Sozialstruktur: Überlegungen bei Gelegenheit der Schriften von Rolf Engelsing

zur Lesergeschichte," *Internationales Archiv für Sozialgeschichte der deutschen Literatur* 2:1977, 134–43.

56. H.E. Bödeker, "Aufklärung als Kommunikationsprozeß," *Aufklärung* 2:1988, 89–111. For more skeptical views, see *Aufklärung und literarische Öffentlichkeit*, ed. C. Bürger, P. Bürger, and J. Schulte-Sasse (Frankfurt: Suhrkamp, 1980); Reinhart Meyer, "Limitierte Aufklärung: Untersuchungen zum bürgerlichen Kulturbewußtsein im ausgehenden 18. und beginnenden 19. Jahrhundert," in *Über den Prozeß der Aufklärung in Deutschland im 18. Jahrhundert*, ed. H.E. Bödeker and U. Herrmann (Göttingen: Vandenhoeck and Ruprecht, 1987), 139–200.

57. Paul Raabe, "Die Zeitschriften als Medium der Aufklärung," in *Wolfenbütteler Studien zur Aufklärung*, 1:1974, 99 ff.

58. Beck, *EGP*, 306–9; Horst Möller, *Aufklärung in Preussen: Der Verleger, Publizist und Geschichtsschreiber Friedrich Nicolai* (Berlin: Colloquium, 1974); Alexander Altmann, *Moses Mendelssohn: A Biographical Study* (London: Routledge and Kegan Paul, 1973), and Altmann, *Die trostvolle Aufklärung: Studien zur Metaphysik und politische Theorie Moses Mendelssohns* (Stuttgart: Frommann-Holzboog, 1982); W. Forster, ed., *Aufklärung in Berlin* (Berlin: Akademie, 1989); M. Fontius and H. Holzhey, ed., *Schweizer im Berlin des 18. Jahrhunderts* (Berlin: Akademie, 1996).

59. Rudolf Vierhaus, "Die aufgeklärten Schriftsteller: Zur sozialen Charakteristik einer selbsternannten Elite," in *Über den Prozess der Aufklärung im 18. Jahrhundert*, 53–65; Herbert Göpfert, "Schriftsteller," in Göpfert, *Vom Autor zum Leser: Beiträge zur Geschichte des Buchwesens* (München: Hanser, 1977), 143–54.

60. Dominik von König, "Lesesucht und Lesewut," in *Buch und Leser*, H. Göpfert, ed. (Hamburg: Hauswedell, 1977), 89–112; Erich Schön, *Der Verlust der Sinnlichkeit oder Die Verwandlung des Lesers: Mentalitätswandel um 1800* (Stuttgart: Klett-Cotta, 1987).

61. "Reading [was a] medium of self-discovery and self-comprehension of the educated" but also "they wished to learn how to understand and interpret their world"; H.E. Bödeker, "Die 'gebildeten Stände' im späten 18. und frühen 19. Jahrhundert: Zugehörigkeit und Abgrenzungen. Mentalitäten und Handlungspotentiale," in *Bildungsbürgertum im 19. Jahrhundert. Teil IV*, ed. J. Kocka (Stuttgart: Klett-Cotta, 1992), 37.

62. It is no coincidence these are Kantian categories. See *Ekλεκtik, Selbstdenken, Mündigkeit*, ed. N. Hinske (Hamburg: Meiner, 1986 = *Aufklärung* 1:1986); Hinske, "Die tragende Grundideen der deutschen Aufklärung: Versuch einer Typologie," in *Die Philosophie der deutschen Aufklärung: Texte und Darstellung*, ed. R. Ciafardone (Stuttgart: Reclam, 1990), 407–58; Hinkse, *Zwischen Aufklärung und Vernunftkritik: Studien zum Kantschen Logikcorpus* (Stuttgart: Frommann-Holzboog, 1998), 17–90.

63. Sergio Moravia, *Beobachtende Vernunft: Philosophie und Anthropologie in der Aufklärung* (München: Hanser, 1973).

64. Monika Linden, *Untersuchungen zum Anthropologiebegriff des 18. Jahrhunderts* (Bern: Lang, 1976); Andreas Käuser, "Anthropologie und Ästhetik im 18. Jahrhundert," *Das achtzehnte Jahrhundert* 14: 1990, 196–206; J. Barkhoff and E. Sagarra, ed., *Anthropologie und Literatur um 1800* (München: Iudicium Verlag, 1992); H-J. Schings, ed., *Der ganze Mensch: Anthropologie und Literatur im 18. Jahrhundert* (Stuttgart: Metzler, 1994); K. Faull, ed., *Anthropology and the German Enlightenment: Perspectives on Humanity* (Lewisburg: Bucknell University Press, 1995).

65. Wolfram Mauser, "Geselligkeit: Zu Chance und Scheitern einer sozialetischen Utopie um 1750," in *Entwicklungsschwellen im 18. Jahrhundert*, ed. K. Eibl (Hamburg: Meiner, 1989 = *Aufklärung* 4:1989), 5–36.

66. Otto Dann, "Die Lesegesellschaften des 18. Jahrhunderts und der gesellschaftliche Aufbruch des deutschen Bürgertums," in *Buch und Leser*, ed. Göpfert, 160–93; O. Dann, ed., *Lesegesellschaften und bürgerliche Emanzipation* (München: Beck, 1981).

67. H.E. Bödeker, "Journals and Public Opinion: The Politicization of the German Enlightenment in the Second Half of the Eighteenth Century," in *The Transformation of Political Culture: England and Germany in the Late Eighteenth Century*, ed. E. Hellmuth (The German Historical Institute London/ Oxford University Press, 1990), 423–45, here 434.

68. "The constitution of associations was at one and the same time the highpoint of the emancipation of the 'educated estates' from the estate tradition and also its self-constitution as a separate segment of society"; Bödeker, "Gebildeten Stände," 41. And see *ibid.*, 43, on the modeling of democratic practices such association made possible.

69. H.E. Bödeker, "Prozesse und Strukturen politischer Bewußtseins *Bildung* der deutschen Aufklärung," in *Aufklärung als Politisierung—Politisierung der Aufklärung*, ed. H.E. Bödeker and U. Herrmann (Hamburg: Meiner, 1984), 10–31; O. Dann, "Die Anfänge politischer Vereinsbildung in Deutschland," in *Soziale Bewegung und politische Verfassung: Beiträge zur Geschichte der modernen Welt*, ed. U. Engelhardt, V. Sellin, and H. Stuke (Stuttgart: Klett, 1976), 197–232; T. Nipperdey, "Verein als soziale Struktur."

70. R. Vierhaus, "'Patriotismus'—Begriff und Realität einer moralisch-politischen Haltung," in *Deutsche patriotische und gemeinnützige Gesellschaften*, ed. Vierhaus (München: Kraus, 1980), 9–29.

71. Mendelssohn, Letter 138, *Briefe, die neueste Literatur betreffend* (1761), 214, cited by Ulrich Im Hof, "German Associations and Politics in the Second Half of the Eighteenth Century," in *The Transformation of Political Culture*, 207–18, at 208 and note.

72. See U. Im Hof, "Die Helvetische Gesellschaft 1761–1768," in *Deutsche*

patriotische und gemeinnützige Gesellschaften, ed. Vierhaus, 223–40; Im Hof, "German Associations and Politics in the Second Half of the Eighteenth Century," esp. 207–8.

73. Beck, *EGP*, 320.

74. M. Welke, "Die Legende vom 'unpolitischen Deutschen': Zeitungslesen im 18. Jahrhundert als Spiegel des politischen Interesses," *Jahrbuch der Wittheit zu Bremen* 25:1981, 161 ff. And see R. Vierhaus, "The Revolutionizing of Consciousness: A German Utopia?" in *The Transformation of Political Culture*, 561–77.

75. R. Vierhaus, "'Patriotismus'—Begriff und Realität einer moralisch-politischen Haltung," 10. Such a conception of political culture and political process is the key to the new approach to German social history and to the political history of Aufklärung. Its triumph can be dated from two collections edited by U. Herrmann and H. E. Bödeker, *Aufklärung als Politisierung* (1987) and *Über den Prozeß der deutschen Aufklärung* (1987), which confirmed the connection of shifts in personal consciousness and social interaction with political culture in the German eighteenth century.

76. Beck, *EGP*, 315.

77. Harcourt Brown, "Maupertuis *philosophe*: Enlightenment and the Berlin Academy," *Studies on Voltaire and the eighteenth century* 24:1963, 255–69; Ronald Calinger, "Frederick the Great and the Berlin Academy of Sciences (1740–1766)," *Annals of Science* 24:1968, 239–49; Cornelia Buschmann, "Philosophische Preisfragen und Preisschriften der Berliner Akademie 1747–1768: Ein Beitrag zur Leibniz-Rezeption im 18. Jahrhundert," *Deutsche Zeitschrift für Philosophie* 29:1987, 779–89; Hans Aarsleff, "The Berlin Academy under Frederick the Great," *History of the Human Sciences* 2:1989, 193–206; Mary Terrall, "The Culture of Science in Frederick the Great's Berlin," *History of Science* 28:1990, 333–64; John Christian Laursen, "The Berlin Academy," *Columbia History of Western Philosophy*, ed. R. Popkin (New York: Columbia University Press, 1999), 490–4.

78. Robert Hering, "Johann Georg Sulzer: Persönliches und Literarisches zur 150. Wiederkehr seines Geburtstages," *Jahrbuch der Freien Deutschen Hochstifts* 1928, 265–325; J. van der Zande, "Orpheus in Berlin: A Reappraisal of Johann Georg Sulzer's Theory of the Polite Arts," *Central European History* 28:1995, 175–208; H. Holzhey, "Die Berliner Popularphilosophie: Mendelssohn und Sulzer über die Unsterblichkeit der Seele," in *Schweizer im Berlin des 18. Jahrhunderts*, ed. M. Fontius and H. Holzhey, 201–16; H.E. Bödeker, "Konzept und Klassifikation der Wissenschaften bei Johann Georg Sulzer (1720–1779), *ibid.*, 325–40.

79. Sulzer, *Kurzer Begriff der Wissenschaften*, cited in Holzhey, "Philosoph für die Welt," 133.

80. Holzhey, *ibid.*, 121.

81. J.D. Michaelis, *Raisonnement über die protestantischen Universitäten Deutschlands* (Frankfurt/Leipzig, 1768), 1:112, cited in Hammerstein,

"Universitäten und gelehrte Institutionen," 311. Schneiders confirms Michaelis' judgment: "The German Enlightenment [figures] belonged generally neither to the grand bourgeoisie nor to the literary aristocracy [French categories]; they were small bureaucrats, preachers and teachers, above all university professors"; Schneiders, "Akademische Weltweisheit," 40.

82. Gerd Ueding, "Popularphilosophie," in *Deutsche Aufklärung bis zur Französischen Revolution, 1680–1789*, ed. R. Grimminger, 605–34, citing from 609.

83. Turner, "Discussion of 'Prussian Professoriate'," 483.

84. Vorländer, *Kant: der Mann und das Werk* (Hamburg: Meiner, 1977), 158.

85. *Ibid.*, 117.

86. Gulyga, *Immanuel Kant: His Life and Thought* (Stuttgart: Birkhäuser, 1987), 40–1.

87. R. George, *op. cit.*, 496.

88. Richard Velkley, *Freedom and the End of Reason: On the Moral Foundation of Kant's Critical Philosophy* (Chicago: University of Chicago Press, 1989); Frederick Beiser, "Kant's Intellectual Development, 1746–1781," in *The Cambridge Companion to Kant*, ed. P. Guyer (Cambridge: Cambridge University Press, 1992), 26–61.

89. Goethe, for example, writing in the early 1770s, listed Kant among the most prominent "popular philosophers" of Germany. It was *Beobachtungen* that gave Kant that reputation.

90. Kant, *Reflexionen* 1497, 1496, 373, 404, 414 (15:769, 712, 148, 163, 167), and *Anthropology*, 7:294. See Wolfgang Ritzel, *Immanuel Kant: Eine Biographie* (Berlin: Walter de Gruyter, 1985), 125.

91. On the question see Bernhard Fabian, "Der Gelehrte als Leser," *Librarium: Zeitschrift der schweizerischen Bibliophilen-Gesellschaft* (Zurich) 19:1976, 160–7.

92. In her introduction and notes to Immanuel Kant, *Bemerkungen in den 'Beobachtungen über das Gefühl des Schönen und Erhabenen'*, ed. M. Rischmüller (Hamburg: Meiner, 1991), Marie Rischmüller discusses this in useful detail; henceforth this work is referred to as *Bemerkungen*.

93. Gulyga, *op. cit.*, 51–2.

94. H. Fischer reported the fifteen reviews; M. Rischmüller, ed., *Bemerkungen*, xvi. Information about the many reprints of this work is to be found in the editorial annotations to the *Akademie* edition.

95. Prussian Ministry of Culture, Directive, August 5, 1765; cited in the *Akademie* edition, 13:24.

96. Braxein to Prussian Ministry, October 19, 1765; cited in 13:24.

97. Ritzel explains Kant's attitude aptly: "Kant knew well that he was not the man to provide ceremonial verse in German and Latin at University occasions or jubilees of the ruling house"; *op. cit.*, 51.

98. Braxein to Prussian Ministry, October 19, 1765; cited in 13:24. Braxein

was speculating that the health of the Ordinary Professor for Logic and Metaphysics might be failing. The awareness of that person’s delicate health is a background factor in Kant’s behavior over the balance of the 1760s.

99. Prussian Ministry, in the name of the King, October 24, 1765; cited in 13:25.

100. Prussian Ministry, in the name of the King, October 28, 1765; cited in 13:25.

101. Kant enclosed a copy of this notification in his application letter for the library position, and it is reprinted in his correspondence as an enclosure to that letter; see 10:49.

102. Josef Heller, *Kants Persönlichkeit und Leben: Versuch einer Charakteristik* (Berlin: Hesse, 1924), 28. It is noteworthy that Kant later conceived economic independence as the indispensable requirement for eligibility for citizenship and political participation. It could be argued that Kant went so far as to consider dependence—broadly construed as *Unmündigkeit*—as a bar to full ethical personhood.

103. Carl Joseph Maria Dénina, *La Prusse littéraire sous Frédéric II*, vol. 2 (Berlin 1790), 305–8.

104. Kant to Lagarde, March 25, 1790 (11:146).

105. Heller, *op. cit.*, 45.

106. Stuckenberg, *The Life of Immanuel Kant* (1882; reprint: Lanham: University Press of America, 1986), 67.

107. Jachman, “Immanuel Kant geschildert in Briefen an einem Freund,” in *Immanuel Kant: Sein Leben in Darstellungen von Zeitgenossen: Die Biographien von L.E. Borowski, R.B. Jachmann und A.C. Wasianski*, ed. F. Groß (Berlin: Deutsche Bibliothek, 1912), 126.

108. Gulyga, *op. cit.*, 24.

109. Jachmann, *op. cit.*, 134.

110. Ritzel, *op. cit.*, 15.

111. Arnoldt, “Kants Jugend und die fünf ersten Jahre seiner Privatdozentur,” *Gesammelte Schriften* III (Berlin: Cassirer, 1908), 103–210.

112. Gulyga, *op. cit.*, 24.

113. Vorländer, *op. cit.*, vol. 2, 8; Kant to Borowski, March 6, 1761 (10:34).

114. Kant to Frederick II, October 24, 1765 (10:49).

115. Kant to Freiherr von Furst und Kupferberg, October 29, 1765 (10:50).

116. “His first appointment to a position with a salary was in 1766, when he became second librarian in the Royal Library. The Government in Berlin, in the letter appointing him, designated him as ‘the able Kant, made famous by his books.’ His salary was 62 taler a year”; Stuckenberg, *op. cit.*, 83.

117. *Ibid.*, 88.

118. At the founding of the University of Halle (1694), the salaries established for the faculties were as follows: 500 florins for the Ordinary Professor of Theol-

ogy (and, exceptionally, for the very distinguished Christian Thomasius in the faculty of Jurisprudence), compared with 100 to 150 florins for the arts faculty. For the University of Jena in 1704, comparable annual salary levels were: 366 florins for theology; between 200 and 300 for jurisprudence, and between 130 and 210 florins for the arts faculty. These, Hammerstein notes, were "typical for the time"; Notker Hammerstein, "Die Universitätsgründungen im Zeichen der Aufklärung," in *Beiträge zu Problemen deutscher Universitätsgründungen der frühen Neuzeit*, ed. P. Bamgart and N. Hammerstein (Nendeln/Liechtenstein: KTO, 1978), 263–98, cited from 267 and note. Elsewhere he argues that it is difficult to generalize about salary levels for the eighteenth century but he notes that one could establish extremes. The highest salary went to jurisprudence, 1500 taler, and the lowest to the arts faculty, between 150 and 300 taler. Hammerstein adds: "in this period one would estimate the minimum for existence around 4–500 taler"; "Universitäten und gelehrte Institutionen von der Aufklärung zum Neuhumanismus und Idealismus," in *Samuel Thomas Soemmerring und die Gelehrten der Goethezeit*, ed. G. Mann and F. Dumont (Stuttgart: G. Fischer, 1985 = *Soemmerring-Forschungen*, vol. 1), 309–29, citing 317 n. Turner notes that around mid-century the Prussian University of Frankfurt am Oder had a salary scale as follows: theologians, 557–338 taler; jurists, 500–200 taler; physicians, 300–100 taler; and philosophical faculty, 175–100 taler; "University Reformers," 499.

119. Turner notes that the jurisprudence faculty at the University of Königsberg automatically became magistrates of the city government upon appointment to the university; *ibid.*, 506.

120. Hammerstein, "Universitäten und gelehrten Institutionen," 317.

121. Turner, "University Reformers," 507.

122. Hammerstein notes that the founding salaries at Göttingen in 1734–7 were substantially better, especially at the level of the philosophical faculty. Also, here the preponderance of the faculty of jurisprudence over that of theology made itself felt: salaries in jurisprudence ranged from 300 to 1000 taler; in theology from 300 to 600, in medicine from 300 to 500, and in the philosophical faculty from 300 to 400; "Universitätsgründungen," 276.

123. G.A. Kelley, "Rousseau, Kant and the Idea of History," *Journal of the History of Ideas* 29:1968, 347–64, citing 364.

124. Goetschel, *Constituting Critique*, 72.

125. The German reception registered Rousseau ubiquitously as "paradoxical." Kant himself would criticize him for this in the *Bemerkungen*, 38.

126. My assessment of Kant's Rousseau reception tallies with that developed by Goetschel, *op. cit.*, esp. 61 and 201 n. 13.

127. Two major contributions are: Josef Schmucker, *Die Ursprünge der Ethik Kants in seine vorkritischen Schriften und Reflexionen* (Meisenheim am Glan: Hain, 1961); Jean Ferrari, *Les source françaises de la philosophie de Kant* (Paris: Klingensieck, 1979), 171–290. See also R. Velkley, *op. cit.*

128. Hamann to Kant, July 27, 1759 (10:7–16).

129. Kant, "Spitzfindigkeit" (2:57). Goetschel makes the same argument (*op. cit.*, 44): "the first marks of Rousseau already appear in the *False Subtlety*. Its title attacks scholastic philosophy just as sharply as Rousseau's own antitheses do, although Rousseau does so more completely and with more virtuosity. Kant echoes Rousseau's acrid sarcasm, which is so often hostile to science." And see *ibid.*, 201 n. 13.

130. Wilhelm Dobbek, *Johann Gottfried Herders Jugendzeit in Mohrungen und Königsberg 1744–1764* (Würzburg: Holzner, 1961), 95.

131. Not only does Kant refer to the novel in his *Bemerkungen*, 91, but there are references to the novel also in the Herder Lecture notes (27:1:18). That Kant read novels, and that they played a significant role in his identity construction, is something that deserves to be underscored, as M. Rischmüller has done in her annotations to the *Bemerkungen*, 146, and *passim*. In this, Kant was a person of his epoch. Novel-reading was crucial in the period just after the middle of the eighteenth century for creating ideologies of "human rights" grounded in a sense of individual identity. All this bears directly on Kant's reception of Rousseau and also on his involvement more generally in the Enlightenment and in "bourgeois emancipation" in Germany.

132. The legend of Kant so embroiled in reading *Emile* that he missed his walk, and thus deprived the housewives of Königsberg of an essential time-check—propounded originally by Borowski (*op. cit.*)—needs to be taken with a grain of salt, since it imputes to the Kant of the early 1760s the methodical ways of the older Kant. Borowski cannot count as a first-hand witness, since he had already left Königsberg in 1762. In fact, he was writing this at a distance of many decades from the event. Yet there is no question that, as J. Schmucker has written (*op. cit.*, 174), no book ever made more of an impact on Kant than the *Emile*.

133. Anna Attridge, "The Reception of *La Nouvelle Héloïse*," *Studies on Voltaire and the eighteenth century* 120:1974, 227–67.

134. Raymond Trousson, "Jean Jacques Rousseau dans la presse périodique allemande de 1750 a 1800 (II)," *Dix-huitième siècle* 2:1970, 233.

135. Contrast this with the two philosophers whom ostensibly he most admired in Germany, Mendelssohn and Lambert, who both took distinctly reserved stances towards Rousseau.

136. The note begins to offer a third thought, but breaks off—leaving us to wonder what this third, "hard to come by and rare" insight into Rousseau might have been. The entire *Bemerkungen*, I suspect, represent Kant's effort to "work through" that third stage.

137. Schmucker, *op. cit.*, 141; for Henrich on the "overwhelming impact" of Rousseau on Kant's philosophical development, see his "Über Kants Entwicklungsgeschichte," 257.

138. Schmucker tends to trace all of Kant's disaffection from *Gelehrsamkeit*

to Rousseau, but our overview of the *Hochaufklärung* in Germany has demonstrated that this was a pervasive motif tracing back at least as far as to Christian Thomasius.

139. This suggestion that science undermines the sexual drive may well account for an otherwise unexplained juxtaposition, in *Bemerkungen*, 76: "On the sexual inclination. On science."

140. Wilhelm Jaitner, *Thomasius, Rüdiger, Hoffmann und Crusius: Studien zur Menschenkunde und Theorie der Lebensführung im 18. Jahrhundert* (Dissertation: Cologne; Bleichrode: Niefert, 1939).

141. Bruno Erdmann, "Kant und Hume um 1762," *Archiv für Geschichte der Philosophie* 1:1888, 62–77, 216–230; Robert P. Wolff, "Kant's Debt to Hume via Beattie," *Journal of the History of Ideas* 21:1960, 117–23; L.W. Beck, *Essays on Kant and Hume*; Manfred Kuehn, "Kant's Conception of Hume's Problem," *Journal of the History of Philosophy* 21:1983, 175–93; Lothar Kreimendahl, *Kant—Der Durchbruch von 1769* (Cologne: Dinter, 1990).

142. Günter Gawlick and Lothar Kreimendahl, *Hume in der deutschen Aufklärung: Umriss einer Rezeptionsgeschichte* (Stuttgart: Frommann-Holzboog, 1987); Reinhard Brandt and Heiner Klemme, *David Hume in Deutschland: Literatur zur Hume-Rezeption im Marburger Bibliothek* (Marburg: Universitätsbibliothek, 1988); Brandt, "Die englische Philosophie als Ferment der kontinentalen Aufklärung," in *Europäische Aufklärung(en): Einheit und nationale Vielfalt*, ed. S. Jüttner and J. Schlobach (Hamburg: Meiner, 1992), 66–79; M. Kuehn, "Hume in the *Göttingische Anzeigen*: 1739–1800," *Hume Studies* 13:1987, 46–73; Shirley Robin Letwin, "Hume: Inventor of a New Task for Philosophy," *Political Theory* 3:1975, 134–58.

143. Compare Hume's essay by that title with the fourth segment of Kant's *Beobachtungen*.

144. Hume, *An Enquiry Concerning Human Understanding* (Indianapolis: Hackett, 1977), 3–4.

145. Susan Shell writes: "Kant is always far more confident than Rousseau of the potential benefits of science"; *Rights of Reason* (Toronto: University of Toronto Press, 1980), 22.

146. "Here indeed lies the justest and most plausible objection against a considerable part of metaphysics, that they are not properly a science; but arise either from fruitless efforts of human vanity, which would penetrate into subjects utterly inaccessible to the understanding, or from the craft of popular superstitions. . . . The only method of freeing learning, at once, from these abstruse questions, is to enquire seriously into the nature of human understanding, and show, from an exact analysis of its powers and capacity, that it is by no means fitted for such remote and abstruse subjects"; Hume, *op. cit.*, 5–6.

147. R. Velkley, "Freedom, Teleology, and Justification of Reason: On the Philosophical Importance of Kant's Rousseauian Turn," in *Rousseau in Deutschland*, ed. H. Jaumann (Berlin: Walter de Gruyter, 1995), 181–96.

148. Gause, the major historian of the city, writes that the Russians did not leave Königsberg until March 1763; *Kant und Königsberg* (Leer/Ostfriesland: Rautenberg, 1974), 52.

149. Kurt Stavenhagen, *Kant und Königsberg* (Göttingen: Deuerlich, 1949), 21–4.

150. J. Kohnen, ed., *Königsberg: Beiträge zu einem besonderen Kapitel der deutschen Geistesgeschichte des 18. Jahrhunderts* (Frankfurt,: Peter Lang, 1994); H. Ischreyt, ed., *Königsberg und Riga* (Tübingen: Niemeyer, 1995 = *Zentren der Aufklärung*), 2.

151. J. Kohnen, "Druckerei-, Verlags- und Zeitungswesen in Königsberg zur Zeit Kants und Hamanns. Das Unternehmen Johann Jakob Kanters," in Kohnen, ed., *Königsberg*, 2. And see H. Ischreyt, "Die Königsberger Freimaurerloge und die Anfänge des modernen Verlagswesens in Rußland (1760–1763)," in *Rußland und Deutschland* (Stuttgart: Kieler Historische Studien, 1974), and Ischreyt, "Streiflichter über die Freimaurerei in Kurland," in *Beförderer der Aufklärung in Mittel- und Osteuropa: Freimaurer, Gesellschaften, Clubs*, ed. E. Balazs, L. Hammermayer, H. Wagner and J. Wojtowicz (Berlin: Camen, 1979).

152. Gause, *op. cit.*, 123.

153. Kohnen, "Unternehmen," 8.

154. Angelo Pupi, "Die Anfänge der Königsbergischen Gelehrten und Politischen Zeitungen (Februar-Mai 1764)," in *Königsberg*, ed. Kohnen, 21–54.

155. Vorländer, *op. cit.*, 118.

156. Hamann to Lindner, February 1, 1764, *Hamanns Briefwechsel*, vol. 2 (Wiesbaden: Insel, 1956), 234.

157. Gulyga, *op. cit.*, 51–2.

158. Ritzel, *op. cit.*, 75, citing Kant, *Praktische Philosophie Herder*, 27.1:81.

159. Vorländer, *op. cit.*, 85.

160. Ritzel, *op. cit.*, 51.

161. Stavenhagen, *op. cit.*, 48.

162. Norbert Weis, *Königsberg: Immanuel Kant und seine Stadt* (Braunschweig: Westermann, 1993), 59–60.

163. Vorländer, *op. cit.*, 138.

164. Gulyga, *op. cit.*, 48.

165. Karl August Boettiger, *Literarische Zustände und Zeigenossen* (Leipzig, 1834), vol. I, 133; cited in N. Hinske, "Kants Idee der Anthropologie," in *Die Frage nach dem Menschen*, ed. H. Rombach (München: Alber, 1966), 423 n.

166. Schöndörffer, "Der elegante Magister," *Reichs Philosophische Almanach auf das Jahr 1924. Immanuel Kant zum Gedächtnis* (Darmstadt, 1924), 65–86, citing 69 n.

167. *Ibid.*

168. Kant, *Bemerkungen*, 7, 127. The editor of Kant's text, Marie Rischmüller, explains the origins and significance of the term in her annotations, 137–8.

169. Stavenhagen, *op. cit.*, 58.

170. Hammerstein, "Universitätsgründungen," 279.

171. Schöndörffer, *op. cit.*, 70.

172. *Ibid.*, 78.

173. Jachman, *op. cit.*, 152–5.

174. "Particularly important [in connection with Königsberg] is the close connection with England and the reception of the modern ideas of English philosophy (empiricism, sensualism, skepticism) as well as literature which it mediated, not to be forgotten the more liberal civic and economic ideas"; Margot Westlinning, "Der junge Herder in Königsberg," in *Königsberg und Riga*, ed. Ischreyt, 73.

175. Cited in Stavenhagen, *op. cit.*, 46.

176. Vorländer, *op. cit.*, 182.

177. Wilhelm Dobbek, *op cit.*, 113.

Index

- abstraction, 35, 102–4, 407, 408
action, 202, 207, 212, 213–14, 216,
217–20, 235–6, 237–8, 239,
241, 245, 248, 262, 287, 302,
305, 348, 349
free, 197–9, 200–5, 207, 209, 210,
212, 213, 215–19, 237
rational, 197–8, 200–1, 205, 215
See also agency, freedom,
rationality
affection, 183–4, 187–9
agency, 7, 198–9, 202–3, 208–9, 211,
305, 306
free, 198–9, 214, 219
rational, 86, 207, 213
See also action, freedom,
rationality
Allison, Henry, 7–8, 15, 168, 202,
203, 205, 207, 210, 211–12,
213–15, 217, 219, 220
Analogy, Second, 199, 202–4, 205–6,
208, 210, 211, 212–13, 214–16,
219, 220
analyticity, 93–102, 106–9
animals, 14, 36–7, 85, 269–71, 408
Anscombe, Elizabeth, 235, 236
antinomy, 27, 210–11, 215
appearance, 30–1, 57, 120, 124–5,
184
See also object, representation
apperception, 8, 19, 121, 183, 198,
200, 201, 202, 212, 214, 215,
216, 217, 218
unity of, 19, 120
See also awareness, consciousness
Aquinas, Thomas, 287, 321
Aristotle, 5, 6, 14, 18, 19, 102, 153,
164, 245, 246, 247, 250, 251,
252, 255, 261, 262, 284, 287
Aufklärung. *See* Enlightenment
Augustine, 153, 284, 287, 290
autonomy, 4, 289, 294, 296, 298–9,
303–4, 316, 323, 374, 383
and heteronomy, 198–9, 200, 298–
9
awareness, 117–19, 121–6, 128–34,
139
See also apperception,
consciousness
Bacon, Francis, 319, 410
beauty, 327, 369, 371
Beck, Lewis White, xvii–xviii, xxv,
36, 41, 59, 71, 88, 117, 176,
197, 202, 211–12, 214–15, 219,
250, 283, 333, 387, 388, 389,
390, 392
conception of man, xvii, xxiii–xxv
conception of freedom, 197, 210–
13, 217
on Kant's strategy, 25–8, 33
beings,
human, 269–70, 271–2, 297–8,
302, 319
rational, 206, 270–1, 276–7, 297–
8, 304, 325
See also man, nature, rationality
Bentham, J., 233, 261

- Bible, 337–8, 345–6, 348, 354, 355–5
See also Christianity, faith, God, religion
- Book of Job, 327, 328
- Bridgman, P.W., 74–80, 82, 87
- Buddhism, 242, 311, 313, 323
- Carnap, Rudolf, 94, 106
- categories, 40, 57, 59–60, 122–3, 201
See also concept
- causality, 9, 51, 177, 188–9, 201, 202, 204, 205, 211, 213, 218–20, 315
See also cause, law
- cause, 201, 204, 208–9, 214–17, 218–19, 301, 303
See also causality, law
- Chisholm, Roderick, 126, 167, 171
- choice, 207, 287, 288, 290, 296, 298–9, 302
See also freedom, will
- Christianity, 235, 286, 311, 314–15, 319, 346, 349–50, 352, 354, 356–60
See also Bible, faith, God, religion
- cognition, 6, 18, 49, 102–3, 312–13, 321, 378–9,
 Kant's conception of, 16, 30, 40, 53–5, 57, 59–60, 64–5, 163, 380–1
See also experience, knowledge
- common sense, 48–9, 52, 54, 59, 60, 228, 372
- compatibilism, 197–9, 201–2, 206, 210, 211, 213, 215, 220
See also incompatibilism
- concept, 7, 22, 38–40, 57, 60, 94, 98–109, 117–18, 202, 375–7, 381
 categorial, 40, 57, 59–60, 122–3, 201
 content of, 38–9, 100, 103
 empirical, 57–8, 102, 104, 109, 201
 form of, 100, 106
 formation of, 100, 101, 102–4
 and intuition, 40, 99, 117, 163, 376–9
 conflict of faculties, 325, 344–5, 347, 356
 consequences, 234–41, 242, 250
 consequentialism, 227–30, 232–3, 238, 240, 242
See also utilitarianism
- content, 38–9, 100, 103, 106, 109, 126
 non-conceptual, 36–7, 38, 39, 40
See also concept, form, matter
- consciousness, 4, 18, 19, 120, 125, 139, 322
See also apperception, awareness
- Critique of Pure Reason*, 25–6, 28, 117, 118, 173, 181–2, 321
- culture, 246, 318
- Davidson, Donald, 33, 37, 210, 215
- deduction, 369–70, 373–4, 380–1, 382–3
See also taste
- definition, 104–5, 107–8, 110, 111
 of truth, 47–8, 50
- Descartes, Rene, 8, 56, 164, 284, 286, 287, 290, 319, 391
- desire, 266–71, 274–5
See also happiness, pleasure
- Dewey, John, 63, 227, 229
- dualism, 29, 31, 312
- duty, 263, 278, 300, 301, 316, 322, 324, 382
See also imperative, law
- education, 295, 352, 255, 395
 Kant on, 333–4, 342, 344, 351, 360
See also university
- Enlightenment, 286, 313, 336, 388–90, 392, 393, 396, 397, 405,

- 411, 412, 417
 epistemology, 99, 111, 287
 See also cognition, knowledge, philosophy
 ethics, 228, 230–1, 234, 240, 241, 261, 273, 287–8, 293, 298,
 See also morality, philosophy
eudaimonia. *See* happiness
 eudaimonism, 263–4, 266, 272, 276, 277, 279
 See also happiness, pleasure
 Evans, G., 33–40
 event, 203–5, 206, 208, 212, 214–17
 See also cause, causality
 evil, 287, 290, 312, 326
 and good, 263, 287, 299, 301, 318, 322, 323, 334, 353
 existence, 79, 161
 of God, 318, 320, 321
 of external world, 288, 294, 320
 experience, 4, 13–14, 20, 37, 38–9, 48, 52, 78, 177, 188, 256, 286, 319, 320, 321, 324, 327
 cognitive, 10, 12, 21, 27, 34–5, 40, 76, 80–1
 objective, 9, 28, 32
 See also cognition, knowledge
 facts, 369
 and values, 86, 241–2
 faculties,
 conflict of, 325, 344–5, 347, 356
 harmony of, 371, 372, 376–9
 faith, 314, 319, 346, 349–50, 356
 See also Bible, Christianity, God, religion
 Fichte, J.G., 167, 290–1, 315
 form, 109, 164, 202, 203, 317, 371, 376
 of concept, 100, 106
 of intuition, 9, 184
 See also content, matter
 freedom, 17, 180–1, 199, 200–1, 210, 211, 212, 213, 214–17, 220, 312, 315, 318, 319, 323, 333, 341, 348
 Beck on, 197, 210–13, 217
 Kant on, 4, 10, 197–9, 204–6, 209, 210–13, 217–21, 270, 273, 296, 298–9, 322, 351, 360, 379
 transcendental, 8–9, 199, 204–5
 See also choice, reason, will
 Frege, Gottlob, 5, 93–4, 96, 97, 98
 Freud, Sigmund, 291, 317
 God, 185–7, 284, 287–8, 289, 291, 292–4, 306, 312, 313, 315, 318–22, 324, 327, 329, 348, 352, 353, 357, 358–9
 See also Bible, Christianity, faith, religion
 good, 241, 261, 262–4, 266, 288, 290, 292–3, 307, 311, 325, 329, 348, 370
 and evil, 263, 287, 299, 301, 318, 322, 323, 334, 353
 highest, 263, 289, 305–7
 See also happiness, pleasure
 Guyer, Paul, 168, 383
 Hamman, J.G., 405, 413, 414, 415, 416
 Hampshire, Stuart, 228–9, 231, 232
 happiness, 261–4, 266–79, 299–300, 305, 306, 318, 324, 408
 See also eudaimonism, good, pleasure
 harmony, 271, 276, 305, 319, 381, 409
 of faculties, 371, 372, 376–9
 pre-established, 181, 185, 207, 214, 305
 Hegel, G.W.F., 169, 284, 290–1, 307, 314, 315, 316, 320

- Heidegger, Martin, 14, 291
 Henrich, Dieter, 369, 388
 Herder, J.G., 397, 405, 406, 412, 414, 416, 417
 heteronomy, 301, 329
 and autonomy, 198–9, 200, 298–9
 highest good. *See* good
 history, 290, 306
 of the will, 283, 286, 307
 Hudson, Hud, 207–8, 211, 220
 human nature. *See* beings, man, nature
 Hume, David, 400, 410–11, 416
 Husserl, Edmund, 5, 17

 idealism, 76, 78, 167–8, 172, 175, 183, 189
 transcendental, 64, 118–19, 133, 167, 168, 171, 172, 178, 182
 See also metaphysics, realism
 identity, 76, 77, 79–80, 81, 95, 121, 128–9, 162, 206, 396
 token-token, 215, 217–18, 219–20
 illusion, 55, 312
 imagination, 5, 11, 21, 269, 275, 276, 326, 327, 371
 and understanding, 376–9
 imperative, 251, 275, 316, 324, 355
 categorical, 85, 198–200, 245, 247, 297–8, 316, 328
 See also duty, law, normativity, reason
 inclination, 249, 266–70, 272, 276, 299–300, 302, 390–1
 incompatibilism, 205, 210, 213
 See also compatibilism
 indexicals, 135–7
 intellect, 62, 209, 287, 289, 293, 294, 305, 306, 313
 See also reason, understanding
 interaction,
 of substances, 201, 203, 204
 and truth, 54, 65
 interest, 250, 262, 296, 303, 327, 342, 370
 of reason, 295–6, 309
 self-, 264–5, 322, 324
 intuition, 17, 99, 117–18, 120, 121, 124–5, 138, 182, 323, 326, 373
 and concepts, 40, 99, 117, 163, 376–9
 empirical, 201, 202
 forms of, 9, 184
 irrationality, 274, 277, 287, 290, 313
 See also rationality

 James, William, 63–4, 227, 242
 Jews, Kant's attitude toward, 347, 350, 352, 356–9
 judgment, 55–60, 62, 65, 73, 74, 82, 86, 96, 98–9, 101, 215, 318, 370–80
 analytic, 93–4, 95–101, 105–10
 cognitive, 201, 369, 371–2, 375–6, 382
 determinative, 47, 56–7, 65, 201–3, 215–16, 316
 of experience, 11–12
 faculty of, 376–7, 381
 moral, 83, 85, 87, 241, 316, 326
 of perception, 11–12
 synthetic, 93–4, 96–8, 100
 of taste, 369–71, 374–80, 382–3

 Kant, Immanuel, xvii, xxiii–xxv, 7–11, 18, 19, 30, 32, 35–6, 39–40, 47–8, 53–4, 59–60, 71–2, 75–6, 79–80, 81–3, 86–8, 94–111, 117–39, 167–89, 206–7, 241, 249–54, 261–3, 311–29, 369–83, 387–417
 on analyticity, 94–111
 on categorical imperative, 85, 198–200, 245, 247, 297–8, 316, 328

- on cognition, 16, 30, 40, 53–5, 57, 59–60, 64–5, 163, 380–1
- his economic circumstances, 401–4
- on determinative judgments, 47, 56–7, 65, 201–3, 215–16, 316
- on education, 333–4, 342, 344, 351, 360
- on freedom, 4, 10, 197–9, 204–6, 209, 210–13, 217–21, 270, 273, 296, 298–9, 322, 351, 360, 379
- as gallant *Magister*, 400–1, 414–15
- on happiness, 261–4, 261–4, 266–79, 299–300, 305, 306, 318, 324, 408
- on Jews, 347, 350, 352, 356–9
- on knowledge, 25–6, 28, 35, 78–9, 117–19, 168, 173–4, 176, 178, 320–1
- on maxims, 245–7, 249, 252–7, 269, 300
- on metaphysics, 167–73, 174–8, 179–89, 205–6, 208–9, 219, 411
- on moral law, 84, 187, 245–6, 272–3, 278, 292, 296–7, 306, 325, 326, 327–9, 333, 357
- on philosophy, 6, 13, 32, 72, 173, 322, 339–40, 343, 345, 399
- on practical reason, 198–9, 200–1, 210, 212, 245, 290, 294, 313, 314, 315–16, 321, 324, 370
- on self-awareness, 117–26, 128–39
- on taste, 369–78, 381–3
- on transcendental idealism, 64, 118–19, 167–89
- on truth, 47–8, 53–4, 55, 59–60, 61, 64–5
- on *Wille* and *Willkür*, 291, 294, 295, 296, 297, 299, 300–4
- on university, 334–7, 339, 345, 352, 356, 359–60
- knowledge, 11, 16, 33–4, 39, 78, 178, 180, 183, 210, 211, 220, 256, 319–21, 327, 340, 343, 374, 391
- Kant's conception of, 25–6, 28, 35, 78–9, 117–19, 168, 173–4, 176, 178, 320–1
- limits of, 178, 181, 188, 319
- self-, 123, 126, 183, 184, 301, 320
- See also* cognition, experience
- Koestler, Arthur, 229, 231, 232, 234, 240
- language, 39, 93, 94, 111, 123, 248
- Langton, Rea, 171–81, 183, 185–9
- law, 84–5, 245–6, 247, 250, 252, 274, 275, 295, 296, 382
- causal, 200, 201, 203–5, 209, 216, 218
- moral, 84, 187, 245–6, 272–3, 278, 292, 296–7, 306, 325, 326, 327–9, 333, 357
- natural, 186, 200, 245
- See also* cause, causality, imperative, normativity, reason
- Leibniz, G.W., 175, 176–7, 185, 219, 288–9, 304
- Lessing, G.E., 393, 395, 405
- man, 36, 39, 317–19, 340, 348, 353, 354, 357, 408, 411
- as creator, xvii, xxiii–xxv
- as an end in itself, 317, 319
- as a rational animal, 14, 16, 18
- See also* beings, nature
- Marx, Karl, 290, 291, 298
- mathematics, 13, 72, 73–4, 156
- matter, 176, 316, 380
- See also* content, form,
- maxims, 245, 246–8, 249–57
- Kant on, 245–7, 249, 252–7, 269, 300
- See also* law, principle, rule
- McDowell, John, 33–41
- meaning, 59–60, 110–11, 317
- Mendelssohn, M., 171, 388, 395, 396, 397, 417

- metaphysics, 391
 - Kant's conception of, 167–73, 174–8, 179–89, 205–6, 208–9, 219, 411
 - See also* cognition, knowledge, object, philosophy, thing, world
- mind, 20, 241
- modernity, 286, 288, 291
 - See also* philosophy
- monism, anomalous, 215–18, 219–20
- Moore, G.E., 233, 236, 238, 239
- morality, 84, 227, 228, 230, 232, 235, 239, 272, 276, 278–9, 286, 291–2, 295, 297, 300, 305–6, 312, 323, 348, 353, 410
 - See also* ethics, philosophy
- motion, 159–60, 162, 164
- mysticism, 311, 314, 322, 323, 328, 352, 353, 355
- nature, 77, 186, 200, 201, 207, 209, 264, 273–5, 306, 316–17, 324, 408, 409
 - human, 261, 264, 266, 271–3, 274, 275, 323, 348, 353, 408, 411
 - See also* reality, world
- necessity, 4, 9, 197, 202, 215, 216, 288, 374
- needs, 276, 317
- Newton, Isaac, 75, 321, 409
- Nietzsche, Friedrich, 291, 315
- normativity, 59, 292, 294, 369–71, 376–8, 380–3
 - See also* imperative, law, reason
- noumena, 208–9, 244
 - See* object, phenomena, thing, world
- object, 4, 12, 16, 29–31, 73, 76, 77–8, 84, 85, 87–8, 101, 104, 117–18, 120–1, 156–7, 186, 209, 212, 324, 371, 372, 376–8, 380–1
 - of cognition, 11, 51, 60, 83, 120, 121, 122–4, 163, 305, 320, 379
 - of experience, 4, 12
 - and subject, 30, 61–3, 321, 371
 - of thinking, 7, 13, 14, 15, 17, 20–1
 - transcendental, 29–30
 - See also* appearance, noumena, phenomena, representation, thing, world
- objectivity, 31, 58–9, 73–4, 82, 84, 87
- ontology, 206–7
 - See also* metaphysics, philosophy
- organism, 316–17
- perception, 7, 16, 20, 22, 37–9
 - judgment of, 11–12
- phenomena, 176, 177, 178–80, 207, 209, 294
 - See also* appearance, noumena, object, representation, thing, world
- philosophy, 48, 87, 94, 286, 319, 320, 339, 360, 389–92, 393–8, 411
 - Kant's conception of, 6, 13, 32, 72, 173, 322, 339–40, 343, 345, 399
 - modern, 13, 28, 33, 283, 284, 286
 - moral, 84, 245–7
 - popular, 389, 392, 395, 399, 411, 417
 - Schulphilosophie*, 389, 391–4, 397–9, 417
 - and theology, 335, 337–8, 343, 347
 - See also* ethics, metaphysics, morality, ontology
- Plato, xxiv, 18, 64, 232, 241, 245, 261, 264, 301, 325, 329
- play, free, 372–3, 378–9, 382
- pleasure, 261–2, 264, 265, 266–9, 270–2, 274–5, 317, 376
 - See also* good, happiness, inclination
- pragmatism, 61, 63–4, 291

- predicate, 94, 95, 106, 109–111, 175
 principle, 210, 231–2, 233
 constitutive and regulative, 210,
 211, 220, 372, 382
 of sufficient reason, 197, 201
 of taste, 372, 375
 See also law, rule
 prudence, 245, 253, 279
 psychology, 31–2, 36, 133, 202–3, 204
 punishment, 239, 273
 purpose, 273, 274, 318, 409
 purposiveness, 202, 209, 372, 373,
 376, 380

 Quine, W.V., 94, 106–7

 rationality, 5, 10, 12, 198, 209, 292–
 3, 294–5, 297, 298, 299, 304,
 312–14
 See also intelligence, irrationality,
 reason, spontaneity
 Rawls, John, 81–4, 87–8, 167, 254
 realism, 33, 75, 79, 84, 167, 172
 empirical, 29–30, 53, 78, 81, 87
 transcendental, 29–30, 53, 78
 See also idealism, metaphysics
 reality, 39, 50–1, 53, 64, 312–13, 317
 noumenal, 35, 40
 See also nature, world
 reason, 201, 249, 254, 264, 269, 271,
 275, 295–6, 301–2, 304, 311,
 312, 313, 314, 318–19, 323,
 327, 328, 341, 348, 349, 353,
 360, 369
 Kant's view on, 8, 27, 205–6, 340,
 341, 411
 practical, 198–9, 200–1, 210, 212,
 245, 290, 294, 313, 314, 315–
 16, 321, 324, 370
 principle of sufficient, 197, 201
 theoretical, 313, 315–16, 321
 See also intellect, understanding

 receptivity, 8, 173–4, 184
 See also spontaneity
 reductivism, 71–4, 76–7
 reference, 50–1, 58, 93, 126–31, 136,
 139
 Reinhold, K.L., 167, 168, 169, 170
 religion, 284, 285, 286, 293, 294,
 305, 312–13, 315, 316, 319,
 346, 349–51, 356–7, 359
 See also Bible, Christianity, faith,
 God
 representation, 4–6, 29–31, 73, 103,
 119–20, 121–3, 125, 129–32,
 134–6, 170, 174, 371, 376–8
 See also appearance, object,
 phenomena
 rule, 3–4, 6, 11, 13, 18, 19–20, 105,
 245, 247–50, 252, 380–1
 See also law, maxim, principle
 Rousseau, Jean Jacques, 271, 289,
 290, 298, 334, 339, 360, 400,
 404, 405–7, 409, 410, 416

 Schelling, F.W.J., 311, 315, 320
 Schopenhauer, Arthur, 291, 311
 science, 11, 48, 72, 75–81, 391, 408–
 9, 410, 411
 Second Analogy. *See* Analogy
 self, 119, 123, 138, 264, 296–7, 303–
 5, 314
 rational, 264, 298–9, 323
 self-awareness, 117–19, 121–39
 self-interest, 264–5, 322, 324
 self-knowledge, 123, 126, 183, 184,
 301, 320
 self-worth, 265–6, 273
 semantics, 13, 93–4, 106, 109
 sensation, 380, 381
 sense, 35, 38–41
 inner, 118, 123–4, 125, 134, 137–
 8, 184
 sensibility, 9, 163, 184, 323

- skepticism, 171, 265, 342, 358
 Socrates, 261, 262, 264–5, 276, 333
 soul, 208, 265, 311, 319, 354
 space, 374
 and time, 335, 337–8, 343, 347
 Spinoza, Baruch de, 177, 288, 290, 301
 spontaneity, 3–8, 10, 12–13, 14–16,
 18, 119, 198–9, 200, 215, 303,
 316
 See also choice, freedom,
 receptivity, thinking
 Strawson, P.F., 126, 128, 171, 173,
 174
 subject, 61, 76, 119, 121, 122–3, 125,
 139, 164, 286, 303
 and object, 30, 61–3, 321, 371
 subjectivity, 58–9
 sublime, 324, 325–6
 substances, 159, 175, 176–7, 180–1,
 182, 185–6
 interaction of, 201, 203, 204
summum bonum. *See* good
 synthesis, 162–3

 taste, 369–78, 381–3, 411
 teleology, 274, 289, 294
 theology, 335, 337–8, 343, 347
 See also Bible, God
 thing, 322
 in itself, 6, 9, 15, 171–4, 176, 177–
 81, 183, 185–8, 207, 208–9,
 213–14, 319
 See also noumena, object, world
 thinking, 5–7, 10, 11, 17, 18–20, 22,
 311
 spontaneous, 3–6, 10, 12–13, 14–
 16, 20
 See also spontaneity, thought
 Thomasius, Christian, 390, 391, 394,
 410
 thought, 94, 101, 117, 119, 120–6,
 129–32, 134–6, 139

 See also spontaneity, thinking
 time, 153–5, 157–64, 327
 and space, 335, 337–8, 343, 347
 truth, xxiii–xxiv, 12, 47, 49, 52–3,
 55–6, 58, 60–3, 73, 83–4, 88,
 105–6, 228, 288, 345
 as coherence, 54, 59, 61, 63
 commonsense, 47–50, 54, 58, 60,
 61, 65
 as correspondence, 50, 53–4, 61, 63
 definition of, 47–8, 50
 Kant's conception of, 47–8, 53–4,
 55, 59–60, 61, 64–5
 pragmatist conception of, 61, 63–4

 understanding, 35, 38–41, 99, 163, 320
 and imagination, 376–9
 See also intellect, reason
 university, 390–1, 393–4, 398
 Kant's idea of, 334–7, 339, 345,
 352, 356, 359–60
 utilitarianism, 228–30, 233, 234, 277,
 285, 292, 322
 See also consequentialism

 values, 83–4, 86, 88, 261
 and facts, 86, 241–2
 virtue, 277, 333

 will, 10–11, 199, 204–5, 208, 217,
 218–20, 245–6, 276, 283–4,
 286, 287–90, 292–303, 304–7,
 316, 321, 374
 history of, 283, 286, 307
 rational, 208, 291, 292, 297, 300
 as *Wille* and *Willkür*, 291, 294, 295,
 296, 297, 299, 300–4
 See also choice, freedom, reason
 Wittgenstein, Ludwig, 18, 26, 247–9
 Wolff, Christian, 390–4
 Wood, Allen, 207–8, 209, 211, 214,
 215, 217, 218, 220

- world, 5, 6, 11, 37, 49, 52, 56, 63, 93,
206, 284, 285–6, 288–9, 290,
291, 305, 306, 320, 396
noumenal, 34, 181–2, 206, 294, 299
phenomenal, 206–7, 208–9, 212,
299
- See also* noumena, object,
phenomena, reality, thing
worth, 273
self-, 265–6, 273
Zeno, 153, 325